PHYSICAL ACTIVITY IN A MULTIETHNIC CANADIAN POPULATION

PHYSICAL ACTIVITY IN A MULTIETHNIC CANADIAN POPULATION

By

AMANDEV AULAKH, BHSC.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

© 2008, Amandev Aulakh

MASTER OF SCIENCE (2008)

McMaster University

Clinical Epidemiology and Biostatistics

Hamilton, ON

TITLE: Physical activity in a multiethnic Canadian population: The association of physical activity with cardiovascular risk factors and exploration of physical activity patterns

AUTHOR: Amandev Aulakh, BHSc. (McMaster University)

SUPERVISOR: Dr. Sonia Anand

NUMBER OF PAGES: xxii, 218

Abstract

Obesity has become a global epidemic in children and adults and a decline in physical activity (PA) has contributed to this phenomenon. Low PA has been identified as a risk factor for various diseases, including cardiovascular disease. Past literature has identified variations in PA by sex, ethnicity and socioeconomic status (SES), but the majority of studies have been restricted to American populations. This thesis explored trends in total, work related and leisure time PA in a multiethnic Canadian population.

Analyses were conducted using data from the SHARE, SHARE-AP and SHARE-AP Action studies. The reliability and validity of physical activity questionnaires (PAQs) were assessed using Pearson and Spearman's rank correlation. Risk factors by PA level were compared using analysis of variance (ANOVA) for continuous variables and chisquare tests for categorical variables. Fisher's exact test was used to determine whether risk factor clustering by PA level was more evident when using PAQs or the RT3 accelerometer (RT3), and Cohen's Kappa coefficient was used to determine the agreement between the PAQs and the RT3 in classifying participants. The Chi-square goodness of fit test was used to determine differences in PA trends and multiple binary logistic regression was performed to determine variable association with low total PA.

Reliability coefficients of all PAQs ranged from 0.06 to 0.80 while validity assessments ranged from -0.07 to -0.17 against systolic blood pressure for the modified ARIC/Baecke Questionnaire and -0.10 to -0.30 for the PAQs used in SHARE-AP Action. Greater risk factor clustering was seen in low PA groups compared to high PA in SHARE and SHARE-AP data. However, no such association was seen in SHARE-AP Action and there was poor agreement in PA level classification between the PAQs and the RT3. Differences by sex, ethnicity and SES were apparent in different contexts of PA. As well, low PA was associated with South Asian and Chinese ethnicity, increasing age, low SES and Aboriginals in the low SES category. These findings show that differences in PA exist between different groups. Identifying populations prone to inactivity can assist in the development of health promotion strategies that target individuals susceptible to low PA.

ACKNOWLEDGEMENTS

I realize that this may resemble a verbose Oscar acceptance speech more than it does a conventional acknowledgements section. However, I am indebted to many people who supported me in the journey to completing my Masters degree. I never expected to encounter the obstacles that I did in finishing my thesis. However, now that this chapter of my life has come to a close, I know that I would not have been able to reach this point without the support of numerous individuals.

I would like to thank my thesis committee members, Dr. Cameron Blimkie, Dr. Changchun Xie and Dr. Andrea Kriska, for their advice, expertise and encouragement throughout this process. To my supervisor, Dr. Sonia Anand, thank you for all of your support, time and encouragement over the past three years. I feel fortunate to have had found such a strong role model in a supervisor.

To the women and men of the CARING network, it has been an absolute pleasure to work with such compassionate, intelligent and generous individuals. I would also like to thank Kathy Stewart for all of her help throughout this process. Andrew, your statistical expertise has been invaluable, particularly in moments when I was ready to offer a sacrifice to the statistical gods. Larry, thank you for all of your help with deciphering data. To the Catherines (Kreatsoulas and McGorrian), you are such strong women and you have both taught me so many invaluable lessons. I will miss afternoon teas, impromptu statistical lessons and our discussions on virtually all topics (academic and

v

otherwise). Words are really quite inadequate at this point to fully express my gratitude to you both.

To my faithful group of friends from the Ridge, Vanathy, Laura, Serena, Katie and Melissa, may there be many more nights of nachos, laughs and rained out barbeques. Thank you for always being there to listen to my rants and to give me perspective. To the HRM women, your encouragement has meant so much to me. Andrea, Debika, Ruth and Sonya, you were always confident that I could accomplish this and that has helped me through this journey.

To Kristina, your infinite wisdom, insight and humour have been invaluable to me not only over the course of my Masters but for the past five years. Whenever I have doubted my abilities, you have been there to renew my faith in myself and to put everything into perspective. Marianne, my Broatian sister, we have seen each other through so much over the years and if there is one thing we have learned, there is nothing that cannot be achieved with a little "power, pressure and focus". Paul and Adrienne, getting through the past year would have been infinitely more difficult without you both and I am so thankful that our paths crossed again at Queen's. Tricia, I am sure that we must have stumbled upon solutions for world peace and the meaning of life during our many walks – thank you always for your willingness to listen. Jen and Stash, thank you for always making me laugh and for believing in me. I will miss walking down the hall to visit you.

Finally, I would like to thank my sisters, Suman, Gagan and Pawan, for their continued love and support. To my parents, you have taught me the importance of family, love,

vi

hard work and staying committed to my goals. Your unwavering support of all my endeavours has meant the world to me. This thesis would not have been possible without you.

TABLE OF CONTENTS

Abstract	iii
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	viii
LIST OF TABLES	xi
LIST OF FIGURES	xvi
LIST OF ABBREVIATIONS	xviii
MSC THESIS SUMMARY	xxi
CHAPTER ONE: BACKGROUND AND LITERATURE REVIEW	1
1.1 – Physical activity	1
1.1.1 – Physical activity as a modifiable risk factor for cardiovascular disease,	2
diabetes mellitus and obesity	
1.2 – Methods of measuring physical activity	20
1.2.1 – Physical activity questionnaires	21
1.2.2 – Properties of physical activity questionnaires	32
1.2.2.1 – Reliability	32
1.2.2.2 – Validity	33
1.2.3 - Criterion measures of physical activity	40
1.2.3.1 – Accelerometers	46
1.2.3.2 – RT3 triaxial accelerometer	50
1.3 - Physical activity measurement in non-white ethnic groups	56
1.3.1 – Literature review	58
1.4 - Sex, ethnicity and socioeconomic status and physical activity	64
1.4.1 – Physical activity patterns	66
1.4.2 – Literature review	67
Chapter Two: Study Design	75
2.1 – SHARE and SHARE-AP overview	75
2.1.1 – Recruitment and study population	75

2.1.2 - Clinic visit and measurements	78
2.1.3 – Physical activity measurement	81
2.2 – SHARE-AP Action overview	87
2.2.1 – Recruitment and study population	87
2.2.2 – Clinic visits and measurements	91
2.2.3 – Physical activity measurements	94
2.3 – Threats to validity	98
2.4 – Limitations of cross sectional studies	99
2.5 – Selection bias	99
2.6 – Social desirability bias	100
2.7 – Reverse causality bias	101
2.8 – Self-reporting of behaviours in epidemiologic and intervention studies	102
Chapter Three: Methods and Analysis	103
3.1 – Assessment of reliability and validity in SHARE, SHARE-AP and	103
SHARE-AP Action	
3.1.1 – SHARE/SHARE-AP	103
3.1.2 – SHARE-AP Action	126
3.2 – Risk factor clustering in SHARE and SHARE-AP	137
3.2.1 – Methods	137
3.2.2 – Analysis	140
3.3 – Risk factor clustering and physical activity categorization in SHARE-AP	161
Action	
3.3.1 – Methods	161
3.3.2 – Analysis	164
3.4 – Physical activity trends in SHARE and SHARE-AP	167
3.4.1 – Methods	167
3.4.2 – Results	169
CHAPTER FOUR: DISCUSSION AND CONCLUSION	181
4.1 – Validity and reliability of physical activity assessment tools	181

181
183
185
189
192
194
196
200
201
203
204
205

LIST OF TABLES

Table 1.1: Studies supporting the association between physical activity and	11
cardiovascular disease	
Table 1.2: Studies supporting the association between physical activity and	14
cardiovascular disease risk factors	
Table 1.3: Studies supporting the association between physical activity and	16
diabetes mellitus	
Table 1.4: Studies supporting the association between physical activity and	18
obesity	
Table 1.5: Comparison of physical activity questionnaires and objective	26
measures of physical activity	
Table 1.6: A selection of physical activity questionnaires that have been used	27
in epidemiologic research	
Table 1.7: Comparison of objective measures of physical activity	45
Table 1.8: A list of fitness components and the methods by which they are	46
commonly evaluated	
Table 1.9: Validation of the RT3 triaxial accelerometer	54
Table 1.10: Studies exploring ethnic issues in physical activity measurement	62
Table 1.11: Studies exploring variations in physical activity by ethnicity,	72
socioeconomic status and sex	
Table 2.1: Data collected in SHARE and SHARE-AP	79
Table 2.2: Physical measurements taken during the clinic visit	80
Table 2.3: Time and proportion values for question 36 in the modified	83
ARIC/Baecke Questionnaire	
Table 2.4: Calculation of the simple sports score for the modified	84
ARIC/Baecke Questionnaire	
Table 2.5: Scoring of the modified ARIC/Baecke Questionnaire used in	85
SHARE and SHARE-AP	
Table 2.6: Visit map for SHARE-AP Action	92

Table 2.7: Physical measurements taken in SHARE-AP Action	93
Table 2.8: Outcomes for SHARE-AP Action	94
Table 3.1: Assessment for normality of data using the Shapiro-Wilks test	105
Table 3.2: Descriptive statistics of the variables included in the analysis	106
Table 3.3: Test-retest reliability of baseline and 1-month administrations of	106
the ARIC/Baecke Questionnaire	
Table 3.4: Descriptive statistics of the variables included in the criterion	112
validity analysis	
Table 3.5: Assessment of the normality of the variables included in the	112
criterion validity analysis	
Table 3.6: Criterion validity of the modified ARIC/Baecke Questionnaire	113
using systolic and diastolic pressure as criterion measures	
Table 3.7: Sensitivity analysis comparing criterion validity of the work related	123
physical activity index	
Table 3.8: Assessment of criterion validity in three different populations using	124
systolic blood pressure as a criterion	
Table 3.9: Assessment of criterion validity in three different populations using	125
diastolic blood pressure as a criterion	
Table 3.10: Assessment of normality for data included in test-retest reliability	127
analyses	
Table 3.11: Descriptive statistics of variables included in test-retest reliability	127
analyses	
Table 3.12: Test-retest reliability of baseline and 3-month administrations of	128
the physical activity questionnaires	
Table 3.13: Test-retest reliability of baseline and 3-month administrations by	128
control and intervention groups	
Table 3.14: Normality tests of the data included in criterion validity analyses	132
Table 3.15: Descriptive statistics of variables included in criterion validity	133
analyses	

xii

Table 3.16: Criterion validity of the physical activity questionnaires using the	133
RT3 as a criterion measure	
Table 3.17: Definition of risk factors	139
Table 3.18: Missing data	140
Table 3.19: Baseline characteristics including diabetic participants	142
Table 3.20: Baseline characteristics excluding diabetic participants	142
Table 3.21: Mean values of continuous measures including diabetic	143
participants	
Table 3.22: Mean values of continuous measures excluding diabetic	144
participants	
Table 3.23: Crude prevalence (%) of continuous risk factors by physical	145
activity level	
Table 3.24: Crude prevalence (%) of continuous risk factors by physical	146
activity level	
Table 3.25: Crude and adjusted prevalence of participants with 3 or more risk	146
factors (%)	
Table 3.26: Age, sex and ethnicity adjusted means for systolic blood pressure	151
(in mmHg)	
Table 3.27: Age, sex and ethnicity adjusted means for diastolic blood pressure	152
(in mmHg)	
Table 3.28: Age, sex and ethnicity adjusted means for waist to hip ratio	153
Table 3.29: Age, sex and ethnicity adjusted means for body mass index (in	154
kg/m ²)	
Table 3.30: Age, sex and ethnicity adjusted means for fasting glucose	155
including diabetic participants (in mmol/L)	
Table 3.31: Age, sex and ethnicity adjusted means for fasting glucose	156
excluding diabetic participants (in mmol/L)	
Table 3.32: Age, sex and ethnicity adjusted means for HbA1c including	157
diabetic participants (in %)	

Table 3.33: A	Age, sex and ethnicity adjusted means for HbA1c excluding	158
diabetic partie	cipants (in %)	
Table 3.34: A	Age, sex and ethnicity adjusted means for ApoB/A1	159
Table 3.35: 0	Odds ratio of smoking adjusted by age, sex and ethnicity	160
Table 3.36: H	Risk factor definitions	162
Table 3.37: 0	Comparisons of mean waist to hip ratios in children and	163
adolescents		
Table 3.38: I	Definitions for low, moderate and high physical activity as	164
assessed by p	hysical activity questionnaires (in MET-minutes/week)	
Table 3.39: I	Definitions for low, moderate and high physical activity as	164
assessed by th	he RT3 accelerometer at baseline (in MET-min/day)	
Table 3.40: I	Descriptive statistics	165
Table 3.41: H	Fisher's Exact Test values	166
Table 3.42: I	Definitions of low, moderate and high physical activity categories	168
by context of	physical activity	
Table 3.43: I	Prevalence of total physical activity by social disadvantage index	170
score		
Table 3.44: 1	Prevalence of leisure time physical activity by social disadvantage	170
index score		
Table 3.45: 1	Prevalence of work related physical activity by social	171
disadvantage	index score (excluding people indicating "I do not work" on	
question 34)		
Table 3.46: 1	Prevalence of total physical activity by ethnicity	172
Table 3.47: 1	Prevalence of leisure time physical activity by ethnicity	173
Table 3.48: 1	Prevalence of work related physical activity by ethnicity	173
(excluding pe	eople indicating "I do not work" on question 34)	
Table 3.49: 1	Prevalence of total physical activity by sex	174
Table 3.50: 1	Prevalence of leisure time physical activity by sex	175
Table 3.51: 1	Prevalence of work related physical activity by sex (excluding	175

people indicating "I do not work" on question 34)	
Table 3.52: Univariate binary regression results	176
Table 3.53: Model assessing main effects of age, sex, ethnicity and social	178
disadvantage index score as predictors of low physical activity	
Table 3.54: Assessment of model fit using likelihood ratio test	179
Table 3.55: Summary of the final logistic regression model	180

LIST OF FIGURES

Figure 1.1: A spring mass system	49
Figure 1.2: Triaxial planes	50
Figure 3.1: Scatterplot of baseline and 1-month administrations of total	107
physical activity index scores	
Figure 3.2: Scatterplot of baseline and 1-month administrations of work	108
related physical activity index scores	
Figure 3.3: Scatterplot of baseline and 1-month administrations of leisure time	109
physical activity index scores	
Figure 3.4: Scatterplot of baseline and 1-month administrations of sports	110
related physical activity index scores	
Figure 3.5: Criterion validity of total physical activity using systolic blood	114
pressure as a criterion measure	
Figure 3.6: Criterion validity of total physical activity using diastolic blood	115
pressure as a criterion measure	
Figure 3.7: Criterion validity of leisure time physical activity using systolic	116
blood pressure as a criterion measure	
Figure 3.8: Criterion validity of leisure time physical activity using diastolic	117
blood pressure as a criterion measure	
Figure 3.9: Criterion validity of work related physical activity using systolic	118
blood pressure as a criterion measure	
Figure 3.10: Criterion validity of work related physical activity using diastolic	119
blood pressure as a criterion measure	
Figure 3.11: Criterion validity of sports related physical activity using systolic	120
blood pressure as a criterion measure	
Figure 3.12: Criterion validity of sports related physical activity using	121
diastolic blood pressure as a criterion measure	
Figure 3.13: Summary of responses to question 34 on the modified	122
ARIC/Baecke Questionnaire	

Figure 3.14: Test-retest reliability of the IPAQ in adults using baseline and 3-	129
month administrations	
Figure 3.15: Test-retest reliability of the modified IPAQ in adolescents using	130
baseline and 3-month administrations	
Figure 3.16: Test-retest reliability of the 24-hour physical activity recall	131
questionnaire in children using baseline and 3-month administrations	
Figure 3.17: Criterion validity of the 24-hour physical activity recall	134
questionnaire in children using the RT3 accelerometer as a criterion measure	
Figure 3.18: Criterion validity of the modified IPAQ in adolescents using the	135
RT3 accelerometer as a criterion measure	
Figure 3.19: Criterion validity of the IPAQ in adults using the RT3	136
accelerometer as a criterion measure	
Figure 3.20: Distribution of risk factor scores for participants with RT3	166
accelerometer and physical activity questionnaire data	

LIST OF ABBREVIATIONS

AMI	Acute Myocardial Infarction
АроВ	Apolipoprotein B
ApoB/A1	Apolipoprotein B/A1 Ratio
ARIC/Baecke	Atherosclerosis Risk in Communities/Baecke Index
BF	Body Fat
BMI	Body Mass Index
BP	Blood Pressure
CCHS	Canadian Community Health Survey
CHD	Coronary Heart Disease
CPAF	Children's Physical Activity Form
CVD	Cardiovascular Disease
DBP	Diastolic Blood Pressure
DLW	Doubly Labelled Water
DM	Diabetes Mellitus
EE	Energy Expenditure
FM	Fat Mass
GQ	General Questionnaire
Hb	Hemoglobin
HbA1c	Hemoglobin A1c
НС	Hip Circumference
HDL-C	High Density Lipoprotein Cholesterol
HR	Heart Rate
IGT	Impaired Glucose Tolerance

IPAQ	International Physical Activity Questionnaire
KAB	Knowledge, Attitude and Behaviour
kcal	Kilocalories
LDL	Low Density Lipoprotein
Lp(a)	Lipoprotein (a)
LTPA	Leisure Time Physical Activity
LV	Left Ventricular
MAQ	Modifiable Activity Questionnaire
MET	Metabolic Equivalents of Task
MI	Myocardial Infarction
NHANES	National Health and Nutrition Examination Survey
OR	Odds Ratio
PA	Physical Activity
PAI-1	Plasminogen Activator Inhibitor 1
PAQ	Physical Activity Questionnaire
RCT	Randomized Controlled Trial
RR	Relative Risk
RT3	RT3 Triaxial Accelerometer
SBP	Systolic Blood Pressure
SDI	Social Disadvantage Index
SES	Socioeconomic Status
SHARE	Study of Health Assessment and Risk in Ethnic Groups
SHARE-AP	Study of Health Assessment and Risk Evaluation in Aboriginal Peoples
TC	Total Cholesterol
TEE	Total Energy Expenditure

TPA	Total Physical Activity			
TriTrac	TriTrac R3D Triaxial Accelerometer			
VO ₂ max	Maximal Oxygen Uptake			
WC	Waist Circumference			
WHR	Waist to Hip Ratio			
WRPA	Work Related Physical Activity			

MSC THESIS SUMMARY

TITLE

Physical Activity in a Multiethnic Canadian Population

OBJECTIVES

- To assess the validity and reliability of physical activity questionnaires (PAQs) used in adults in the SHARE, SHARE-AP and SHARE-AP Action studies and in children in the SHARE-AP Action study.
- To determine whether associations between clustering of risk factors and different levels of physical activity (PA) are more apparent when assessing PA using PAQs compared to criterion methods of assessing PA.
- To determine if differences in physical activity patterns exist by sex, ethnicity (European, South Asian, Chinese and Aboriginal) and socioeconomic status (SES).

Hypotheses

- It is hypothesized that the PAQs will demonstrate good validity and reliability because they are based upon existing questionnaires. The questionnaires utilized in SHARE and SHARE-AP are modified versions of the ARIC/Baecke Index that were piloted before use. In SHARE-AP Action, adult activity was assessed using the International Physical Activity Questionnaire (IPAQ) while physical activity assessment was done in children using a modified IPAQ and a 24 hour physical activity questionnaire recall.
- 2. It is hypothesized that a stronger relationship between risk factor clustering and physical activity will be apparent when using criterion methods rather than PAQs.

The rationale for this hypothesis is that overreporting on PAQs can misclassify individuals, which may indicate more risk factor clustering at higher physical activity levels than actually exists (when one would expect to see less clustering). This is typically not a concern in criterion methods.

3. It is hypothesized that higher SES will be associated with increased leisure activity while lower SES will be associated with increased work and sedentary activity across sex and different ethnic groups. This hypothesis is based on prior studies completed in children and adolescents that have found differences in patterns between males and females, by different ethnic groups and by socioeconomic status.

TYPE OF THESIS

Analysis and interpretation of data from three epidemiologic population studies

SUMMARY

In this thesis, I will provide an overview of PA and describe the associations between PA and cardiovascular disease (CVD), CVD risk factors, diabetes mellitus and obesity that have been documented in the literature. I will then describe the various techniques developed to measure PA, with an emphasis on PAQs and accelerometers. I will also include a review of the issues associated with measuring PA in non-white ethnic groups. The literature review will conclude with a discussion of the trends by sex, ethnicity and SES in PA. I will describe the study designs of SHARE, SHARE-AP and SHARE-AP Action. I will then assess the validity and reliability of the PAQs used in the various studies. I will also perform analyses to determine whether clustering of risk factors is more apparent when measuring PA with a PAQ or a criterion measure. I will also investigate differences in total PA, work related PA and leisure time PA by sex, ethnicity and SES. As well, I will also use logistic modelling to determine which variables are associated with low PA.

CHAPTER ONE: BACKGROUND AND LITERATURE REVIEW

This chapter defines physical activity and describes the various methods of assessing physical activity, with an emphasis on physical activity questionnaires and accelerometers. As well, issues in assessing physical activity in non-white ethnic groups and trends in physical activity by sex, ethnicity and socioeconomic status are discussed.

1.1 - Physical activity

Accurate measurement of physical activity (PA) is important in epidemiologic research to ensure that valid conclusions are made. Central to this is a concise definition of PA including an understanding of how it differs from similar concepts, such as exercise and physical fitness. A commonly accepted definition of PA was established by Caspersen et al. (1) which includes four defining components: 1) movement of the body by the skeletal muscles, 2) results in energy expenditure (as measured in kilocalories or kilojoules), 3) energy expenditure that varies continuously and 4) correlates with measures of physical fitness. Exercise shares all of these attributes, but the distinguishing factor is that exercise is activity that is performed with the intent of increasing or maintaining physical fitness. Consequently, it can be considered a subcategory of PA (1). Physical fitness refers to those characteristics that are developed and maintained through PA and exercise. These are broadly divided into health-related and skill- related fitness. Aerobic endurance, muscle strength and endurance, flexibility and body composition are encompassed in health-related fitness, while skill-related fitness includes attributes such as speed, agility, power, coordination, reaction time and

balance. Physical fitness characteristics are often employed as a surrogate measure for PA (1).

PA is one of the three components comprising total energy expenditure (TEE). The basal metabolic rate, which is the caloric expenditure associated with maintaining basic function in a resting person, accounts for 50-70% of TEE (2). The thermic effect of food, which is the increase in heat production attributable to food consumption up to 8 hours after eating (3), accounts for 7-10% of TEE. PA accounts for the remaining energy expenditure (EE) and is the component that varies most between people. Consequently, it has been the focus of many PA measures (2).

1.1.1 – Physical activity as a modifiable risk factor for cardiovascular disease,diabetes mellitus and obesity

There are many theoretical reasons why increased PA may be associated with lower cardiovascular disease (CVD) burden. PA is believed to directly affect the circulatory system and induce metabolic changes that contribute to the development of CVD risk factors. Aerobic activity impacts the sensitivity of adrenalin and insulin, which assist in the metabolism of carbohydrates and fats by oxidative enzymes. This can change fat distribution, blood pressure (BP), cholesterol and triglycerides, all factors that can contribute to CVD (4). While PA is associated to CVD via various risk factors, PA is also believed to be associated with coronary heart disease (CHD) independent of these factors. A review by Kohl found that in 20 of 31 reviewed studies, an inverse doseresponse relationship exists between PA and CHD (5).

Physical Activity and Cardiovascular Disease

One of the first studies to demonstrate an association between PA and CVD was published in 1953 by Morris et al., which focused on work related physical activity (WRPA) and CHD (6). The study included 31 000 British males working at London Transport and 110 000 men working at the Post Office. Individuals in active jobs (e.g. bus conductors and postmen) were at lower risk of developing CHD compared to bus drivers and office employees, who were more sedentary in their work related activities. The standardized rate of CHD per 1000 persons was 2.7 in drivers compared to 1.9 in conductors and 2.4 in office workers compared to 1.8 for postmen (6). Since the study by Morris et al., the association between PA and clinical outcomes of CVD has been explored in cohort, case-control and randomized controlled trials (RCTs) (see Table 1.1).

An association between PA and CVD has been explored in numerous cohort studies. The Framingham Heart Study explored the association between PA risk factors and CVD in a cohort of men aged 45-64 years by taking into account their activity over a 24-hour period in addition to their WRPA (7). The period of follow-up was 24 years and PA was measured using an index. Increased PA levels were associated with reduced cardiovascular and coronary mortality, observed across all age categories. The mortality rate from CVD per 1000 was 226 for participants scoring greater than 34 on the PA index compared to 367 for those scoring less than 29 on the index (7). Subsequent studies, including the Honolulu Heart Program, Physicians' Health Study and the Harvard Alumni Health Study, found that increased levels of PA in men conferred lowered risk of developing CVD. The relative risk (RR) of developing CVD for the most active individuals in these studies ranged from 0.43-0.81 (8-10). Similar findings emerged from the Women's Health Study and the Nurses' Health Study which focused on female cohorts. The RR of CVD incidence for the most active individuals relative to the least active was 0.75 in the Women's Health Study while the Nurses' Health Study found a RR of 0.69 (11,12). Furthermore, a meta-analysis conducted by Berlin and Colditz (13) explored the association between PA and CHD in 26 cohort studies. The pooled RR for moderately active individuals compared to highly active individuals was 1.1 (95% CI: 1.0-1.3) for the incidence of CHD while the pooled RR for sedentary groups relative to highly active groups was 1.3 (95% CI: 1.1-1.5).

Case control studies have also been conducted and they have found that PA has a protective effect against CVD development. The Northern Manhattan Stroke Study explored the relationship between leisure time physical activity (LTPA) and the incidence of ischemic stroke in 1047 individuals. A dose-response relationship was evident as participation in heavy PA (e.g. aerobic dance, jogging) was more protective against stroke than light PA (e.g. walking, golf, gardening). The odds ratio (OR) of heavy PA was 0.23 compared to 0.39 for light PA (14). Rastogi et al. found a similarly protective effect for LTPA against the risk of CHD in a population of Indians (15). The RR of developing CHD was 0.45 in participants performing more than 145 MET-minutes of PA

4

per day compared to sedentary participants. In addition, sedentary activity increased the risk of CHD. The RR was 1.88 for participants performing more than 3.6 hours of sedentary activity per day relative to those performing less than 70 minutes daily (15).

Rothenbacher et al. also looked at LTPA but examined summer and winter PA separately in a German population (16). Similar trends were found for summer and winter PA with participation in more than 2 hours of PA per week conferring a RR of 0.39 and 0.27 compared to sedentary individuals, respectively. In contrast, work related physical strain actually increased the risk of CHD. Participants with the highest levels of physical strain had a RR of 4.86 compared to individuals in occupations with no physical strain (16). More recently, the INTER-HEART study identified PA as one of nine modifiable risk factors for acute myocardial infarction (AMI). Participation in strenuous PA had a protective effect against the risk of AMI, with an OR of 0.72 adjusted for age, sex and smoking (99% CI: 0.65-0.79) (17).

Physical Activity and Cardiovascular Disease Risk Factors

Although RCTs have been conducted to determine the association between PA and CVD risk factors, no RCTs could be found that looked specifically at CVD and PA (5). Several risk factors have been used as outcomes, including waist circumference (WC), hip circumference (HC), waist to hip ratio (WHR), % body fat (BF), BP and high density lipoprotein cholesterol (HDL-C) (see Table 1.2). The First Step Program aimed to change sedentary behaviour through group meetings and the use of pedometers to help motivate participants to engage in PA. In a study of sedentary individuals that were either overweight or obese, there were no significant differences in WC and HC between the control group and those receiving the First Step Program (18). In contrast, a study by Eriksson et al. found significant changes in WC and WHR in those undertaking a lifestyle intervention. Participants were enrolled in a program that aimed to improve diet and increase PA for a year and changes in specific outcomes were evaluated. The mean change between groups was significant for WC (-1.9 cm, P < 0.001), WHR (-0.01cm, P < 0.01) and diastolic blood pressure (DBP) (-2.3 mmHg, P < 0.05) (19). Other studies employing various lifestyle programs aiming to increase PA as interventions have found significant changes in systolic blood pressure (SBP), %BF, HDL-C and body composition (20-22).

Physical Activity and Diabetes Mellitus

While the association between PA and diabetes mellitus (DM) has been noted in numerous cohort studies (23), the relationship between PA and DM has been established most convincingly in RCTs (see Table 1.3). The Qa Ding Impaired Glucose Tolerance (IGT) and Diabetes Study evaluated the risk reduction of diet modification, exercise and a combination of diet and exercise treatment in a Chinese population with IGT. While diet reduced the risk of developing DM by 31%, a combination of diet and exercise had a risk reduction of 42% and exercise alone reduced the risk by 46% (24).

Similar findings were evident in the Finnish Diabetes Prevention Study. Individuals with IGT were randomized to either a lifestyle intervention or a control group. The intervention group received counselling aimed to improve their diet and exercise, and participants were followed for an average of 3.2 years. The lifestyle intervention conferred a 58% risk reduction in the development of DM (25).

The Diabetes Prevention Program also noted a risk reduction in participants in the lifestyle intervention group. Individuals with elevated fasting and post load plasma glucose were randomized to either a lifestyle intervention that aimed for a 7% reduction in weight and 150 minutes of PA weekly, taking 850 mg of metformin twice daily or a control group. While the lifestyle group had a 58% risk reduction after an average of 2.8 years of follow up, a 31 % reduction in the incidence was noted in the metformin group comparatively (26).

Physical Activity and Obesity

Body mass index (BMI) is typically used as a measure of obesity. A BMI between 25 and 29.9 kg/m² is considered overweight while a BMI of 30 kg/m² and above is considered obese (27). Studies examining PA and obesity prevention have failed to consistently demonstrate an association (see Table 1.4). In a review by Wareham et al., prospective cohorts show that those who participate in more PA are less likely to experience weight gain. However, this association may be attenuated by a number of factors, including directionality (28). While low levels of PA may contribute to obesity, obesity may also influence individuals to be less active (28). A study by Petersen et al. failed to show a conclusive relationship between physical inactivity and obesity, but found that inactive individuals were more likely to become obese later. In women, the OR of being obese was 0.81 for medium PA and 1.16 for high PA compared to inactive participants. Similarly in men, the OR for medium PA was 1.28 while it was 1.65 for high PA (29).

Relatively few studies have conducted RCTs exploring PA and obesity prevention. Donnelly et al. explored the impact of an exercise supervision program on weight, BMI and fat mass (FM) (30). Seventy-four moderately obese and overweight men and women were enrolled for 16 months. At the conclusion, it was found that while the program was successful in helping women maintain their current weight, men experienced weight loss. Men in the intervention group averaged 5.2 kg in weight loss, a reduction of 1.6 kg/m² in BMI and decreased FM by 4.9 kg. No significant changes were found in the current weight, BMI and FM of women in the intervention group while their control counterparts experienced increases in all three outcomes (30).

In contrast, a RCT conducted by Reilly et al. in a population of preschool children failed to show significant changes in BMI. Students were randomized to an enhanced PA program or control group and were assessed for changes at 6 and 12 months. No significant mean differences were found in the intervention group (31). These findings may be explained by the shorter period of intervention compared to Donnelly et al.'s study and the young age of participants. Further studies are needed to confirm the association between PA and obesity prevention.

While BMI is often used as a measure for obesity, recent studies have shown that WHR, or abdominal obesity, may be a better measure for CVD risk. Yusuf et al. explored the association between various measures of obesity, including BMI and WHR, and risk of myocardial infarction (MI) in 27 000 participants from the INTERHEART study (32). The OR for MI in participants in the highest quintile of BMI compared to the lowest quintile was 1.44 (95% CI: 1.32-1.57). Adjustment for WHR and other risk factors attenuated this association. In contrast, the OR for the highest WHR quintile relative to the lowest WHR quintile was 2.52 (95% CI: 2.31-2.74). Even after adjustment for BMI and other risk factors, the OR was 1.75 (95% CI: 1.57-1.95) (32). Price et al. also compared BMI and WHR as predictors of mortality in a population of older adults (aged \geq 75). There was no association between circulatory mortality and BMI in males (P = 0.667), but WHR was positively related to mortality (P = 0.001). In women, BMI was negatively associated with circulatory mortality (P = 0.004) and there was a positive relationship between WHR and mortality (P = 0.005) (33).

Reviews conducted by Ross and Janssen (34) and Kay and Fiatarone Singh (35) indicate there is limited evidence to suggest a strong relationship between WHR as a measure of abdominal obesity and PA. Ross and Janssen included RCTs and nonrandomized trials that explored the effects of PA on different obesity measures, including WHR. In the 7 RCTs included, there was no significant reduction in WHR associated with PA. Of the 6 nonrandomized trials included, only 2 found significant reductions in WHR in participants (34). Kay and Fiatarone Singh noted that an inverse association between PA and WHR has been established in cross sectional studies (35). However, similar to Ross and Janssen, WHR does not appear to be a good indicator of body composition change resulting from PA. In 10 RCTs that used WHR as a measure

of change, only one group experienced significant WHR changes relative to control groups. Although WHR is associated with PA, it is a poor measure of change (35).

Table 1.1: Studies supporting the association between physical activity and

		4.1
cardiovascu.	lar	disease

Authors,	Study	Intervention/	Population	Outcome	Ν
Year	Design	Exposure			
Kannel et al., 1986	Cohort	PA level	Males in the Framingham Heart Study aged 45-64 years	Improved cardiovascular and coronary morbidity and mortality with increasing PA across all age groups	1166
Donahue et al., 1988	Cohort	Tertiles of PA	Males in the Honolulu Heart Program aged 45-69	0.69 RR for active men aged 45-64 and 0.43 RR for active men (in highest PA tertile) aged 65-69 for incidence of CHD compared to inactive men (in lowest PA tertile)	7644
Sacco et al., 1998	Case Control	Leisure time PA (level and hours per week (h/wk))	Males and females in the Northern Manhattan Stroke Study	PA was protective against stroke across all ages, both sexes and ethnicity (adjusted OR = 0.37). A dose- response relationship was evident, with heavy PA being more protective	1047 (369 cases and 678 controls)

			2	than light PA (OR for heavy PA = 0.23 vs. OR for light PA = 0.39) and h/wk (OR for <2 h/wk = 0.42 vs. OR for ≥ 5 h/wk = 0.31)	
Sesso et al., 2000	Cohort	PA in kilojoules per week (kJ/wk)	Males in the Harvard Alumni Health Study aged 39-88	RR of CHD decreased as kJ/wk increased: 0.90 RR for 2100-4199 kJ/wk , 0.81 RR for 4200-8399 kJ/wk, 0.80 for 8400-12599 kJ/wk and 0.81 RR for \geq 12600 kJ/wk compared to <2100 kJ/wk	12 516
Lee et al., 2001	Cohort	Kilocalories per week (kcal/wk) expended on all activities	Females in the Women's Health Study aged 45 and older	RR for CHD decreased as EE increased: 0.79 RR for 200-599 kcal/wk, 0.55 RR for 600-1499 kcal/wk and 0.75 RR for \geq 1500 kcal/wk compared to <200 kcal/wk	39 372
Rockhill et al., 2001	Cohort	Hours of PA per week (h/wk)	Females in the Nurses' Health Study aged 30-55	RR of cardiovascular death declined with increasing PA: 0.80 RR for 1-1.9 h/wk, 0.74 RR for 2-3.9	80 348

				h/wk, 0.62 RR for 4-6.9 h/wk and 0.69 RR for \geq 7 h/wk compared to <1 h/wk.	
Rothenbacher et al., 2003	Case Control	Leisure time PA	German males and females aged 40-68	RR of CHD decreased with increasing summer leisure time PA compared to sedentary individuals: RR 0.85 for <1 h/wk, 0.60 for 1- 2 h/wk and 0.39 for >2 h/wk. Similar trends were noted for winter PA.	791 (312 cases and 479 controls)
Yusuf et al., 2004	Case Control	Leisure time PA	Males and females from 52 countries	Participation in strenuous activity had a protective effect against risk of AMI (OR adjusted for age, sex and smoking = 0.72)	29 972 (15 152 cases and 14 820 controls)
Rastogi et al., 2004	Case Control	Leisure time PA	Indian males and females aged 21-74	RR of CHD for participants in highest PA level (≥ 145 METs) was 0.45 compared to participants completing no leisure time PA	1050 (350 cases and 700 controls)
Table 1.2: Studies supporting the association between physical activity and

cardiovascular disease risk factors

Authors,	Study	Intervention/	Population	Outcome	Ν
Year	Design	Exposure			
Sacco et al., 1998	Case Control	Leisure time PA (level and hours per week (h/wk))	Males and females in the Northern Manhattan Stroke Study	PA was protective against stroke across all ages, both sexes and ethnicity (adjusted OR = 0.37). A dose- response relationship was evident, with heavy PA being more protective than light PA (OR for heavy PA = 0.23 vs. OR for light PA = 0.39) and h/wk (OR for <2 h/wk = 0.42 vs. OR for ≥ 5 h/wk = 0.31)	1047 (369 cases and 678 controls)
Eriksson et al., 2006	RCT	Lifestyle intervention	Individuals with moderate to high risk of CVD	Significant mean changes in mean WC, WHR and DBP in intervention group	151 (75 intervention and 76 control)
Araiza et al., 2006	RCT	Pedometer- based PA program	Individuals with DM	Increased HDL- C and resting EE and lowered PAI-1 activity in intervention	30 (15 intervention and 15 control)

				group	
Murphy et al., 2006	RCT	45 minutes of walking twice a week	Sedentary individuals	Reduction in SBP and maintenance of BF levels in intervention group	33 (21 intervention and 12 control)

Table 1.3: Studies supporting the association between physical activity and diabetes

mellitus

Authors,	Study	Intervention/	Population	Outcome	N
Year	Design	Exposure			
Manson et al., 1992	Cohort	Frequency of vigorous PA	Males in the Physicians' Health Study aged 40-84	Age adjusted RR of developing non- insulin dependent DM for men participating in PA at least once a week was 0.64	21 271
Pan et al., 1997	RCT	Diet, Exercise or Diet and Exercise	Individuals with IGT	Risk reduction for DM was 31% in the diet group, 46% in the exercise group and 42% in the diet and exercise group	530 (130 diet, 141 exercise, 126 diet and exercise and 133 control)
Tuomilehto et al., 2001	RCT	Lifestyle intervention program	e Individuals 58% risk tion with IGT reduction of D in lifestyle intervention group		522 (265 intervention and 257 control)
Diabetes Prevention Program Research Group, 2002	RCT	Lifestyle modification program or metformin	Individuals with elevated fasting and post-load plasma glucose	58% reduction in incidence of DM in lifestyle intervention group	3234 (1079 lifestyle, 1073 metformin and 1082 control)
Villegas et al., 2006	Cohort	Leisure time PA and daily living PA	Females in the Shanghai Women's	The risk of developing type 2 diabetes	70 658

dy declined with increasing PA: the RR for the highest level of leisure time PA was 0.83 and the RR for the highest level of daily living PA was 0.88 relative to the lowest levels.
lowest levels.

Authors,	Study	Intervention/	Population	Outcome	N
Year	Design	Exposure			
Donnelly et al., 2003	Donnelly et RCT S al., 2003 e p		Overweight and moderately obese individuals	Men in the intervention lowered weight (5.2 kg), BMI (- 1.6 kg/m ²) and FM (-4.9kg) while women maintained current weight, BMI and FM.	131 (87 intervention and 44 control)
Petersen et al., 2004	n et Cohort Leisure time Males 4 PA femal 20-78		Males and females aged 20-78	The OR of developing obesity increased with PA: in women, OR for medium PA was 0.81 and 1.16 for high PA (relative to inactive participants) and in men, OR for medium PA was 1.28 and 1.65 for high PA	6279
Tudor- Locke et al., 2004	RCT	First Step Program	Sedentary obese/ overweight individuals	Non-significant differences in WC and HC in intervention group	47 (24 intervention and 23 control)
Yin et al., 2005	RCT	After-school PA program	Grade three students	Increased attendance associated with decreased %BF and FM and	278

Table 1.4: Studies supporting the association between physical activity and obesity

.

				increased fat free mass	
Reilly et al., 2006	RCT	Enhanced PA program and health education	Scottish preschool children	No significant changes in BMI 6 or 12 months after randomization	545 (268 intervention and 277 control)

1.2. - Methods of measuring physical activity

The measurement of PA is challenging, as a number of parameters must be considered. These include the context, type, intensity, frequency and duration of the activity to be measured. Overall, PA can be divided into categories on the basis of when activities take place during the course of a day. Typical divisions of PA are by workplace and leisure time. LTPA encompass PA that is both structured (e.g. sports) and nonstructured (e.g. household duties) (1, 2). In addition, sedentary behaviour can be considered in measuring PA, although it is distinct from PA. Sedentary behaviour is low intensity and can include activities such as watching television, playing video games and motorized transport (36). PA can also be categorized by the type of activity performed. The intensity of PA varies depending on the type of activity performed, which can include aerobic, weight bearing, flexibility and strength activities (37).

A variety of techniques have been developed to measure PA. These can be broadly divided into two categories: subjective and objective measures. Subjective measures, such as physical activity questionnaires (PAQs), tend to focus solely on PA and are commonly used in epidemiologic research. Objective instruments, conversely, measure TEE and/or body motion and include motion sensors and doubly labelled water (DLW). These measures are often used as gold standards to validate PAQs. The success of these tools in accurately measuring PA is variable, largely due to the difficulty of measuring such a complex construct (38).

20

1.2.1 - Physical activity questionnaires

PAQs are commonly employed in population studies and may be selfadministered or administered by an interviewer. They typically require participants to recall activities that they participated in over a given period of time. PAQs can vary in length from a single item to several pages. Questionnaire items ask about details of the participant's activity such as the duration of activity (measured in minutes or hours), intensity (measured in metabolic equivalents of task (METs) or kilocalories (kcal)) and frequency (the number of times per week). This information is then used to obtain a continuous measure of the participant's PA (e.g. METs) or to categorize their level of activity (e.g. low, moderate and high) (37).

PAQs are widely used in epidemiologic studies because they have several advantages over objectives measures (see Table 1.5). The costs associated with PAQs are low relative to most objective measures as these are usually limited to printing. However, interviewer-administered PAQs are more costly than self-administered PAQs because of the added cost of interviewers. In comparison, the costs associated with using motion sensors include the sensor instrument, computer software and expertise (39).

Furthermore, PAQs provide information about the nature of PA being performed by asking the participant to discuss the frequency, duration, intensity, context and type of activity (37). The Baecke Questionnaire, Modifiable Activity Questionnaire (MAQ) and Framingham Physical Activity Index all measure PA taking place in the context of leisure and work while the CARDIA Physical Activity History also looks at household related activities in addition to leisure and occupation activities (40-43).

From a practical perspective, PAQs are easy to use because they can be selfcompleted by participants or administered by an interviewer (see Table 1.5). They typically require participants to recall their activities over a given period of time that can span from a 24-hour period, like the Framingham Physical Activity Index, to a year in the case of the CARDIA Physical Activity History (42, 43). As well, PAQs do not influence or impede the behaviour of participants because they simply require participants to recall activity. Other objective methods of measurement, such as a pedometer or accelerometer, can impede regular activity. Participants may find it inconvenient to wear the device and consequently compliance may be compromised (38).

PAQs can also be tailored to a specific population. This may be achieved by altering the length of the questionnaire, changing the language in which the questionnaire is given and ensuring that questionnaire items are relevant and take cultural norms into consideration. The International Physical Activity Questionnaire (IPAQ), for example, was developed to allow for comparison of standardized measured PA levels between different countries using the same questionnaire. It is available in a shorter 9-item form and a longer 31-item form and has been translated into different languages for use in countries including Japan, Sweden, the Netherlands, Brazil, Finland and Guatemala (44, 45).

22

However, there are limitations of using PAQs. Ambiguously worded items on a PAO may yield inaccurate or incorrect responses because participants are unsure of the meaning of the items. It is important for questions to be worded clearly and as simply as possibly, particularly when being administered in a pediatric population (46). As well, utilization of a questionnaire in a population unlike the one in which the PAQ was developed and validated renders it an invalid form of measurement. The reliability and validity of the PAQ must be reassessed in the new study population (46). The Baecke Questionnaire was initially developed and validated in a young Dutch population. A study by Florindo et al. (47) assessed the validity of the Baecke Questionnaire in a population of individuals with HIV or AIDS for the measurement of PA. The validity of the Baecke Questionnaire was determined using maximal oxygen uptake (VO₂ max) and peak workload as criterion measures. The Pearson Correlation Coefficient was 0.41 with VO₂ max and 0.43 for peak workload, both associations being statistically significant. Consequently, the PAQ was considered a valid PA measurement tool in this population (47).

Sources of error may also compromise the accuracy of PA measurement by PAQs. Recall bias is a concern because PAQs requires participants to retrospectively report their PA over a given period of time. As the length of the period of recall increases, recall bias is more likely (2). In a study exploring the association between arterial stiffness and self-reported PA or cardiorespiratory fitness, a weaker association was found between arterial stiffness and PA compared to the association between arterial stiffness and cardiorespiratory fitness. This may be attributable to recall bias and consequent misclassification or measurement of PA in study participants (48).

PAQs are also vulnerable to floor effects, meaning that they are not able to measure low levels of PA. Many PAQs do not measure activities below a defined level (37). For example, the CARDIA Physical Activity History only asks participants to record activities that they spent a minimum of one hour a month performing for most items (43). Similarly on the IPAQ, only activities that have been performed for a duration of 10 minutes or longer are recorded (45). This has implications for individuals that perform very low levels of physical activity because their level of PA may be underestimated by PAQs (39).

Social desirability bias may also introduce error in PAQs because they are a selfreported measure of PA. In the context of PAQs, social desirability bias occurs when study participants overestimate their PA because it is perceived to be a positive behaviour. As well, they can underestimate their sedentary activity because it is considered a negative behaviour (49, 50). A study by Warnecke et al. looked at social desirability, as measured by the Marlowe-Crowne Social Desirability Scale, and PA measurements in an ethnically diverse population. When predictors of reported PA were modeled in a multiple regression, social desirability was the strongest predictor (49). However, a study by Motl et al. suggests that social desirability bias may not be a source of overreporting in PAQs. The Stanford Usual Activity Questionnaire was administered in a college population of males and females along with the Marlowe-Crowne Social Desirability Scale. While the correlation between the two scales was statistically significant, its magnitude was small (0.08, p<0.05) suggesting that social desirability bias does not greatly influence PAQ measurements (50).

In spite of these issues, PAQs remain valuable and valid tools in the measurement of PA in epidemiologic studies. Population studies are not concerned solely with the overall estimate of EE. PAQs can provide valuable information about the determinants of and factors influencing PA, information that is not available using objective PA measures. As well, they are the only feasible tool to measure PA in large population studies and are relatively low cost. Furthermore, they have demonstrated their validity against multiple objective measures, including DLW, %BF, indirect calorimetry and VO₂ max (see Table 1.6). **Table 1.5:** Comparison of physical activity questionnaires and objective measures of

 physical activity

Characteristic	Physical Activity Questionnaires	Objective Measures
Quantity Measured	Physical activity	Total energy expenditure
		Total body movement counts
Monetary Cost	Low	Low to high
Information on Type of Activity	Yes	No
Time Cost to Participant	Low to moderate	Low to high
Interference with Participant Activity	No	Yes
Output	Categorical or continuous measure	Continuous measure
Sources of Error	Recall bias Floor effects Social desirability bias	Malfunctioning of device Poor compliance

Table 1.6: A selection of physical activity questionnaires that have been used in epidemiologic research

PAQ	Population	Time	Administered	Criterion	Validity	Retest	Reliability
	Developed	Frame		Measure		Period	
ARIC/Baecke	American	Not	Self	48-hour PA	Men = 0.59*	1 month	Total Leisure
Questionnaire	males and	specified		records	Women = $0.33*$		Time $PA = 0.86$
	females aged			Caltrac	Men = 0.24*		-0.92
	45-64			accelerometer	Women = $0.19*$		
				VO_2 max	Men = 0.57*		
					Women = 0.46*		
				% BF	Men = -0.30*		
					Women = -		
					0.51*		
Baecke	Dutch males	Not	Self	% BF	Total Activity =	3 month	Work Index =
Ouestionnaire	and females	specified			-0.49**		0.88*
	aged 20-32	1			Work Activity =		Sport Index =
	e				-0.17**		0.81*
					Sport Activity =		Leisure Index =
					-0.35**		0.74*
					Leisure Activity		
					= -0.39**		
				Caltrac	Total Activity =		
				Activity	0.19**		
				Monitor	Work Activity =		
					0.11**		
					Sport Activity =		
					0.32**		
					Leisure Activity		
					= 0.01**		

							and the second sec
				4-week activity history	Total Activity = 0.37** Work Activity = 0.05** Sport Activity = 0.40** Leisure Activity = 0.28**		
				VO ₂ max	Total Activity = 0.54** Work Activity = 0.23** Sport Activity = 0.52** Leisure Activity = 0.26**		
CARDIA Physical Activity History	American males and females aged 18-30	Past year	Interviewer in person or by telephone	% BF	Heavy Activity = -0.35** Moderate Activity = - 0.09**	1 month	Heavy Activity = 0.91 Moderate Activity = 0.66 Total Activity =
				Caltrac Activity Monitor	Heavy Activity = 0.31** Moderate Activity = 0.11**		0.88
				4-week activity history	Heavy Activity = 0.54** Moderate		

.

				VO ₂ max	Activity = 0.08** Heavy Activity = 0.63** Moderate Activity = 0.08**		
Framingham Physical Activity Index	Participants in the Framingham Heart Study	Past 24 hours	Interviewer	HR METs	Ages 45-54RuralParticipants = - 0.17 UrbanParticipants = - 0.11 Ages 55-64RuralParticipants = - 0.21 UrbanParticipants = - 0.15 RuralParticipants = - 0.65 UrbanParticipants = 0.65UrbanParticipants = 0.55	2-3 year	Total Activity = 0.30-0.59

.

ID IO	M-11	D47	0.10	00.4	T C	1 1	T C
IFAQ	iviales and	Past /	Sell or	CSA	Long form	1-week	Long form
	remates aged	days or	interviewer by	Accelerometer	(poolea) = 0.33	test-	(poolea) = 0.81
	25-49 in 14	average	telephone		Short form	retest	Short form
	countries	week			(pooled) = 0.30		(pooled) = 0.76
MAQ	Males and	Past	Interviewer	Caltrac	Including	1-3 week	Past Year
	female Pima	week and		Activity	<u>Walking</u>		Leisure Activity
	Indians aged	past year		Monitor	Leisure Activity		Ages 10-20 =
	10-59				Past Week =		0.37
					0.80		Ages 21-36 =
					Leisure Activity		0.92
					Past Year $= 0.69$		Ages 37-59 =
					Work Activity		0.88
					Past Year $= 0.41$		Past Week
					Total Activity		Leisure Activity
					Past Week =		Ages 10-20 =
					0.59		0.35
					Excluding		Ages 21-36 =
					Walking		0.62
					Leisure Activity		Ages 37-59 =
					Past Week =		0.77
					0.62		
					Leisure Activity		
					Past Year $= 0.44$		
					Work Activity		
					Past Year $= 0.41$		
					Total Activity		
					Past Week =		
					0.53		

Stanford Usual	American	Usual	Interviewer	% BF	Moderate	1 month	Moderate
Activity	males and	activity			Activity $= -0.33$		Activity $= 0.77$
Questionnaire	females aged	or past 3			Vigorous		Vigorous
-	25-64	months	,		Activity $= -0.16$		Activity $= 0.67$
				Caltrac	Moderate		-
				Activity	Activity $= 0.23$		
				Monitor	Vigorous		
					Activity $= 0.22$		
				4-week	Moderate		
				activity history	Activity= 0.05		
					Vigorous		
					Activity=0.28		
				VO ₂ max	Moderate		
					Activity $= 0.27$		
					Vigorous		
					Activity = 0.38		

* Pearson Correlation

**Spearman Correlation

1.2.2 - Properties of physical activity questionnaires

1.2.2.1 - Reliability

Reliability refers to the consistency of a questionnaire (46). This can be determined by examining the internal consistency and the test-retest reliability of the PAQ. Internal consistency refers to the homogeneity of a questionnaire and can be assessed using a variety of methods including the split halves method, item-total correlation, the Kuder-Richardson formula 20 and Cronbach's alpha (46). The splithalves method involves dividing a questionnaire in half and then assessing the correlation between the halves (51). In item-total correlation, an item is removed from the questionnaire and it is then correlated against the total score. The Kuder-Richardson formula 20 is used when items are dichotomously scored while Cronbach's alpha is employed for items with more than two answer options (46, 51).

One of the aims of PAQs is to yield consistent scores upon repeated administration of the questionnaire, which can be assessed using test-retest reliability. Typically, a questionnaire is administered twice and the correlation of the scores is calculated using the Pearson Correlation Coefficient if the data are normally distributed. In a study evaluating the test-retest reliability of the IPAQ, the Spearman correlation coefficient was used instead of the Pearson Correlation Coefficient because the data were not normally distributed (45). For categorical or dichotomous outcomes, a weighted Kappa coefficient is used to determine test-retest reliability (46). The interval of time between the test and retest can vary widely from a day to a year or longer. The assumption underlying test-retest reliability is that no actual changes in the true score have occurred during the time period between tests (46). As the period of time between the test and retest increases, the reliability of the measure typically declines as the activities that an individual reports in their retest may be dissimilar from those originally reported in the first test. However, if the questionnaire is re-administered too quickly, the participant may remember their original responses and record those answers from memory, resulting in an inflated estimate of reliability (46, 52).

1.2.2.2 – Validity

While it is essential that a PAQ yields reproducible results, it is equally important that the PAQ must be measuring the domain, PA, that it intends to measure (46). Determining the validity of PAQs in a population is important because a PAQ with poor validity will yield results that are not useful. There are three main areas that are taken into consideration when assessing the validity of an instrument: content validity, criterion validity and construct validity.

Content Validity

The content validity of an instrument is the ability of the instrument to include all relevant content that is encompassed by the construct being measured. An instrument that includes a wide range of domains encompassed by the construct should be better able to assess the construct being measured. For example, a PAQ that only measures PA in the context of work would be a less accurate measure of PA compared to a PAQ that measures PA in a variety of contexts, such as leisure time and household activities (46, 52). A PAQ with good content validity can make inferences that are valid in different circumstances because it accounts for and includes various domains relevant to measuring PA (46).

Developing a PAQ with good content validity can be challenging because it should include the activities performed by the study population that account for the majority of EE. This is particularly challenging in a multiethnic population since the activities that account for most of the EE in one ethnic group may not account for much of the activity in another ethnic group. PAQs should have a sufficient number of items to adequately include the relevant aspects of PA. A PAQ comprised of a single item will have poor content validity because it cannot adequately cover the breadth of PA (46). For example, the IPAQ demonstrates good content validity because it takes activity in multiple contexts into consideration, measures different features of PA and can be culturally adapted (45). Furthermore, PAQs should ask participants to give information on the specifics of the PA they perform. This includes information on the context, type, duration, frequency and intensity of their activity (53).

To ensure that a questionnaire does include all relevant facets of PA, items may be taken from a variety of sources. This can include taking items from an existing questionnaire. One of the advantages of this approach is that the individual items have already been validated before and are likely still useful. As well, using existing questions can save time in generating and wording questionnaire items (46). The population in which the questionnaire will be administered can improve the content validity of a PAQ, particularly if the PAQ is being modified for use in a population different from the one in which it was originally developed. By piloting the PAQ in a sample of the study population, missing relevant domains can be identified and then added to the questionnaire (46).

Criterion Validity

Once a PAQ has been developed, its properties must be assessed to ensure that it is a valid tool for measuring PA. This is referred to as the criterion validity of the PAQ and it involves determining the correlation of the PAQ against a measure that is considered a criterion standard (46). There are two forms of criterion validity: predictive validity and concurrent validity. Predictive validity evaluates the ability of an instrument to predict the construct that it aims to assess while concurrent validity determines the correlation of the measure against a criterion measure while both measures are simultaneously administered. The validation of PAQs is primarily concerned with the concurrent validity of the questionnaire (46, 52).

The criterion validity of PAQs can be assessed against a gold standard, such as DLW, but can also be assessed against objective measures of PA such as motion sensors and measures of physical fitness (54). Criterion validation of PAQs poses challenges because there are multiple existing PA standards. For example, a PAQ emphasizing cardiorespiratory fitness may have poor association against fitness measures of strength

35

(53). The validation of a questionnaire is useful not only in determining whether an instrument is a valid measure, but also in identifying other methodological issues of interest, such as overreporting of PA. VO₂ max was employed as a criterion measure in a study validating the IPAQ and revealed that overreporting was an issue. Nearly 10% of the participants had reported high levels of physical activity on the IPAQ but did poorly on the VO₂ max test (44).

Construct Validity

When PA is being measured by indirect means, such as a PAQ, the construct validity of that approach needs to be assessed (51). The construct validity of an instrument must be determined when the quantity being measured cannot be "operationally defined" or when no criterion measure exists that completely encompasses the construct being measured. The construct validity of a PAQ assesses the extent to which a PAQ measures the hypothetical construct of PA (46, 55). The construct validity of a questionnaire measuring self-reported historical PA was assessed by Bowles et al. in a population of males and females aged 18-80. The association between cardiorespiratory fitness (treadmill test performance) and self-reported PA was determined. As well, health markers (glucose, cholesterol, triglycerides, SBP and BMI) were used to assess the ability of the PAQ to categorize individuals as sufficiently or insufficiently active (53). Recalled PA was significantly associated with performance on the treadmill test (P < 0.001). Significant differences between sufficiently active and

insufficiently active groups were found for triglyceride level and BMI in both men and women. Based on these results, the PAQ was considered to have construct validity (53).

The construct validity of an instrument can be assessed through different methods. These include group differences, factor analysis and experimentally assessing the construct validity of the questionnaire (51, 56). Administering a PAQ in two groups with known differences, such as active and sedentary individuals, can help determine its construct validity. The expectation is that active individuals are more likely to receive higher scores on the PAQ compared to the sedentary individuals. If the PAQ can distinguish between these two groups, it can be inferred that the questionnaire is measuring PA (51).

Factor analysis is a statistical technique based on the theory that the hypothetical construct being measured consists of smaller factors. For example, aerobic activity and strength activity could be considered factors of PA. Items on the PAQ should load onto a factor and items that load onto the same factor should have good correlation with each other (46, 51). The construct validity of a PAQ may also be determined by using the questionnaire experimentally, as was done in Bowles et al.'s study. The PAQ evaluated was able to categorize individuals by PA status (sufficiently or insufficiently active) using health markers and was well associated with a cardiovascular fitness measure (53).

37

Seasonal Variation

Seasonal variations in PA have been noted in cross sectional and longitudinal studies (57). The Seasonal Variation of Blood Cholesterol Study explored patterns in total, leisure time, work related and household PA over the course of a year. Male and female participants aged 20-70 completed three 24-hour PA recalls at five different points during the course of a year and information on environmental information, such as weather, temperature, and daylight was also collected to determine whether these factors influenced PA. Total PA was higher in both men and women during the summer (1.4 and 1.0 MET-hours/day, respectively). Further examination of PA patterns revealed that moderate non-occupational PA, including household and LTPA, was 2.0-2.4 METhours/day higher in men and women (57). Uitenbroek also explored seasonality in LTPA in a Scottish population aged 18-60 (58). For a three-year period, participants completed a daily interview that asked about their PA for the past month and the number of times they participated in PA for a minimum duration of 20 minutes. PA was highest from June to August and lowest from December to February. The percentage of participants engaging in PA three times a week or more during the winter was 23% in contrast to 32% during the summer (58).

These seasonal trends are also evident in pediatric populations. Rifas-Shiman et al. investigated differences in reported PA by males and females aged 9-14 (59). Participants completed two different PAQ formats: annual and seasonal. Similar to Uitenbroek's study, higher PA was reported during the summer and the lowest rates of PA occurred in the winter. Girls performed 8.2 hours/week of PA in the winter compared to 17.1 hours/week in the summer while boys spent 9.2 hours/week in the winter and 20.6 hours/week in the summer. Median PA estimates were over-reported on the annual PAQ compared to the seasonal PAQ. Estimates for girls were 2.3-4.7 hours/week higher on the annual PAQ and 1.9-4.4 hours/week higher for boys. This discrepancy may be a result of misreporting seasonal activities as though they are performed throughout the year (59). These studies highlight the fluctuations that occur in PA and the importance of taking seasonality into account when measuring PA.

The MAQ is designed to measure PA performed over a shorter period of time (past year and past week) in addition to PA performed over the course of a participant's life (60). Developed by Kriska et al. in a Native American population, it measures both leisure and occupational PA and also measures low-level PA, such as watching television. Estimates of PA are obtained by multiplying the number of hours spent performing a given activity by its MET value. MET-hours per week or total PA hour estimates are then calculated (60).

One of the unique features of the MAQ is that it measures PA over three time frames, which has several advantages. One of the concerns about measuring PA over a short period of time is that activity may not necessarily be representative or typical of the participant as factors including seasonality can influence activity. Conversely, activity recall over longer periods of time is prone to recall bias. By estimating PA in both contexts, in addition to an estimate of lifetime activity, a more accurate measure of participants' PA can be obtained (60).

1.2.3 - Criterion measures of physical activity

DLW, motion sensors such as pedometers and accelerometers, and physical fitness measures are among the criterion measures that have been used to measure TEE (see Table 1.7). They are commonly used in kinesiology studies where a precise measure of TEE is required. Because they are objective measures and not subject to reporting errors and recall biases, these criterion measures are used in the validation of PAQs (38).

Doubly Labeled Water

DLW has been frequently employed to validate PAQs in various populations (61-63). DLW was developed by Nathan Lifson in the 1950s. The technique involves ingestion of known concentrations of two stable isotopes, deuterium (²H) and heavy oxygen (¹⁸O), which are dissolved in water. Deuterium is excreted from the body as water and oxygen is excreted as both water and carbon dioxide (CO₂) (64). Urine samples are then periodically collected to determine changes in isotope concentrations in the body. The difference in elimination rates between ²H and ¹⁸O provides an estimate of CO₂ production: $R_{H2O} = TBW (K_H)$

$$R_{H2O} + 2 r_{CO2} = TBW (K_O)$$

 $r_{CO2} = TBW (K_0 - K_H) / 2$

where r_{H2O} is the water output, r_{CO2} is the CO₂ output, TBW is total body water, K_O is the rate of ¹⁸O elimination and K_H is the rate of ²H elimination (65). TEE is then calculated using the CO₂ output and yields an estimate in kcal (65). The difference in the rates of ²H and ¹⁸O production is a measure of oxygen consumption, which can be estimated through the output of CO₂ (66).

Because of its accuracy, DLW is considered the gold standard in assessing TEE and is frequently used to validate PAQs (61). Furthermore, it is considered a safe procedure because the isotopes used, ²H and ¹⁸O, are stable and not radioactive (67). However, it is impractical to employ DLW in epidemiologic studies because it is costly and difficult to administer to a large population. Furthermore, it does not offer insights into the duration, context, type and intensity of activity being performed which are often important to note in epidemiologic studies (64).

Pedometers

Pedometers are instruments that measure the numbers of steps taken or the total distance walked, and the participant typically wears them at the waist. Movement is measured by a lever that is suspended by a spring. The lever moves up and down in

response to vertical acceleration and each movement is counted and recorded by the pedometer (68). Earlier pedometers were mechanical in nature and because they were inaccurate and prone to measurement error, they were not considered a valid measure of PA. However, current pedometers are typically electronic and demonstrate greater accuracy in measuring the number of steps taken (39). They have demonstrated variable criterion validity against other measures, including direct observation, HR, accelerometers and oxygen consumption (54, 69). In a systematic review of the convergent validity of pedometers by Tudor-Locke et al., pedometers correlated most strongly with accelerometers (median correlation = 0.86) and observed activity (median correlation = 0.68). However, pedometers did not correlate as well with self-reported PA (median correlation = 0.33) and EE (median correlation = 0.68). Pedometers had negative correlations with observed inactivity (median correlation = -0.44) and time spent sitting (median correlation = -0.38) (69).

Pedometers are a lower cost option when compared to accelerometers. The average cost of a pedometer is between \$10-30 per unit while accelerometers range in cost from \$50-400 (estimated 2001 costs) (69). Because pedometers are sensitive to vertical acceleration, they are best able to measure ambulatory activities. As a result, they are well suited to studies primarily concerned with the measurement of walking activity or a study where the majority of PA performed by participants is walking (70).

However, there are limitations in employing pedometers to measure PA. Their utility in determining EE is limited because only the number of steps taken is recorded.

There is no information measured regarding the length of strides taken, the total displacement of the movement or intensity of activity. Furthermore, pedometers do not offer any information on the context in which activities are being performed and cannot assess activity involving the upper body or non-load bearing activity (54, 70). Malfunctioning remains a concern, particularly for movement sensors, as they will not work when damaged or dropped (38). Patient compliance is also an issue. Although pedometers are small and lightweight, they may hinder the activity of some participants who may be concerned that the pedometer may fall off or break. Consequently, participants may choose to restrict their activity or not wear the sensor at all (38).

Health-Related Fitness Measures

Various health-related fitness indicators have been used in the validation of PAQs. There are five main components of health-related fitness: cardiorespiratory fitness, body composition, muscle endurance, muscle strength and flexibility (1). The rationale for employing fitness measures as a criterion measure is that increased levels of PA are one of the factors influencing physical fitness measures. For example, it can be inferred that an individual reporting high levels of PA should accordingly have a high level of fitness (1, 71).

The fitness components used most frequently to validate PAQs are cardiorespiratory measures, muscle endurance and body composition (see Table 1.8). Because fitness encompasses numerous components, validation of a PAQ should not be done solely using a measure of fitness but instead, should be assessed in conjunction with other objective measures of PA. A questionnaire designed to measure a specific area of health-related fitness should employ an appropriate fitness criterion. For example, the validity of questionnaire items pertaining to vigorous physical activity should be assessed using a fitness measure of aerobic capacity, such as $VO_2 \max (72)$. The IPAQ is intended to measure activities pertaining to cardiorespiratory fitness. In a study validating the IPAQ against fitness measures in a population of Finnish men, $VO_2 \max$ was used as the main criterion measure (44).

One of the limitations of physical fitness as a criterion measure is that it is an indirect measure of PA. Physical fitness is reliant upon other factors including genetics, age, obesity, smoking and diet and consequently may not accurately reflect PA (72). Furthermore, some physical fitness measures may have good correlation with structured PA but have poor association with unstructured activity (47). The Baecke Questionnaire is concerned primarily with habitual activity. Its validity against VO₂ max was assessed in Brazilian men and women with HIV/AIDS. While VO₂ max had good correlation with physical exercise taking place in the context of leisure (r = 0.41), it was not significantly correlated with leisure and locomotion activities (r = 0.19), work related activities (r = 0.14) or the total questionnaire score (r = 0.27) (47).

Characteristic	Doubly Labeled Water	Pedometers	Health-Related Fitness Measures	Accelerometers
Quantity Measured	CO ₂ production	Number of steps taken	Varies	Vector magnitude
Monetary Cost	High to Very High	Low	Moderate to Very High	Low to Moderate
Information on Type of Activity	No	No	No	No
Time Cost to Participant	Moderate	Low	Moderate to High	Low
Interference with Participant Activity	Low to High	Low to Moderate	Low	Low to Moderate
Output	Continuous measure	Continuous measure	Continuous measure	Continuous measure
Sources of Error	Error if food quotient not known	Malfunction Lack of precision	Indirect measure of PA	Malfunction

 Table 1.7: Comparison of objectives measures of physical activity

Table 1.8: A list of fitness components and the methods by which they are commonly

evaluated

Fitness Component	Evaluation Method		
Cardiorespiratory Fitness	VO ₂ max via Ergometer cycle		
Body Composition	BMI		
	%BF		
	Skinfolds		
Muscular Strength	Dynamometer		
	Muscular Fitness Tests (e.g. sit ups, push		
	ups, squats)		

1.2.3.1 – Accelerometers

Accelerometers electronically detect and record body movement through the use of piezoresistive or piezoelectric sensors in a spring mass system (see Figure 1.1) (64). When a mass within the accelerometer experiences acceleration, it compresses or expands a spring and this displacement is recorded by the accelerometer's sensor (73). The theoretical basis for using accelerometers to assess EE is that acceleration and muscle exertion are directly proportional to each other. Consequently, acceleration should be related to EE (70).

There are three main types of accelerometers: uniaxial, biaxial and triaxial. Uniaxial and triaxial accelerometers are most commonly used in epidemiologic research (74, 75). Uniaxial accelerometers detect acceleration of the body in one plane, typically the vertical plane, while biaxial accelerometers detect movement in two planes. Triaxial accelerometers measure movement in three planes: vertical, horizontal and mediolateral (see Figure 1.2). Because triaxial accelerometers are able to detect movement in multiple planes, they are considered a better assessor of EE than uniaxial accelerometers (70).

Accelerometers share some advantages with pedometers, including being portable and nonreactive (54). However unlike pedometers, accelerometers are able to assess the frequency and intensity of movement to some extent. Accelerometers record changes in velocity which can be considered a measure of intensity. When information from the accelerometer is downloaded into computer software, the duration of activity at a given frequency can be determined (39, 73). As well, some accelerometers may provide a more accurate measure of TEE because they take resting EE and EE from recorded activities into account. For example, the TriTrac R3D triaxial accelerometer (TriTrac) (Professional Products, Madison, WI) uses sex, age, weight and height to compute resting EE, which is then used in its estimation of TEE (76). Accelerometers can also record activity for a period of up to two weeks, which may provide a more representative estimate of EE than would a shorter time period (73).

There are practical issues prohibiting the use of accelerometers in larger studies. These include cost and time. The cost of accelerometers can range from \$50-400 (39). As well, downloading accelerometer data into the software requires some level of computer knowledge. Including time to train staff can make lengthen the process of data management (39). Accelerometers are also prone to malfunctioning if they are dropped and cannot provide information on the context of activity, like pedometers (38). One of the concerns about accelerometers is that they do not have adequate sensitivity to free-living activities. The CSA, TriTrac and Biotrainer accelerometers were evaluated in young to middle-aged men and women. Participants performed ambulatory activities, including walking and jogging on a treadmill, as well as household activities, including vacuuming, sweeping, raking, mowing, shoveling and stacking. Using VO₂ max as a criterion measure, all three accelerometers tended to overestimate treadmill activity while underestimating the household activities (77).

Accelerometers also may not have sufficient sensitivity to take changes in grade into account. A study looking at the validity of the TriTrac in 60 young adults found that accelerometers demonstrated poor sensitivity to changes in grade while walking or running on a treadmill. Estimated EE values by accelerometer data for 0% incline and 5% incline were 0.1167 and 0.1159 kcal/ kg*min, respectively, while walking at a speed of 6.4 km/h. Comparatively, EE measured by indirect calorimetry was 0.0926 kcal/kg*min for 0% incline and 0.1266 kcal/kg*min for 5% incline at the same speed (76).

There are also issues specific to the use of accelerometers in pediatric populations. Accelerometers collect data in intervals and the length of intervals can influence PA measurement in children. Children typically do not to engage in activity at a given intensity for a lengthy period of time. Instead, they tend to engage in PA at low and moderate levels while sporadically increasing the intensity of their activity (78).

48

Consequently, accelerometers with a longer recording interval may fail to record the short periods of intense activity (79).





Note: From Mathie M.J., Coster A.C.F., Lovell N.H., and Celler B.G. (2004).

Accelerometry: providing an integrated, practical method for long-term, ambulatory

monitoring of human movement. Physiological Measurement. 25(2);R1-R20.
Figure 1.2: Triaxial planes



1.2.3.2 - RT3 triaxial accelerometer

The RT3 triaxial accelerometer (RT3) (Stayhealthy Inc., Monrovia, CA), which is based on the R3D TriTrac (Professional Products, Madison, WI), was used in the SHARE-AP Action study. The TriTrac has been used in research since 1992 and has been validated in many populations for the assessment of PA (80). While it is considered an accurate PA measurement tool, the TriTrac has weaknesses that the RT3 was designed to address.

The accelerometers used in the two devices are different. The TriTrac uses three separate accelerometers to measure movement and they are manually assembled which

may result in inaccurate measurement by variant placement of the accelerometers. In contrast, the RT3 employs a single chip that integrates the measurement of acceleration in three planes, diminishing the likelihood of error resulting from construction (81). As well, the RT3 is considerably smaller and lighter than the TriTrac. The RT3 weighs 62.5g and its dimensions are $7.1 \times 5.6 \times 2.8 \text{ cm}^3$ (82) while the TriTrac is 170 grams in weight and its dimensions are $11.0 \times 6.9 \times 3.3 \text{ cm}^3$ (76).

The RT3 is worn on a belt at the waist by an individual, has four operation modes and is battery powered. Data can be collected and stored up to 21 days. Activity can be measured at 1 second or 1 minute intervals and measurements of the three individual axes or the vector magnitude may be recorded. The vector magnitude value is the sum of the three vectors (VM = $[x^2 + y^2 + z^2]^{1/2}$) (83). Prior to wearing the motion sensor, the participant's sex, age, weight and height is entered to initialize the sensor (83). The units of output are METs, kcal or activity units (81).

Reliability

Because the RT3 was recently developed, its reliability and validity properties have not been extensively investigated. Powell and Rowlands studied the reliability of the RT3 using eight accelerometers (83). A single female participant wore four RT3s on each hip while walking, running, sitting and standing. Two trials of all activities were conducted. The intermonitor coefficient of variation (CV) for walking and running at different speeds ranged from 1.5-6.0%. However, the reliability of the RT3 while sitting and standing was lower, with the CV ranging from 8.7-25.6% (83). Kozub et al. also assessed the reliability of the RT3, but in a population of children and adolescents with visual impairments (84). Participants wore a RT3 device on each hip while performing a variety of activities including scooters, horseshoes, archery, basketball, volleyball, dance and bowling. The intraclass correlation for the entire sample was 0.98 (P < 0.001). The reliability was still high when looking at the coefficient by sex. The reliability estimate for males was 0.96 (P < 0.001) and 0.98 for females (P < 0.001) (84).

Validity

Studies by Kozub et al. and DeVoe et al. found moderate validity estimates for the RT3 (see Table 1.9) (84, 80). Kozub et al (84) used the Children's Physical Activity Form (CPAF) as a criterion measure. The CPAF is an observational system where PA is monitored each minute and the activity recorded is then ranked using a four category coding system. The intraclass correlation for all data was 0.89 (P < 0.001) and high correlations were also found when looking at males and females separately. The validity estimate for females was 0.87 and 0.94 in males (both P < 0.001). Intraclass correlation for specific activities varied from 0.50 - 0.72 (84).

The RT3 also demonstrated good validity in a study by Rowlands et al (82). Males in primary school and university participated in structured and unstructured activities while wearing an RT3 attached to a belt. The RT3 demonstrated good validity against oxygen consumption, which was the criterion measure,. The correlation between vector magnitude and oxygen magnitude ranged from 0.831 to 0.854 for all participants (82). In contrast, DeVoe et al. found the RT3 had varying validity against oxygen consumption and heart rate (HR) as criterion measures for walking and jogging on a treadmill (80). Higher correlations were found when participants used the treadmill at 0% grade (oxygen consumption: 0.91, P < 0.001; HR: 0.83, P < 0.001) but correlations were lower at higher grades (oxygen consumption: 0.32, P<0.001; HR: 0.45, P<0.001). Overall, the RT3 vector magnitude was only moderately correlated with oxygen consumption (0.57, P<0.001) and HR (0.51, P<0.001). The RT3 had very good correlation against the TriTrac (0.96, P<0.001) (80).

Authors, Year	Criterion Measure	Population	Ν	Outcome (correlation)
Kozub et al.,	et al., Children's Children and 19 (9 boys		19 (9 boys	Scootering = 0.50
2005 P. A F	Activity Form	adolescents with visual impairments	and 10 girls)	Recreational Activities = 0.58
		aged 6-18		Games/Sports = 0.56
				Dance =0.66
				Bowling $= 0.72$
				Spearman correlation
Rowlands et	Rowlands et al., 2004Oxygen consumptionBoys (mean age 9.5 ± 0.8 years) and men (mean age 20.7 ± 1.4 years) $34 (19 \text{ boys})$ and 15 men)	34 (19 boys	All Activities	
al., 2004		9.5 \pm 0.8 years) and men (mean age 20.7 \pm 1.4 years)	and 15 men)	Vector Magnitude = 0.854
				VM = 0.848
				HM = 0.837
				MLM = 0.817
				Treadmill Activities
				Vector Magnitude = 0.831
				VM = 0.824
				HM = 0.789
				MLM = 0.761
				Unstructured Activities
				Vector Magnitude = 0.852
				VM = 0.848

Table 1.9: Validation of RT3 triaxial accelerometer

				HM = 0.853
				MLM = 0.805
DeVoe et al.,	TriTrac,	Males (mean	17 (12	TriTrac = 0.96
2003	Oxygen consumption and HR	age 22.7 <u>+</u> 2.3 years) and females (mean	males and 5 females)	Oxygen consumption = 0.57
		age 24.2 <u>+</u> 1.5 years)		HR = 0.51

VM = Vertical Magnitude

HM = Horizontal Magnitude

MLM = Mediolateral Magnitude

1.3 - Physical activity measurement in non-white ethnic groups

Differences in PA among different ethnicities have been noted in both American and Canadian studies (85, 86). However, whether these differences are actually true differences or the consequence of measurement error of PA in ethnic groups is uncertain. While variations in rates of CVD, obesity and diabetes across ethnic groups suggest that actual differences may exist; however, these estimates may be inflated by error resulting from measurement issues related to ethnicity (87). Kriska suggests that measurement error may be attributed to ethnicity-related issues or the types of activities that are included and excluded in PAQs (88).

One of the measurement issues related to ethnicity is cross-cultural equivalence. While it is a multidimensional concept and various forms of equivalence have been identified, most of these can be defined as pertaining to either interpretive or procedural equivalence (89). Interpretive equivalence refers to whether people of different ethnicities perceive the same question in an identical manner or meaning. This concept is crucial to the assessment of PA across different ethnic groups because comparisons between groups cannot be made if they are not assessing the same concept (89). Procedural equivalence, conversely, is concerned with the process of the methods and measures used that facilitate comparison by ethnicity. This encompasses how questionnaire items are worded, the validity of PAQs and the process of comparing results between different ethnic groups (89). The effectiveness of PAQs to accurately measure PA is partly contingent on their ability to ask participants about the activities that comprise the majority of their EE (88). Misclassification of the PA level of participants is likely to occur if activities contributing significantly to EE are not included in PAQs (90). Many surveys focus on LTPA which white ethnic groups participate in most compared to non-white groups. However, rates of WRPA are highest among non-white groups and it is important to account for this activity (91). Similar trends apply to household PA, which non-white groups are more likely to perform compared to white groups (92).

Kriska also raises two important concerns that must be addressed when discussing variability in PA among ethnic groups. Non-white ethnic groups tend to have a disproportionate amount of people in lower socioeconomic status (SES) groups compared to white ethnic groups. Consequently, perceived differences between ethnic groups may not be a consequence of ethnicity but instead, of SES and confounding of ethnicity and SES must be considered (88). Ford et al. looked at PA patterns in high and low SES American individuals (93). Differences in total physical activity (TPA), LTPA, WRPA, household PA and walking were explored. High SES men and women were more likely to engage in LTPA than their low SES counterparts. Comparatively, low SES individuals (93). As well, while differences exist among ethnic groups, they also exist within groups. Acknowledging the variation within groups may shed light on the factors influencing participants to engage in particular forms of PA (88).

57

1.3.1 - Literature review

Qualitative studies comprise the majority of research exploring issues in PA measurement in non-white ethnic groups (see Table 1.10). Warnecke et al. explored four tasks that are executed when an individual answers a questionnaire item: interpretation, memory retrieval, judgment formation and response editing, and whether ethnicity influenced these processes (49). Individuals of African American, Mexican American, Puerto Rican and non-Hispanic white ethnicity were included in the study. While recall strategies were not influenced by ethnicity, differences by ethnicity were apparent in interpretation of questions, judgment formation and response editing. African Americans were most likely to perceive household activities as PA compared to white ethnic groups. Additionally, African Americans also perceived yard work as PA relative to Hispanics. Interpretation was also influenced by ethnicity depending on the structure of the question. When questions were open-ended, ethnicity influenced how questions were answered. In contrast, when the question included broad categories of types of PA, less variation was noted in responses because participants were given cues on how to answer the question (49).

Non-white ethnic groups were more likely to select extreme categories when responding to questions that asked participants to indicate their response on a scale from "strongly disagree" to "strongly agree". While white participants averaged 6.7 extreme responses, African Americans averaged 7.8 responses, Mexican Americans 8.1 responses and Puerto Ricans averaged the most extreme responses, 8.7. Response editing was also evident in particular ethnic groups. Based on the results of the Marlowe-Crowne social desirability questionnaire that all participants completed, African Americans and Mexican Americans were most likely to exhibit the social desirability trait and edit their responses to appear to demonstrate desirable behaviour (49).

Johnson et al. explored whether ethnic differences in comprehension of questionnaire items existed across four groups. African-American, Mexican American, Puerto Rican and non-Hispanic white participants answered questions taken from health surveys that varied by length, level of abstraction, reading level, response and qualified judgements (94). As question length increased, white and African Americans had the most difficulty in question comprehension. All groups experienced comprehension difficulty as the level of abstraction increased and there were no differences among groups. Puerto Ricans were most sensitive to changes in reading level. As levels increase, the probability of having difficulty in interpreting the question also increased. Non-white groups were more likely to have difficulty understanding questions when the question required a yes/no answer or a verbal label, such as "excellent/good/fair/poor". In particular, African Americans were more likely to have difficulty with yes/no responses and African Americans, Puerto Ricans and Mexican Americans were all more likely than whites to have difficulty responding to questions requiring verbal labelling (94). As well, Mexican Americans and Puerto Ricans were most likely to have difficulty understanding questions that required qualified judgements. Qualified judgements

included questions that asked participants to exclude certain items from consideration or specified a time frame.

Focus groups conducted by Tortolero et al. suggested that interpretation of terminology was influenced by culture. Hispanic and African American women completed a questionnaire that asked open-ended questions about their views of PA (95). The participants did not make large distinctions between the terms PA and exercise, although some participants thought that PA encompassed a broader range of activities than exercise, such as daily tasks like walking and childcare. Most women articulated that they did not participate in sports and felt that only men and children participated in sports. Some also expressed that it may not be culturally acceptable for women to engage in sports activity. Furthermore, leisure activity was negatively perceived with Hispanic women expressing strong opinions against it because they perceived it to be synonymous with being lazy or inactive (95).

The study also found that SES may be a stronger predictor of differences in PA among different ethnic groups. The types of PA performed did not vary greatly by ethnic group nor did most participants believe that their culture influenced their activity as much as economic factors, which may limit the opportunity to participate in activities. However, at the same time, most women felt it was important for surveys to include cultural activities like ceremonies and traditional dancing so participants feel like the PAQ is relevant to them (95).

60

The Cross-Cultural Activity Participation Study explored perceptions of PA in a population of African American and Native American women through semi-structured interviews (90). Some of the participants felt their ethnicity has shaped their views of PA, citing marginality, history and cultural pride as factors playing a role in their perceptions. The concept of PA in leisure time, for example, was meaningless for an African American participant because she was raised under oppressive conditions where the opportunity for free time never existed. In contrast, some Native women expressed that PA was part of their cultural ceremonies and traditions but that these had been partially lost while adapting to modern urban culture (90). This highlights the need for PAQ items to be unambiguous and specific since interpretation of the questions may render a range of results. Consequently, PAQs developed in a monoculture population may not adequately address these issues.

Table 1.10:	Studies exploring	ethnic issues i	in physical	activity measurement

Author, Year	Population	Ν	Study Design	Outcome
Warnecke et al., 1997	American adults aged 18-50 of African, Puerto Rican, Mexican and non- Hispanic white descent	423	Cross sectional	Ethnic differences were found in interpretation, judgement formation and response editing. In interpreting questions, African Americans were most likely to perceive household activities and yard work as PA compared to other ethnic groups. Puerto Ricans were most likely to select extreme categories (8.7 extreme responses on average) while whites were least likely to do so (6.7 average extreme responses). African Americans and Mexican Americans were most likely to exhibit the social desirability trait when completing questionnaires.
Tortolero et al., 1999	American women aged 35-75 of African and Hispanic descent	84	Focus group	Participants did not make large distinctions between the terms PA and exercise, although some participants thought that PA should include daily tasks. Most women did not participate in sports and said that only men and children participated in sports. Hispanic women negatively perceived leisure activity because they equated it with being lazy or inactive.
Henderson and Ainsworth, 2003	American women aged 40 and older of African and Native American descent	56	Semistructured interviews	Perceptions about the influence of ethnicity on views of PA varied among participants. Some participants did not believe they had been influenced while others found their ethnicity shaped their concept of what was considered PA.
Johnson et al., 2006	American adults aged 18-53 of African, Mexican, Puerto	345	Interviews	Ethnic differences were noted for 4 characteristics of questionnaire items: question length, reading level, qualified

Rican and non-	judgements and type of response.
Hispanic white	
descent	

1.4 – Sex, ethnicity and socioeconomic status and physical activity

Variations in PA by demographic factors, including sex, ethnicity and SES have been observed in multiple cross sectional studies, including national surveys such as the National Health and Nutrition Examination Survey (NHANES) and the Canadian Community Health Survey (CCHS) (85, 86). Although these appear to be unambiguous and straightforward categorizations, there is debate over the use of these terms and what they actually mean.

Sex differences in PA and PA patterns are frequently explored in population studies. The term 'sex' refers to a classification of male and female on the basis of biological traits that are determined by sex chromosomes. The term 'gender' is often interchangeably used with sex, although they are not synonymous. Classification by gender refers to the social or cultural characteristics that have traditionally been ascribed to a particular sex (96). Consequently, the terms should not be considered equivalent because they refer to different aspects of being male and female.

The term 'race' was originally formulated as a biological construct to categorize people under the assumption that genetic differences existed among difference races. However, studies have shown that the genetic variation within a race is greater than the variation among groups (97). Consequently, the utility of distinguishing people by race is limited and should not be employed. The concept of race has evolved to become a social construct that is grounded largely in perceptions of physical appearance, behaviour and biology. In contrast, the term 'ethnicity' is a socio-cultural construct that is based on multiple facets, including ancestral origin, language, religion and cultural traditions (97). Because it includes potential determinants of disease, such as environmental, cultural and behavioural factors, ethnicity is an important dimension to explore in research (98). Most of the studies included in the literature review (see Table 1.11) failed to make a distinction between race and ethnicity, often lumping the terms together. Only two studies used ethnicity to categorize participants.

SES refers to the social and economic factors that influence a person's standard of living. Different variables have been used as measures of SES including income, wealth, occupation and education (99). Studies have used different approaches to determine SES. Some studies have used only one measure. For example, Sallis et al.'s study in a high school population used school district as a measure of SES (100) while SES was measured in He et al.'s exploration of PA patterns as years of education completed (91). Selecting one variable for SES limits the utility of the measure, as there is a greater likelihood that participants may be misclassified or that the variable selected does not adequately differentiate participants. Conversely, SES can be assessed by looking at a number of variables. The CCHS observed variations in PA by education and household income level to determine the effect of SES (85). Crespo et al.'s study assessed social class using education, family income, poverty status, employment status, occupation and marital status (86). Another approach in measuring SES in studies involves developing an index or scale that includes multiple indicators. The Social Disadvantage Index (SDI),

65

developed by Anand et al. using the Study of Health Assessment and Risk in Ethnic Groups (SHARE) and Study of Health Assessment and Risk Evaluation in Aboriginal Peoples (SHARE-AP) study populations, uses employment status, marriage status and income to yield a maximum score of 5 (101).

1.4.1 – Physical activity patterns

The aim of PAQs is to measure the activity that comprises the majority of EE in a population. Earlier PAQs focused on the measurement of work related and leisure activities such as the Baecke and the MAQ because it was thought these activities comprised the majority of EE in most populations (55, 102). However, measuring activity in only these contexts may not account for the majority of EE and it is necessary to measure PA in other contexts such as household activities (92). Consequently, assessing PA in different contexts including occupational, leisure time and household may provide a more accurate measure of PA.

In addition, the measurement of sedentary activity should also be performed when estimating PA, although less attention has been devoted to the assessment of sedentary activity. Sedentariness is not the absence of PA but rather, the participation in activities that involve low or very low levels of EE (36). Spanier et al. suggest that an increase in sedentary activity, and not necessarily a decrease of PA or increase in calorie consumption, has contributed to an increase in obesity. Trends in PA show slight increases in participation rates and caloric intake, neither of which can solely account for current obesity patterns. Consequently, further investigation into sedentary behaviour is warranted and necessary to better understand the role that sedentary activity may play in obesity (36).

1.4.2 – Literature review

Ethnic differences in PA have been explored primarily in American populations through cross sectional studies (see Table 1.11). He et al.'s study examined variations in PA patterns by race/ethnicity and education as a measure of SES (91). LTPA, occupational PA and TPA were measured using in-home interviews and scored on a scale. TPA across all groups was similar, ranging from a score of 11.3 for Englishspeaking Hispanic women to 12.9 for English-speaking Hispanic men (on a scale of 0 -42). However, trends in LTPA and WRPA differed by ethnicity. White males and females reported the highest LTPA (6.9 and 6.1, respectively, on a scale of 0 - 12) while Spanish-speaking Hispanic males and females had the lowest levels (4.9 and 4.1, respectively). In contrast, Spanish-speaking Hispanics reported the highest WRPA (7.3 for males and 7.9 for females on a scale of 0 - 24) while WRPA was lowest in Whites (4.7 and 5.6 for males and females, respectively). Differences by SES were also noted. Males and females completing more than 12 years of formal education performed more LTPA than those with 12 or fewer years. As well, the highest levels of WRPA were found in individuals with the fewest years of education. The study found that education was a stronger determinant of PA, both leisure time and work related, than ethnicity (91). PA patterns in children and adolescents by ethnicity were examined in the National Longitudinal Study of Adolescent Health. Adolescents of Hispanic, Asian, non-Hispanic White and non-Hispanic black ethnicity used questionnaires to record their hours of inactivity (based on television/video watching and computer/video game playing) and times per week engaging in moderate to vigorous activity (103). Both sex and ethnic differences were noted in PA and physical inactivity. Males were more likely to participate in high levels of PA than females across all ethnic groups and the highest rates of activity were in White ethnic groups (35.2% participating in moderate to vigorous activity 5 or more times per week) compared to Hispanics, Asians and Blacks (29.8, 29.6 and 28.5%, respectively). Non-Hispanic whites also reported the lowest hours of physical inactivity per week (19.3 hours) when compared to Asian, Hispanic and non-Hispanic black groups (21.3, 22.2 and 29.7 hours, respectively). Males also spent more time engaged in physical inactivity (24.0 hours) than females (18.7 hours) (103).

Sex differences were also noted in Sallis et al.'s study in a high school population comprised of Latino, Anglo, African-American and Asian/Pacific Islander students. Participants completed a questionnaire that recorded their participation in specific activities and the number of times they engaged in exercise in different contexts (out of school, on sports teams, activity lessons, physical education classes and vigorous activity in physical education classes). Overall, males participated in activity for an average of 860 minutes a week compared to 489 minutes per weeks for females. Socioeconomic differences were also noted, as students of higher SES were more likely to participate vigorously in physical education, take physical education classes and join in activity lessons outside of school than low-SES students. Ethnic differences were noted in the activity lessons, physical education classes per week, vigorous activity in physical education classes and total frequency of vigorous activity (100).

Physical inactivity during leisure time and the influence of race/ethnicity and social class was examined in adults completing the third NHANES. Crespo et al. determined the prevalence of physical inactivity in Caucasian, African-American and Mexican-American ethnic groups by various social class indicators including education, income occupation, being above or below the poverty line, current employment status and marital status (86). Caucasians had the lowest age-adjusted prevalence of physical inactivity at 18% (95% CI: 17.0 – 20.8), while 35% of African Americans (95% CI: 30.4 – 36.4) and 40% of Mexican Americans (95% CI: 34.3 – 39.9) reported being inactive. Women had a higher prevalence of physical inactivity in all three ethnic groups. As well, the higher prevalence of physical inactivity among African Americans and Mexican Americans compared to Caucasians was evident across all but two categories by social class indicators, suggesting that social class cannot solely account for differences in inactivity (86).

Similarly, Marshall et al.'s study found ethnic differences in leisure time physical inactivity among non-Hispanic white, non-Hispanic black and Hispanic ethnic groups (104). Prevalence of inactivity was lower in males (9.9-20.9%) than females (12.0-27.3%) in all ethnic groups. Whites reported the lowest physical inactivity while

Hispanics had the highest prevalence of physical inactivity. However, this disparity among ethnic groups diminished when prevalence was adjusted for social class, suggesting that social class may be a moderator of the association between inactivity and ethnicity. Physical inactivity prevalence in all ethnic groups was highest among low social class groups (104).

Only one study exploring differences in PA by ethnicity in a Canadian population was found. Bryan et al. used data from the CCHS, a questionnaire administered every two years, to determine the prevalence of PA among different ethnic groups, including White, North American Aboriginal, Latin American, East/Southeast Asian, Black, West Asian/Arab and South Asian (85). The prevalence of moderate PA was highest in the White ethnic group (49%) while South Asians had the lowest prevalence at 34%. Males were more active than women across all ethnicities. While the study adjusted results by household income and education, it did not look specifically at SES as a determinant of PA, which may account for the discrepancy of PA among ethnic groups. Furthermore, the North American Aboriginal ethnic group was composed only of off-reserve Aboriginals, which may not be indicative of PA in Aboriginals living on a reserve (85). Results from the 1991 Aboriginals Survey found that differences in health status and access to health care existed between Aboriginals living on reserves and those living offreserve in urban and rural areas (105). Aboriginals living on reserves were more likely to self-report their health as "fair or poor" and less likely to have seen a physician in the

70

past year compared to those living off-reserve. These noted differences in on and offreserve Aboriginals may influence participation in PA.

Author, Year	Population	Ν	Study Design	Outcome
Bryan et al., 2006	Canadians aged 20-64 of White, North American Aboriginal, Latin American, East/Southeast Asian, Black, West Asian/Arab, South Asian and "Other" ethnicity	171 513	Cross sectional	Prevalence of PA differs by ethnicity: White (49%), Other (48%), North American Aboriginal (47%), Latin American (40%), East/Southeast Asian (39%), Black (38%), West Asian/Arab (36%) and South Asian (34%).
Sallis et al., 1996	California high school students of Latino, Anglo, African American and Asian/Pacific Islander ethnicity	1871	Cross sectional	Males participated in more activity than females (860 vs. 489 minutes per weeks). Students of higher SES were more likely to participate vigorously in physical education, take physical education classes and join in activity lessons outside of school than low-SES students. Ethnic differences were noted in the activity lessons, physical education classes per week, vigorous activity in physical education classes and total frequency of vigorous activity.
Gordon- Larsen et al., 1999	American adolescents of non-Hispanic white, non- Hispanic black, Hispanic	13157	Longitudinal	Physical inactivity was highest among males and in non-Hispanic black ethnic groups (males and females) and lowest in non-Hispanic white groups (males and

Table 1.11: Studies exploring variations in physical activity by ethnicity, socioeconomic status and sex

	and Asian ethnicity			females). Sex differences in physical activity were evident with males being more active than females. Ethnic differences were also present, though these were greater among females compared to males, with non-Hispanic white females participating in the most vigorous activity and Asian females the least.
Crespo et al., 2000	American adults of Caucasian, African American and Mexican American ethnicity	18885	Cross sectional	Prevalence of physical inactivity differed by ethnicity: Caucasians (18%, 95% CI: 17.0-20.8), African-Americans (35%, 95% CI: 30.4-36.4) and Mexican- Americans (40%, 95% CI: 34.3-39.9). Physical inactivity was higher in women than men across all ethnic groups and African-Americans and Mexican- Americans reported higher inactivity than men across all but two social class indicators
Sternfeld et al., 1999	American females aged 20-65 of African American, Asian, Hispanic and White ethnicity	2636	Cross sectional	PA patterns differed by ethnicity and SES. The highest quartile of sports/exercise and active living was associated with white ethnicity and higher education (graduated college). Hispanic ethnicity was associated with the highest quartile of household/caregiving activity (OR: 1.58, 95% CI: 1.10-2.28) and having

				a high school education or less was associated with the highest quartile of occupational activity (OR: 2.45, 95% CI: 1.86-3.21).
Marshall et al., 2007	American men and women of non-Hispanic white, non-Hispanic black and Hispanic ethnicity	11211	Cross sectional	Inactivity differed by ethnicity and SES. Inactivity prevalence was lower in males (9.9-20.9%) than females (12.0-27.3%) in all ethnic groups. Whites reported the lowest physical inactivity while Hispanics reported the highest inactivity. In all ethnic groups, inactivity was more prevalent for low SES.
He and Baker, 2005	American men and women aged 51-60 of non-Hispanic white, non- Hispanic black, English- speaking Hispanic and Spanish-speaking Hispanic ethnicity	9621	Cross sectional	Ethnic and education differences in leisure time PA and WRPA were evident. Whites were most likely to perform LTPA but least likely to perform WRPA. In contrast, Spanish-speaking Hispanics reported the most WRPA but the least LTPA. LTPA declined with fewer years of education while WRPA increased.

CHAPTER TWO: STUDY DESIGN

This chapter outlines the objectives and study designs of SHARE, SHARE-AP and SHARE-AP Action. It also describes the physical activity questionnaires that were used in these studies to measure physical activity. This chapter also includes a discussion of the threats to validity in this thesis.

2.1 – SHARE and SHARE-AP overview

SHARE and SHARE-AP were cross sectional studies conducted in a multiethnic Canadian sample to explore the determinants of CVD. Similarities and differences in the prevalence of CVD, atherosclerosis, conventional risk factors (glucose, smoking, hypertension, body fat, body weight) and novel risk factors (homocysteine, fibrinogen, plasminogen activator inhibitor 1 (PAI-1), lipoprotein (a) (Lp(a))) were examined across four ethnic groups. SHARE included individuals of European, South Asian and Chinese ethnicity while Aboriginal Canadians comprised the SHARE-AP study population.

2.1.1 – Recruitment and study population

SHARE

Individuals were recruited from three Canadian cities: Edmonton, Hamilton and Toronto. These cities were selected because they had considerable Chinese and South Asian populations and facilities (carotid ultrasound machines and technicians) that were required to conduct measurements reflecting atherosclerosis. Surname analysis, a previously validated method, was used to establish the sampling frame (106). A comprehensive list of South Asian and Chinese surnames was compiled from community-specific telephone directories and group lists in addition to existing lists of surnames. Public telephone directories were then merged with the list of surnames (107) and the South Asian and Chinese samples were selected using probability systematic random sampling. Participants of European ethnicity were selected from similar geographic regions as South Asian and Chinese participants, and this was established through matching by postal codes. When an individual of South Asian or Chinese ethnicity enrolled in SHARE, a list of individuals with the same postal code was generated. Surnames not identified as South Asian and Chinese were presumed to be European. In addition to matching by geographic region, participants were also sex and age (within 5 years) matched to South Asian and Chinese individuals.

Potential study participants were initially contacted by personalized letters for recruitment and then subsequently contacted by telephone. Up to twelve telephone calls were made and the times of calls were varied to maximize the likelihood of contact with potential participants. In instances where more than one eligible individual was identified in a selected household, the individual with the earliest birthday in terms of the calendar year was selected for enrolment. Eligible individuals completing the telephone interview but declining to enrol in SHARE were considered eligible non-responders. Information collected during the interview, including demographic, linguistic and socioeconomic information, was analyzed to determine non-respondent bias. Individuals agreeing to participate were asked to attend a clinic visit at their closest study centre (University of Alberta Hospitals, Edmonton; Hamilton General Hospital, Hamilton; and St. Michael's Hospital, Toronto) and were sent a Food Frequency Questionnaire (FFQ) and confirmation letter in advance of the clinic visit.

Participants were aged 35-75 and must have resided in Canada for a minimum of five years to ensure sufficient exposure and adaptation to a Canadian lifestyle. While surnames were used to identify potential participants of an ethnic group, ethnicity was confirmed by origin of ancestry. Participants were considered to be of European ethnicity if both parents and grandparents were from Canada, the United States or Europe. South Asians were identified as such if both parents and grandparents were from India, Pakistan, Sri Lanka or Bangladesh, and individuals were considered to be Chinese if both parents and grandparents originated from China, Taiwan or Hong Kong. Individuals were excluded from participation if they had a history of cancer or other chronic diseases (e.g. renal or liver failure) as these may influence determinants of CVD. However, participants with clinical vascular disease were not excluded. The final SHARE population was comprised of 326 European, 317 Chinese and 342 South Asian Canadians.

SHARE-AP

Participants in SHARE-AP were recruited from the Six Nations reserve in Ontario. A comprehensive list of Six Nations Band members was used to identify potential participants and contact information was obtained from telephone directories. Similar to the recruitment procedure used in SHARE, participants were randomly selected and contacted first by mail and then by telephone for recruitment. Participants agreeing to enrol were invited for a clinic visit at the Gan Yohs Community Health Centre. A total of 301 participants were enrolled. Individuals included in the study were between the ages of 35 and 75 and lived on the reserve for a minimum of five years. The exclusion criteria for SHARE-AP and SHARE were the same.

2.1.2 - Clinic visit and measurements

Participants for both SHARE and SHARE-AP were asked to fast for a minimum of 8 hours prior to their clinic visit. Written informed consent was obtained at the beginning of the visit, which included consent to DNA analysis and record linkage. The following was completed at the clinic visit (see Table 2.1):

- Fasting blood samples
- Blood and urine samples two hours after ingestion of a 75 g glucose load (in nondiabetes participants)
- General Questionnaire (self-administered)
- Physical measurements (see Table 2.2)
- B-mode carotid ultrasound
- Echocardiogram (SHARE only)
- Resting 12-lead electrocardiogram

- Nutritional assessment: Food Frequency Questionnaire (FFQ) and 7 day food recall (self-administered) and teaching participants how to fill out food records
- SHARE Well Being Questionnaire (self-administered)

Table 2.1: Data collected in SHARE and SHARE-AP

Measurement Instrument	Outcomes Measured
General Questionnaire	Ethnic group, socioeconomic status, physical activity, alcohol consumption, smoking, medical history, health care utilization and menopause
Physical Examination	Systolic and diastolic blood pressure, ankle blood pressure, waist and hip circumference, height and weight
Food Frequency Questionnaire	Macronutrient and micronutrient intake, foods consumed, portion size, diet habits, and use of vitamin and mineral supplements
Well Being Questionnaire	Job stress, general well being, coping, social support, hostility/anger and equality
Laboratory Analyses	Total cholesterol (TC), HDL, low density lipoprotein (LDL), triglycerides, glucose, Lp(a), Apolipoprotein B (ApoB), homocysteine, PAI-1, fibrinogen
Carotid Ultrasound	Intima Media Thickness of carotid arteries
Echocardiography	Left Ventricular (LV) Mass, LV dimensions, ejection fractions
Electrocardiogram	Prior myocardial damage and LV hypertrophy

Table 2.2: Physical measurements taken during the clinic visit

Measurement	Description
Arm and Ankle blood pressure	Measured using a standard mercury sphygmomanometer from the upper part of the right arm and the right ankle. Participants were to sit quietly for a minimum of five minutes. Two measurements, one at the beginning and at the end of the interview, were averaged.
Height	Participants were measured in bare feet while standing against a wall. Measurements using a tape measure were rounded to the nearest 0.5cm.
Weight	Participants were weighed in light clothing on a platform scale and measurements were rounded to the nearest 200g.
Waist to Hip Ratio	Calculated by dividing the waist circumference measurement by the hip circumference measurement. Waist measurements were taken at the smallest diameter between the costal margin and iliac crest over light clothing. Hip measurements were taken at the widest diameter of the buttocks over light clothing. Hip and waist measurements were taken twice.

Laboratory Measurements

Fasting blood samples were frozen at -70° C and central analysis of the samples was conducted in the core laboratory at the Hamilton General Hospital. Enzymatic methods were used to measure TC, glucose and triglycerides (108, 109). The Friedewald formula was used to calculate LDL cholesterol (110). HDL-C was measured by adding phosphotungstic acid and magnesium chloride to the sample to precipitate very low density lipoprotein and LDL (111). Lp(a) was measured with automated immunoprecipitation and turbidimetric detection on the Ciba-Corning 550 Express analyzer (Oberlin, OH) using the Incstar Lp(a) test kit (Stillwater, MN). The rate nephelometic method was employed to measure ApoB, using the Beckman Array Protein System and reagents from Beckman Instruments (Brea, CA). Hemoglobin (Hb) was measured using the Coulter Counter. The 765 Glycomat analyzer was used to measure hemoglobin A1C (HbA1C) which separates different subtypes of hemoglobin by lowpressure cation exchange chromatography and gradient elution.

2.1.3 – Physical activity measurement

Self-reported PA was measured using the modified Atherosclerosis Risk in Communities/Baecke Index (ARIC/Baecke). The index measures habitual PA in three contexts: WRPA, sports related PA and LTPA (112) (see Table 2.5). The Baecke Questionnaire was modified for the ARIC study through the addition of questions related to work and sports related PA. The original Baecke Questionnaire included only one question to describe the participant's occupation and the ARIC/Baecke has seven questions related to work, including tasks performed while working, the type of occupation and their employer (40). As well, participants completing the ARIC/Baecke listed and described a maximum of four sports activities while the original questionnaire only permitted listing a maximum of two activities (112). The ARIC/Baecke was shortened slightly to be incorporated into the General Questionnaire (GQ) used for SHARE and SHARE-AP. The ARIC/Baecke allows participants to list and describe up to four of the sports or activities they participate in most frequently while the GQ asked participants to describe two sports or activities. As well, a question pertaining to the participant's perception of their LTPA relative to others of the same age was omitted from the GQ (113).

The PAQ used in SHARE and SHARE-AP was scored according to the original scoring scheme for the Baecke Questionnaire. Most questionnaire items were scored on a 5-point Likert scale (Never (1) – Seldom (2) – Sometimes (3) – Often (4) – Always/Very Often (5)). The index score for each type of PA (work, sport and leisure) was calculated by taking the mean of all questions answered and the total index score was obtained by summing the work, sport and leisure indices. Participants were classified as participating in low, moderate or high levels of PA by tertiles.

Work Index

The Work Index was computed using the following formula:

$$[13 + (6 - 34a) + 34b + 34c + 34d + 34e + 34f + 35] / 8$$

Participants filled out their occupation for question 13 while questions 34 and 35 were answered on a Likert scale. Occupation was ranked according to intensity (low = 1, medium = 3, high = 5). Occupations assigned low intensity values included clerical work, teaching, administrators, physicians and lawyers. Occupations assigned medium intensity values included carpentry, cooks, factory work, registered nurses and plumbers. Occupations assigned high intensity values included construction work, farmers, heavy equipment mechanics, janitors and waiters/waitresses. Individuals not specifying an occupation were given a score of 1.

Sport Index

Questions 37a and 37b were answered using a Likert scale. For question 36, participants indicated the two activities that they performed most frequently and then described the activity performed, the number of times per week they performed the activity and the number of months per year they performed the activity. Activities were assigned an intensity value (light = 0.76 MJ/h for sports less than 3 METS, moderate = 1.26 MJ/h for sports between 3 and 5 METs inclusive and vigorous =1.76 MJ/h for sports greater than 5 METs) based on MET values taken from the Compendium of Physical Activities Tracking Guide (114). The intensity value was multiplied by time values, proportion values and a factor of 1.25. Time and proportion values were both converted to a scale (see Table 2.3) (105):

Table 2.3:	Time and	proportion	values	for c	question	36	in the	modified	ARIC	Baecke
Questionna	ire									

Time		Proportion					
On Questionnaire	Value	On Questionnaire	Value				
<1 hour	0.5	<1 month	0.04				
1-2 hours	1.5	1-3 months	0.17				
2-3 hours	2.5	4-6 months	0.42				
3-4 hours	3.5	7-9 months	0.67				
>4 hours	4.5	>9 months	0.92				

83

This process was performed for both activities indicated. These scores were summed and converted to a score from 1 to 5 (see Table 2.4).

 Table 2.4:
 Calculation of the simple sports score for the modified ARIC/Baecke

Questionnaire

On Questionnaire	Value
0	1
0.01-<4	2
4-<8	3
8-<12	4
>12	5

The Sport Index was computed using the following formula (40, 112):

[(36a + 36b) + 37a + 37b] / 3.

Leisure Index

The Leisure Index was computed using the following formula (all items were answered on a Likert scale):

$$[(6-37c)+37d+37e]/3$$

Table 2.5: Scoring of the modified ARIC/Baecke Questionnaire used in SHARE and

SHARE-AP

Work Index		
Question Number	Question	Scoring
13	What was/is your usual occupation?	1-3-5
34a	At work I sit:	1 - 2 - 3 - 4 - 5
34b	At work I stand:	1-2-3-4-5
34c	At work I walk:	1-2-3-4-5
34d	At work I lift heavy loads:	1 - 2 - 3 - 4 - 5
34e	At work I am physically tired:	1 - 2 - 3 - 4 - 5
34f	At work I sweat:	1 - 2 - 3 - 4 - 5
35	In comparison with others my own age, I think my work is physically?	1 - 2 - 3 - 4 - 5
Sport Index		
36a	Most frequently played sport/exercise?	Sport Intensity (0.76 - 1.26 - 1.76) x Time $(0.5 - 1.5 - 2.5 - 3.5 - 4.5)$ x Proportion $(0.04 - 0.17 - 0.42 - 0.67 - 0.92)$
36b	Second most frequently played sport/exercise?	Sport Intensity (0.76 - 1.26 - 1.76) x Time $(0.5 - 1.5 - 2.5 - 3.5 - 4.5)$ x Proportion $(0.04 - 0.17 - 0.42 - 0.67 - 0.92)$
37a	During your leisure time do you play sports or exercise?	1 - 2 - 3 - 4 - 5
---------	--	-------------------
37Ь	During your leisure time do you sweat?	1 - 2 - 3 - 4 - 5
Leisure	Index	
37c	During your leisure time do you watch television?	1 - 2 - 3 - 4 - 5
37d	During your leisure time do you walk outside?	1 - 2 - 3 - 4 - 5
37e	During your leisure time do you do household chores?	1-2-3-4-5

2.2 – SHARE-AP Action overview

SHARE-AP Action was an open RCT exploring the effectiveness of a behavioural and lifestyle intervention intended to alter PA and dietary intake in an Aboriginal population. Families on the Six Nations reserve were randomized to either the intervention group or the control group. The intervention included a behavioural and lifestyle modification program while families in the control group received information on healthy lifestyles. The intervention was administered for a period of six months and differences between the two groups were ascertained for the main outcomes: daily caloric intake (kcal per day) and PA (minutes per week).

2.2.1 – Recruitment and study population

Six hundred individuals that had either participated in SHARE-AP or the heart health screening after SHARE-AP was finished or others interested were approached for enrolment in SHARE-AP Action. Recruitment strategies also included advertising the study through different mediums on the Six Nations reserve, including local papers, the local weekly health radio show, information nights and through the Six Nations Health Services centre. To be eligible for SHARE-AP Action, households must have included a male and female parent with a minimum of one child aged 5–17 and all participants aged 5–65 were eligible to participate. Exclusion criteria for households included:

• Unwillingness to be visited weekly by a SHARE-AP Action Health Counsellor

- An anticipated absence from the reserve for a period greater than one month during the course of the intervention
- An anticipated break-up of the household within the next year

Exclusion criteria for individuals in a household were:

- Having a medical illness preventing the implementation of dietary and physical activity changes
- Having terminal cancer
- Suspected severe alcohol abuse
- Experienced a recent MI or stroke in the past month

Households eligible for enrolment were assigned to the control or intervention group using an automated 24-hour randomization service at the Hamilton Project Office. A toll free number was dialled and once data regarding entry into the study was given, households were then randomized to one of the groups.

Intervention Group

The intervention was comprised of health education/messaging and supplying goods to intervention households. Participants in the intervention group were visited by a health counsellor on a weekly basis to receive health information and set goals for reducing caloric intake and increasing PA. The health counsellors, who had experience in health promotion, were members of the Six Nations community. They were trained formally at the beginning of the study to provide health education to participants. Household members were evaluated individually at the beginning of the study by the health counsellor for their dietary and PA habits. Programs were then tailored and developed for participants to decrease their caloric intake and increase their PA, taking the age of the participant into account. The purpose of the weekly visits was to keep participants motivated and reinforce the intervention.

The Dietary Reference Intakes for macronutrients recommends that 45% of energy be derived from carbohydrates, 30% from fat and 25% from protein, and was used as a goal for participants (115). Changes in dietary intake were promoted by the health counsellor by showing participants how to read nutritional labels, giving them traditional Six Nations recipes and encouraging them to attend food preparation classes.

The goal of the PA program was to increase PA to at least 150 minutes per week in all individuals. The health counsellor developed an individualized program for each participant to help them achieve this goal, taking their schedule and existing PA patterns into consideration. Children and adolescents were encouraged to attend an after-school activity program held three times a week. Varied activities, including soccer, nature walks and bicycling, were organized and run by the health counsellors. As incentive to attend the program, children that attended a minimum of 80% of the sessions were given a bicycle and helmet.

One of the concerns with the Six Nations population was the high consumption of fruit juices and pop, which comprises 8% of their caloric intake. This may be exacerbated by the belief of some individuals on the reserve that their tap water is not fit

for consumption. Consequently, intervention households received a water cooler, a pack of 24 bottles of spring water and two 18L containers of spring water each week to encourage the consumption of water. As well, the health counsellor promoted drinking more water and low fat milk and less pop and fruit juice.

In addition to the primary outcomes of changes in energy consumption and PA, changes in knowledge, attitudes and behaviour (KAB) towards food and PA were explored as secondary outcomes. These were measured using a modified version of the Pathways' KAB questionnaire (116). Social Learning Theory outlines concepts that are important in influencing behavioural changes, and the randomization at the family level and design of the SHARE-AP Action trial incorporated these ideas. In a family setting, parents can act as role models for children through observational learning. By modeling behaviours, such as engaging in PA or eating healthily, the behaviour of children can be influenced by parents. Reinforcement of health behaviours can also occur within the family unit. Parents and children can reciprocally encourage and motivate each other to continue engaging in positive health behaviours. In addition, self-efficacy is important in changing lifestyle habits. Participants may gain confidence in their ability to perform specific habits when they are able to observe changes. Weekly visits with health counsellors, making dietary changes and increasing their PA can build confidence and affect change (117, 118).

Control Group

Households in the control group received a 30-minute introductory session and written materials including Canada's Physical Activity Guide to Healthy Active Living (119) and Canada's Food Guide to Healthy Eating (120). During the baseline clinic visits, participants were given Health Canada's Healthy Weight Chart and their BMI and WHR measurements were taken but they did not receive an individualized program. Children in the control group were allowed to participate in the after-school program with children from the intervention group. This was done to alleviate feelings that those in the intervention group were receiving better treatment relative to those in the control group.

2.2.2 – Clinic visits and measurements

Participants were asked to attend baseline, 3 month and 6 month assessments and the measurements taken at each visit are summarized in Table 2.6. Blood samples were taken for analysis at the core lab at the Population Health Research Institute. Blood was analyzed for fasting glucose and lipids through the use of standardized assays. Physical measurements were also taken and the methods to measure weight, height, BMI, hip and waist circumference are described in Table 2.7. In addition, bioelectrical impedance was used to determine the body composition of participants (FM, lean mass and total body water).

Table 2.6: Visit map for SHARE-AP Action

Baseline Visit •	Study Explained			
•	Consent			
•	Baseline Blood Sample			
•	Randomization			
•	24-hour nutrition recall			
•	KAB/Self-efficiency questionnaire			
•	IPAQ/Modified IPAQ/24-hour physical activity recall			
•	Physical measurements (Table 2.4)			
•	RT3 Accelerometer			
3-Month Visit	24-hour nutrition recall			
•	KAB/Self-efficiency questionnaire			
•	 IPAQ/Modified IPAQ/24-hour physical activity recall 			
•	Repeat physical measurements			
•	Repeat Blood Sample			
	Exit Survey			
Final 6-Month Visit	24-hour nutrition recall			
•	KAB/Self-efficiency questionnaire			
•	IPAQ/Modified IPAQ/24-hour physical activity			
	recall			
•	 RT3 Accelerometer 			
•	• Repeat physical measurements			
•	Repeat Blood Sample			
	Exit Survey			

Table 2.7: Physical measurements taken in SHARE-AP Action

Measurement	Description
Blood pressure	Measured using a standard mercury sphygmomanometer from the upper part of the right arm and participants were to sit quietly for a minimum of five minutes. Two measurements, one at the beginning and one at the end of the interview, were averaged.
Height	Participants were measured in bare feet while standing against a wall. Measurements were taken using a tape measure and measurements were rounded to the nearest 0.5cm.
Weight	Participants were weighed in light clothing on a platform scale and measurements were rounded to the nearest 200g.
Waist to Hip Ratio	Calculated by dividing the waist circumference measurement by the hip circumference measurement. Waist measurements were taken at the smallest diameter between the costal margin and iliac crest over light clothing. Hip measurements were taken at the widest diameter of the buttocks over light clothing. Hip and waist measurements were taken twice.
Body Mass Index	Calculated by dividing weight (in kilograms) by the height squared (in square metres)

The primary outcomes of SHARE-AP Action were daily energy intake and weekly PA (Table 2.8). Participants completed 24-hour nutrition recalls to measure energy intake. One of the advantages of this method, compared to other methods such as FFQs, is that it is sensitive to short-term changes in diet and suitable for the 6-month duration of the intervention (121). ESHA Food Processor Nutrient Analysis software was used to analyze nutritional data. Its database includes a basic nutrient composition database, the Canadian Nutrient File and the United States Department of Agricultural Food Book. The nutritional composition of Canadian fast foods and Aboriginal foods were also added to the database, and this information was taken from 7-day food records from the SHARE-AP dietary validation study (122).

Table 2.8: Outcomes for SHARE-AP Action

Primary Outcomes (Change from baseline to final)	 Daily energy intake (kcal per day) Weekly physical activity (minutes per week)
Secondary Outcomes (Change from baseline to final)	 Knowledge and attitudes of healthy lifestyles Self-efficacy Body fat percentage Body mass index Blood pressure Lipids Glucose

2.2.3 – Physical activity measurements

PA was measured using the IPAQ in adults (>18 years of age), the modified IPAQ in adolescents (11-18 years of age) and the 24-hour physical activity recall in children (5-10 years of age). The IPAQ was developed to facilitate the measurement of PA in different countries using the same instrument. Multiple versions of the IPAQ exist that vary by format (long or short format), administration (telephone or self) and time frame (past 7 days or a typical week). The activity domains covered in the IPAQ include work related, transportation related, household related and leisure time physical activity and sitting (45). The long version of the IPAQ was administered to adults in SHARE-AP Action. The modified IPAQ completed by adolescent participants was shortened, but still included the same domains present in the long IPAQ.

PA in adolescents was computed by summing PA from Section A (Physical Activity) and Section C (Work Related Physical Activity) of the GQ. Estimates of PA were determined by multiplying the MET value of the activity by the duration of the activity and the frequency (days/week) of the activity. The Compendium of Physical Activities Tracking Guide was used to obtain MET values for the activities in Section A. For Section C, moderate activity was assigned a MET value of 4.0 and vigorous activity was assigned a MET value of 8.0.

For adults, the Guidelines for Data Processing and Analysis of the IPAQ were used to determine PA as measured by the IPAQ (123). The duration of activity was converted into minutes from hours and minutes and any duration of activity less than 10 minutes in length was recoded to 0. The rationale for this is that activity should be performed for a minimum of 10 minutes to derive a health benefit (123). As well, values for any given activity in excess of 180 minutes were truncated to 180 minutes. This is in accordance with the instructions presented in the Guidelines. Any participants with TPA in excess of 960 minutes were excluded and were considered outliers. PA was then calculated by multiplying the activity MET value by duration and frequency and summing these totals (123).

A 24-hour physical activity recall was used in children instead of a modified version of the IPAQ. The recall was based on the questionnaire used in the obesity

prevention Pathways study conducted in a population of American Indian children. The questionnaire followed the form of a checklist and included sports activities, television/video watching and playing video/computer games. Instead of asking children to report the duration and intensity of their activity, they were asked to specify how much activity was performed (none – a little – a lot) before, during and after school in the past 24 hours. This method was employed because young children often have trouble estimating the duration and intensity of their activity over long periods of time (124).

PA estimations from the PAQ for children were obtained by multiplying the MET value of an activity by the amount that they performed the activity and then summing the values for all activities (124). MET values for activities were obtained using The Compendium of Physical Activities Tracking Guide (114). For the duration of activity, a value of 0 was assigned to "none", 1 was assigned to "a little" and 2 was assigned to "a lot".

Participants also wore the RT3 accelerometer for 7 days as an objective measure of PA. The accelerometer yields an estimate of PA in either kcal or METs. In addition, a family PA log was kept for intervention households where leisure activity was to be recorded in minutes per day. These logs were reviewed during clinic visits with the health counsellor.

PA as measured by the RT3 was computed by converting the vector sum, which is measured in kcal, into MET-min/day. This was done by multiplying the kcal value by 70kg because the RT3 unit uses 70kg as the standard subject weight, and then dividing by

the participant's weight (in kilograms) (133). Participants were excluded from analysis if they had less than 3 days of activity recorded or if their percentage of zero values was higher than the 75th percentile of all participants (81.1%). 71 participants wore the RT3 at baseline. Of these, 2 had less than 3 days of recorded data and 11 were above the 75th percentile for zero values. One participant failed both criteria. This yielded a final group of 57 participants with RT3 data.

2.3 – Threats to validity

One of the concerns in epidemiologic research is whether the inferences or conclusions yielded by a study are real or valid. This is associated with the internal and external validity of the study. Internal validity explores the extent to which a study is accurate in its conclusions based on the study population. External validity is concerned with the generalizability of the inferences made by the study to a particular population (125).

Two sources that may threaten the validity of a study are bias and confounding. Bias is any systematic error that results in deviations from what is true and most forms of bias in epidemiologic studies can be categorized as either selection bias or information bias. Selection bias is discussed in detail in Section 3.2. Information bias involves incorrect or inaccurate information being collected. This may result from different methods of measuring information between groups or from recall bias, where participants who have the outcome are more likely to report the exposure because they may attribute the outcome to the exposure (125). Confounding occurs when a relationship between an exposure and outcome is actually attributable to a third factor, a confounding variable. The confounder is associated, but unrelated, with the exposure and is causally associated with the outcome (126).

2.4 - Limitations of cross sectional studies

In cross sectional studies, exposure and outcome measures are simultaneously determined. The prevalence of disease can be determined in those with and without exposure, and associations between the exposure and outcome can then be made. Cross sectional studies are ideal when determining the aetiology of a disease or when multiple outcomes are being explored (126). One of the major limitations of cross sectional studies is that cause and effect cannot be determined because the exposure and outcome are measured at the same time. While cross sectional studies can determine an association between exposure and outcome, interpretation of any relationship should be cautiously made. Ascertaining cause and effect can only be determined in prospective cohort studies or randomized controlled trials (126).

Both SHARE and SHARE-AP were cross sectional studies and the exposure, PA, and outcome measures, CVD risk factors, were measured at the same time. An individual may have low levels of PA and high levels of risk factors but it is uncertain whether low PA caused or was caused by the high levels of risk factors.

2.5 – Selection bias

Selection bias is present when there are systematic differences in those that are enrolled for a study compared to those who are not enrolled. When selection bias is present, the study population may not be representative of the larger population (52). In SHARE, the sample was randomly selected from a list of surnames and participants were contacted by telephone. Information from the non-responders who agreed to answer questions was collected, including smoking status, heart disease, cancer, employment and post-secondary education. A comparison of the responders (n = 1527) and non-responders (n = 985) showed that no significant differences existed between the groups in terms of smoking, cancer or heart disease. There were significant differences between the groups for employment (61% in non-responders vs. 70% in responders, P < 0.001) and post-secondary education (53% in non-responders vs. 39% in responders, P < 0.001) (127). Participants in SHARE-AP were randomly drawn from a comprehensive list of residents on the Six Nations reserve. Similar to SHARE, information from the non-responders was collected. No significant differences were found between those enrolling in SHARE-AP and the non-responders, suggesting that a representative sample was selected (122).

Another consequence of selection bias is that differences between groups from the baseline to final visit may be attributable to factors other than the intervention. Random allocation can reduce selection bias from occurring, minimizing the chance that significant differences exist between groups (52). Participants enrolled in SHARE-AP Action were randomized either to the lifestyle intervention group or the control group.

2.6 – Social desirability bias

Social desirability occurs when study participants unintentionally edit themselves and the information they divulge to conform to social norms. Social desirability is different from faking good, which occurs when a study participant deliberately lies or provides misinformation to make a good impression. When social desirability bias is present, it can threaten the validity of a study in two ways (46).

First, the data that was collected may not be accurate, particularly if it could not be verified by another measure (46). In SHARE and SHARE-AP, PAQs were the only measure of PA. These are subjective measures and study participants may unconsciously overestimate the time and frequency of PA while underestimating their sedentary behaviours, such as watching television. PA was also measured using a PAQ in SHARE-AP Action but participants also wore the RT3, an objective measure, to measure their activity over the course of a week. The second issue associated with social desirability is that it threatens the discriminant validity of PAQs (46). A PAQ is only a useful tool if it is able to distinguish study participants by different levels of PA. If all participants are reporting high levels of PA irrespective of their actual PA, the PAQ no longer has the ability to accurately classify study participants.

2.7 – Reverse causality bias

In a causal relationship, the exposure precedes the outcome. Reverse causality occurs when the direction of the association between an exposure and outcome is opposite to what is expected. In this case, the outcome influences or causes the exposure (128). Reverse causality is a concern in SHARE and SHARE-AP because the exposure and outcome measures are determined at the same time, which makes directionality difficult to ascertain. For example, it may be thought that low levels of PA contribute to obesity, but an equally valid explanation is that obesity may result in lower PA levels. Reverse causality is less of a concern in SHARE-AP Action because it is a RCT. For example, changes occurring in the body weight of intervention participants could be attributed to increases in PA.

2.8 - Self-reporting of behaviours in epidemiologic and intervention studies

Self-reporting is frequently used in epidemiologic and intervention studies to measure behaviours such as PA, food consumption and smoking. One of the advantages of self-reporting is that it is a practical and cost effective approach for population-based studies with large numbers of participants. However, this method is subject to error and bias, including social desirability and recall bias (129).

CHAPTER THREE: METHODS AND ANALYSIS

This chapter describes the methods and analyses performed in this thesis. All analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 15.0.

3.1 – Assessment of reliability and validity in SHARE, SHARE-AP and SHARE-AP Action

This analysis determines the reliability and validity of the physical activity questionnaires used in SHARE, SHARE-AP and SHARE-AP Action. A 1-month testretest period is used to determine reliability in the SHARE and SHARE-AP questionnaires while a 3-month test-retest period is used for SHARE-AP Action. Systolic and diastolic blood pressure are used as criterion measures for the questionnaires used in SHARE and SHARE-AP, and the RT3 accelerometer is used as the criterion measure for the SHARE-AP Action questionnaires.

3.1.1 – SHARE/SHARE-AP

Test-Retest Reliability

The test-retest reliability of the Baecke Questionnaire used in SHARE was evaluated by re-administering the PAQ to participants one month after they first completed the questionnaire. The reliability of the entire PAQ as well as the reliability of work related, leisure time and sports related PA scales was assessed by correlation of the test and re-test PA scores. Distributions of the data were assessed for normality using the Shapiro-Wilks normality test, which is the recommended test of normality when the sample size is less than 2000 (see Table 3.1). The distribution of the data for TPA was normal both at baseline and at the 1-month re-administration. Consequently, the Pearson correlation coefficient was used to assess reliability of TPA. LTPA at baseline, WRPA at re-administration and sports related PA at both baseline and re-administration were not normally distributed. The Spearman rank correlation coefficient was employed to assess the test-retest reliability of work related, leisure time and sports related PA. Data were also assessed for outliers and there were no outliers present.

Descriptive statistics of the variables used in the reliability analysis are presented in Table 3.2 and scatterplots of the data are shown in Figures 3.1 to 3.4. Table 3.3 summarizes the correlation coefficients, which ranged from 0.72 - 0.80. This represents a moderate to strong relationship between the baseline and re-test PA scores, and all correlations were significant (P < 0.05).

Variable	Shapiro-Wilks Statistic	df	P Value
Baseline (Visit 003)			
Total Physical Activity	0.99	89	0.88
Work Related Physical Activity†	0.97	64	0.12
Leisure Time Physical Activity	0.97	89	0.04*
Sports Related Physical Activity	0.95	89	0.00*
1-month Re-administration (Vis	sit 004)		
Total Physical Activity	0.98	89	0.37
Work Related Physical Activity†	0.96	64	0.03*
Leisure Time Physical Activity	0.98	89	0.08
Sports Related Physical Activity	0.96	89	0.01*

 Table 3.1: Assessment for normality of data using the Shapiro-Wilks test

df = Degrees of Freedom

† Excludes participants that do not work

*Statistically significant at the P < 0.05 level

Variable	Baseline	1-month Re-administration
Mean Total Physical Activity (SD)	7.37 (1.36)	7.65 (1.54)
Mean Work Related Physical Activity (SD)	1.97 (0.83)	2.00 (0.83)
Mean Leisure Time Physical Activity (SD)	3.23 (0.62)	3.21 (0.69)
Mean Sports Related Physical Activity (SD)	2.17 (0.80)	2.43 (0.84)

Table 3.2: Descriptive statistics of the variables included in the analysis

Table 3.3: Test-retest reliability of baseline and 1-month administrations of the

ARIC/Baecke Questionnaire

Variable	Ν	Correlation Coefficient	P Value
Total Physical Activity	89	0.79	0.00*
Work Related Physical Activity†	64	0.80	0.00*
Leisure Time Physical Activity	89	0.72	0.00*
Sports Related Physical Activity	89	0.75	0.00*

† Excludes participants that do not work

*Statistically significant at the P < 0.05 level















Figure 3.4: Scatterplot of baseline and 1-month administrations of sports related physical activity index scores

Criterion Validity

The criterion validity of the modified ARIC/Baecke used in SHARE and SHARE-AP was assessed by determining the relationship between the various PA index scores and SBP and DBP, which served as criterion measures. The descriptive statistics of the variables used are summarized in Table 3.4. The normality of the data was assessed using the Shapiro-Wilks test and none of the data were normally distributed (see Table 3.5). The Spearman rank correlation was used to determine the criterion validity of the PAQ.

The variables were analyzed for outliers and one participant had a SBP of 115 mmHg and a DBP of 180 mmHg and was consequently excluded from analysis. As well, 4 participants were missing SBP and DBP data and were excluded from analysis. Six participants were given a score of 1 on the leisure index because their score was below 1 to include them in the analysis. A value of 1 for the work index was also imputed for participants that did not work to include participants in the TPA analysis.

The relationship between the PA scales and both SBP and DBP were fairly weak (Table 3.6). Scatterplots of the analyses are shown in Figures 3.5 to 3.12. Criterion validity with SBP ranged from -0.07 to -0.17 and the correlations with DBP ranged from -0.07 to -0.16. All correlations were statistically significant at the P < 0.05 level.

Table 3.4: Descriptive statistics of the variables included in the criterion validity

analysis

PAQ	N	Mean(SD)
Total Physical Activity	1286	7.37(1.36)
Work Related Physical Activity [†]	866	2.17(0.86)
Leisure Time Physical Activity	1286	3.20(0.61)
Sports Physical Activity	1286	2.18(0.74)
Systolic Blood Pressure	1282	118.95(18.57)
Diastolic Blood Pressure	1282	72.80(11.14)

† Excludes participants that indicated in question 34 that they do not work

Table 3.5: Assessment of the normality of the variables included in the criterion validity

 analysis

Variable	Shapiro-Wilks Statistic	Df	P Value
Systolic Blood Pressure	0.96	1282	0.00*
Diastolic Blood Pressure	0.99	1282	0.00*
Total Physical Activity	1.00	1286	0.01*
Work Related Physical Activity†	0.94	866	0.00*
Leisure Time Physical Activity	0.97	1286	0.00*
Sports Related Physical Activity	0.96	1286	0.00*

df = Degrees of Freedom

† Excludes participants that indicated in question 34 that they do not work

*Statistically significant at the P < 0.05 level

Table 3.6: Criterion validity of the modified ARIC/Baecke Questionnaire using systolic

	Criterion Validity vs. Systolic Blood Pressure		Criterion Validity vs. Diastolic Blood Pressure	
	Correlation Coefficient	P value	Correlation Coefficient	P value
Total Physical Activity	-0.17	0.00*	-0.07	0.02*
Work Related Physical Activity	-0.13	0.00*	-0.07	0.04*
Leisure Time Physical Activity	-0.08	0.01*	-0.16	0.00*
Sports Related Physical Activity	-0.07	0.01*	-0.07	0.01*

and diastolic blood pressure as criterion measures

*Statistically significant at the P < 0.05 level



Figure 3.5: Criterion validity of total physical activity using systolic blood pressure as a criterion measure.





Figure 3.7: Criterion validity of leisure time physical activity using systolic blood pressure as a criterion measure.



Figure 3.8: Criterion validity of leisure time physical activity using diastolic blood pressure as a criterion measure.



Figure 3.9: Criterion validity of work related physical activity using systolic blood pressure as a criterion measure.





Figure 3.10: Criterion validity of work related physical activity using diastolic blood

pressure as a criterion measure.

Figure 3.11: Criterion validity of sports related physical activity using systolic blood pressure as a criterion measure.



Figure 3.12: Criterion validity of sports related physical activity using diastolic blood pressure as a criterion measure.



Sensitivity Analyses

Question 34 on the PAQ asked the participant to indicate whether they worked. Figure 3.13 summarizes participant responses and whether participants completed the work PA index. Initially, it was planned that if the participant indicated they did not work on question 34, they were excluded from the analysis assessing the criterion validity
of the WRPA index. However, because a large number of participants who either indicated that they did not work but then filled out the work PAQ or indicated they did work but then failed to complete the questionnaire, a sensitivity analysis was performed. The first analysis assessed the criterion validity of the work PA index when only those individuals that indicated that they worked according to question 34 were included. The second analysis included only those individuals that completed the work PA index, regardless of their answer for question 34. The results of the correlation analysis are presented in Table 3.7. The strength of the relationship was slightly weaker in the second analysis compared to the first. As well, while the correlations between the work PA index and the criterion measures were weak in the first analysis, neither correlation was statistically significant in the second analysis.

Figure 3.13: Summary of responses to question 34 on the modified ARIC/Baecke questionnaire



Table 3.7: Sensitivity analysis comparing criterion validity of the work related physical

 activity index

	Analysis 1	Analysis 2
	N = 866	N = 865
Correlation with Systolic Blood Pressure	-0.13	-0.03
P value	0.00*	0.45
Correlation with Diastolic Blood Pressure	-0.07	-0.05
P value	0.04*	0.17

*Statistically significant at the P < 0.05 level

A sensitivity analysis was also performed to determine whether excluding those participants that had been diagnosed with hypertension or had been medically treated for hypertension would have improved the criterion validity of the PAQ. The results of the criterion validity assessment against SBP are shown in Table 3.8, and Table 3.9 shows the assessment against DBP. The criterion validity when excluding participants either diagnosed with hypertension or taking medication for hypertension was either equal to or worse than the validity estimates when including the entire study population. **Table 3.8:** Assessment of criterion validity in three different populations using systolic

Population	Total Physical Activity	Work Related Physical Activity	Leisure Time Physical Activity	Sports Related Physical Activity
All participants N = 1283	-0.17*	-0.13*	-0.08*	-0.07*
Excluding participants diagnosed with hypertension N = 1032	-0.11*	-0.12*	-0.07*	-0.02
Excluding participants treated for hypertension with medication N = 1062	-0.11*	-0.11*	-0.06*	-0.03

blood pressure as a criterion

Table 3.9: Assessment of criterion validity in three different populations using diastolic

Sample	Total Physical Activity	Work Related Physical Activity	Leisure Time Physical Activity	Sports Related Physical Activity
All participants	-0.07*	-0.07*	-0.16*	-0.07*
N = 1283				
Excluding participants diagnosed with hypertension N = 1032	-0.06	0.05	-0.16*	-0.05
Excluding participants treated for hypertension with medication	-0.06	0.05	-0.15*	-0.06
N = 1062				

blood pressure as a criterion

3.1.2 – SHARE-AP Action

Test-retest Reliability

Results of the PAQs at baseline and 3 months were compared to determine the test-retest reliability. Although data at 6 months was also available, data at 3 months was used to assess reliability because PA was one of the targets of the intervention. It is likely that PA could have changed significantly after 6 months of intervention and that the PA data would not be suitable in assessing reliability. The distribution of the PAQ data was assessed and it was not normally distributed. Because the data deviated from a normal distribution, Spearman Rank Correlation was performed (Table 3.10). Descriptive statistics of the variables used are summarized in Table 3.11.

Results of the reliability analysis are shown in Table 3.12 and Figures 3.14 to 3.16. A moderate relationship existed between baseline and 3-month PAQ scores for adults (0.43, P = 0.00) and children (0.42, P = 0.01), and both analyses were statistically significant at the P < 0.05 level. There was virtually no correlation between baseline and 3-month scores in the adolescent group (0.06, P = 0.75) and this was not statistically significant. Table 3.13 also shows reliability estimates, dividing participants into control and intervention groups to determine whether there were greater changes in PA in the intervention group. While reliability in control adults was higher compared to intervention group adults, reliability in adolescents and children was higher in the intervention group compared to the control group. Based on these analyses, the IPAQ

126

and the 24-hour recall PAQs had better reliability than the modified PAQ used in adolescent participants.

Table 3.10: Assessment of normality for data included in test-retest reliability analyses

Age Group	Shapiro-Wilks Statistic	df	P Value
Baseline Physical Activity			
Adults	0.89	49	0.00*
Adolescents	0.64	30	0.00*
Children	0.77	37	0.00*
3-month Physical Activity			
Adults	0.92	49	0.00*
Adolescents	0.84	30	0.00*
Children	0.84	37	0.00*

df = Degrees of Freedom

*Statistically significant at the P < 0.05 level

Table 3.11:	Descriptive	statistics o	f variables	included in	test-retest	reliability	analyses
-------------	-------------	--------------	-------------	-------------	-------------	-------------	----------

Age Group	N	Baseline Mean PA(SD) in MET- min/week	3-month Mean PA(SD) in MET- min/week
Adults	49	4197.26(3184.89)	5055.79(4328.58)
Adolescents	30	6555.70(9041.17)	9376.88(8236.85)
Children [†]	37	135.96(132.08)	247.93(178.29)

†Measured in METs/day

Table 3.12: Test-retest reliability of baseline and 3-month administrations of the

 physical activity questionnaires

Age Group	Spearman Rank Coefficient	P value
Adult	0.43	0.00*
Adolescents	0.06	0.75
Children	0.42	0.01*

*Statistically significant at the P < 0.05 level

Table 3.13: Test-retest reliability of baseline and 3-month administrations by control

and intervention groups

Age Group	Control Group	Intervention Group	
Adult	0.56*	0.31	
Adolescents	-0.03	0.28	
Children	0.30	0.55*	



Figure 3.14: Test-retest reliability of the IPAQ in adults using baseline and 3-month administrations

Figure 3.15: Test-retest reliability of the modified IPAQ in adolescents using baseline



and 3-month administrations

Figure 3.16: Test-retest reliability of the 24-hour physical activity recall questionnaire in children using baseline and 3-month administrations



Criterion validity

The criterion validity of the PAQs used in SHARE-AP Action was assessed using the MET-min/day output produced by the RT3 as the criterion measure. Because three different PAQs were used, the properties of each were determined separately. Normality of the data being used was assessed and the Spearman's rank coefficient was used to determine criterion validity because the data was not normally distributed (Table 3.14). Summary statistics of the variables used are shown in Table 3.15.

Overall, the PAQs were found to have poor criterion validity with the RT3.

Figures 3.17 - 3.19 graphically present the analyses. The correlations ranged from -0.10

to -0.30 and none of these were statistically significant at the P < 0.05 level (Table 3.16).

Age Group	Shapiro-Wilks Statistic	df	P Value
Physical Activity Question	naire		
Adults	0.90	25	0.01*
Adolescents	0.77	15	0.00*
Children	0.77	25	0.00*
RT3			
Adults	0.94	25	0.16
Adolescents	0.81	15	0.01*
Children	0.94	25	0.15

Table 3.14: Normality tests of the data included in criterion validity analyses

df = Degrees of Freedom

Age Group	N	PAQ Baseline Mean PA(SD) in MET-min/week	RT3 Baseline Mean PA(SD) in MET-min/day
Adults		3962.74(3127.47)	192.02(111.25)
Adolescents		4923.83(4408.73)	295.78(215.89)
Children [†]		130.93(109.08)	287.53(190.73)

Table 3.15: Descriptive statistics of variables included in criterion validity analyses

†Measured in METs/day

Table 3.16: Criterion validity of the physical activity questionnaires using the RT3

accelerometer as a criterion measure

Age Group	Spearman Rank Coefficient	P value
Adult	-0.10	0.65
Adolescents	-0.28	0.31
Children	-0.30	0.15



Figure 3.17: Criterion validity of the 24-hour physical activity recall questionnaire in children using the RT3 accelerometer as a criterion measure.



Figure 3.18: Criterion validity of the modified IPAQ in adolescents using the RT3

accelerometer as a criterion measure.

Figure 3.19: Criterion validity of the IPAQ in adults using the RT3 accelerometer as a criterion measure.



3.2 - Risk factor clustering in SHARE and SHARE-AP

This analysis explored trends in cardiovascular risk factors by physical activity level. Mean levels of continuous variables and proportions of categorical variables were compared by physical activity level. As well, means of continuous variables are adjusted by sex, ethnicity and age and compared by physical activity level. The crude and adjusted prevalence of continuous risk factors is also determined.

3.2.1 – Methods

Analysis of variance (ANOVA) was used to compare the means of the continuous variables for the three levels of PA. Participants were categorized by PA using the Baecke score. Low group was defined as a score ranging from 3.00 to 6.90, moderate PA was defined as a score from 6.91 to 8.00 and high PA was defined as a score from 8.01 to 15.00. A 2x3 chi square test was conducted to determine differences in the proportion of smokers between the three PA groups. For continuous variables where a significant difference (LSD) to determine which groups significantly differed from one another. Univariate analyses were conducted to obtain mean values adjusted by age, sex and ethnicity. Statistical significance was set at P < 0.05 (two-tailed) unless otherwise indicated. A two-tailed test was used instead of a one-tailed test because differences that were either greater or lesser than the mean, and not unidirectional differences, were being explored.

Binary logistic regression was used to obtain odds ratios adjusted by age, sex and ethnicity. Females were employed as the reference for sex comparisons, South Asians were the reference category for comparisons by ethnicity, and individuals 45 years of age and younger were the reference population for age comparisons.

Cutpoints for determining the prevalence of risk factors were set using current guidelines or standards for what denotes increased risk for CVD (Table 3.17). In the absence of established cutoffs for Apolipoprotein B/A1 ratio (ApoB/A1), individuals in the 75th percentile or higher were considered to have the risk factor. In both the crude and adjusted prevalence analyses, participants were considered to possess a risk factor if their value for a given variable was equal to or exceeded the cutpoint. To determine prevalence of a risk factor, the proportion of participants in each PA level that had the risk factor was calculated. Linear regression was used to calculate adjusted means for the adjusted prevalence analysis. The risk factor was the dependent variable and the independent variables were sex, ethnicity and age. Adjusted predicted values were generated and participants were reassessed to determine if they possessed a risk factor based on their new adjusted value. Chi square tests were performed to determine differences in the prevalence of risk factors between PA groups.

Table	3.17:	Definition	of	risk	factors

Risk Factor	Cutpoints/Definition
SBP	\geq 140 mmHg
DBP	\geq 90 mmHg
WHR	\geq 0.85 for women
	\geq 0.90 for men
BMI	≥ 25.0
ApoB/A1	\geq 0.94 (75 th percentile or higher)
Glucose	Fasting glucose \geq 5.7 mmol/L AND
	$HbA1c \ge 5.9\%$
Smoking	Current or former smoker

Missing data by variable is summarized in Table 3.18. Participants with missing data were excluded analysis by analysis instead of casewise exclusion. Casewise exclusion was not used because it would reduce the number of eligible participants included in analyses. Because multiple variables were analyzed, there is a greater likelihood that a participant may be excluded because they are missing data for a given variable.

Variable	Overall	Low PA	Moderate PA	High PA
	N = 1286	N = 437	N = 421	N = 428
ApoB/A1*	314	109	81	124
WHR	7	3 4		0
BMI	1	1	0	0
Smoking	3	3	0	0
SBP	3 2		1	0
DBP	3	2 1		0
HbA1c	8	1	6	1
Fasting Glucose	3	1	1	1

Table 3.18: Missing data

*Includes participants from SHARE-AP without ApoB/A1 data because ApoB and Apolipoprotein A1 were not measured.

3.2.2 – Analysis

Overall Trends

Overall baseline characteristics and PA level are presented for all participants in Table 3.19 and for non-diabetic participants in Table 3.20. Significant differences between groups were found for SBP, DBP, WHR and fasting glucose (P < 0.05). No significant differences were noted for comparisons between the three groups for BMI, ApoB/A1 and HbA1c (see Table 3.21). Post hoc analyses showed that the low PA group had significantly higher mean SBP, DBP, WHR and fasting glucose compared to the moderate and high PA groups (P < 0.05). Mean BMI and Hb1Ac were higher and ApoB/A1 was lower in the low PA group compared to the other PA groups, but these were not statistically significant. However, no significant differences were found when comparing the moderate and high PA groups for SBP (P = 0.79), DBP (P = 0.53), WHR (P = 0.24) and fasting glucose (P = 0.28). The difference in the proportion of smokers by PA level was also statistically significant (P < 0.05).

When diabetic participants were removed from the analysis (see Table 3.22), significant differences between PA levels were still found for SBP, DBP, WHR and smoking (P < 0.05), but no significant differences between groups were found for fasting glucose, although the trend of higher glucose in the low PA group remained (P = 0.08). Post hoc analyses showed that the low PA group differed significantly from both the moderate and high PA groups for SBP (both comparisons P = 0.00) and DBP (Moderate: P = 0.02; High: P = 0.00). The low PA group was also significantly different from the high PA group for WHR (P = 0.01), but not significantly different from the moderate PA group (P = 0.18). No significant differences were noted between moderate and high PA groups for SBP (P = 0.72), DBP (P = 0.42) and WHR (P = 0.17).

Characteristic	Overall	Low \mathbf{PA}^{\dagger}	Moderate PA ^{††}	High PA ^{†††}
N	1286	437	421	428
Mean Age (SD)	50.4 (10.3)	54.0 (11.1) 49.5 (9.8)		47.6 (8.8)
% Males	48.9	48.5	48.7	49.5
Ethnicity (%)				
European	25.3	19.0	28.5	28.7
South Asian	26.6	29.1	25.9	24.8
Chinese	24.7	28.4	27.6	18.0
Aboriginal	23.4	23.6	18.1	28.5

Table 3.19: Baseline characteristics including diabetic participant

[†] Low PA defined as PA score 3.00 – 6.90 ^{††} Moderate PA defined as PA score 6.91 – 8.00 ^{†††} High PA defined as PA score 8.00 – 15.00

Table 3.20: Baseline characteristics excluding diabetic participants

Characteristic	Overall	Low PA	Moderate PA	High PA
Ν	1147	365	386	396
Mean Age (SD)	49.5 (10.0)	52.7 (11.0) 49.0 (9.5		47.1 (8.5)
% Males	50.4	51.2	50.0	50.0
Ethnicity (%)				
European	26.9	20.3	29.8	30.3
South Asian	27.6	30.7	26.4	26.0
Chinese	26.7	32.3	28.8	19.4
Aboriginal	18.7	16.7	15.0	24.2

Tertiles of Physical Activity								
Low (SD)	Moderate (SD)	High (SD)	P value					
123.58 (20.75)	116.73 (16.95)	116.40 (16.79)	0.00*					
74.10 (12.80)	72.51 (11.13)	72.02 (10.42)	0.02*					
0.90 (0.09)	0.88 (0.09)	0.88 (0.09)	0.00*					
27.46 (5.64)	27.12 (5.42)	27.31 (5.50)	0.66					
0.76 (0.25)	0.80 (0.43)	0.79 (0.27)	0.34					
6.06 (2.30)	5.49 (1.44)	5.64 (1.93)	0.00*					
5.97 (1.31)	5.83 (1.02)	5.85 (1.09)	0.17					
178 (40.7%)	153 (36.3%)	195 (45.6%)	0.02*					
	Terti Low (SD) 123.58 (20.75) 74.10 (12.80) 0.90 (0.09) 27.46 (5.64) 0.76 (0.25) 6.06 (2.30) 5.97 (1.31) 178 (40.7%)	Tertiles of Physical Act Low (SD) Moderate (SD) 123.58 (20.75) 116.73 (16.95) 74.10 (12.80) 72.51 (11.13) 0.90 (0.09) 0.88 (0.09) 27.46 (5.64) 27.12 (5.42) 0.76 (0.25) 0.80 (0.43) 6.06 (2.30) 5.49 (1.44) 5.97 (1.31) 5.83 (1.02) 178 (40.7%) 153 (36.3%)	Tertiles of Physical Activity Low (SD) Moderate (SD) High (SD) 123.58 (20.75) 116.73 (16.95) 116.40 (16.79) 74.10 (12.80) 72.51 (11.13) 72.02 (10.42) 0.90 (0.09) 0.88 (0.09) 0.88 (0.09) 27.46 (5.64) 27.12 (5.42) 27.31 (5.50) 0.76 (0.25) 0.80 (0.43) 0.79 (0.27) 6.06 (2.30) 5.49 (1.44) 5.64 (1.93) 5.97 (1.31) 5.83 (1.02) 5.85 (1.09) 178 (40.7%) 153 (36.3%) 195 (45.6%)					

 Table 3.21:
 Mean values of continuous measures including diabetic participants

*Statistically significant at the P < 0.05 level **Number of Smokers

SD = Standard Deviation

Tertiles of Physical Activity							
Risk Factors	Low (SD)	Moderate (SD)	High (SD)	P value			
SBP (mmHg)	121.90 (19.95)	116.06 (16.60)	115.60 (16.54)	0.00*			
DBP (mmHg)	74.57 (12.97)	72.60 (11.09)	71.93 (10.51)	0.01*			
WHR	0.89 (0.09)	0.88 (0.09)	0.87 (0.09)	0.02*			
BMI (kg/m ²)	26.72 (5.04)	26.82 (5.20)	26.97 (5.35)	0.80			
ApoB/A1	0.76 (0.25)	0.79 (0.43)	0.79 (0.27)	0.46			
Fasting Glucose (mmol/L)	5.40 (1.12)	5.24 (0.79)	5.27 (1.09)	0.08			
HbA1c (%)	5.70 (0.89)	5.76 (0.92)	5.68 (0.80)	0.44			
Smokers** (Percent)	143 (39.2%)	134 (34.7%)	173 (43.7%)	0.03*			

Table 3.22:	Mean	values	of	continuous	measures	exclue	ling	diabet	ic par	ticipan	ts
--------------------	------	--------	----	------------	----------	--------	------	--------	--------	---------	----

*Statistically significant at the P < 0.05 level

**Number of Smokers

The crude and adjusted prevalence of risk factors in the three PA levels is presented in Tables 3.23 and 3.24. The crude prevalence of both high SBP and DBP differed significantly between groups (P < 0.05). Means were lowest in the high PA group while the highest means were in the low PA group. Significant differences between PA groups were also found for the prevalence of individuals with high WHR and dysglycemia. There were no significant differences between groups for the prevalence of elevated BMI or ApoB/A1. When prevalence rates were adjusted for by age, sex and ethnicity, only the prevalence of high WHR differed significantly between groups (P = 0.01). The prevalence of all other risk factors was not significantly different by PA group. None of the participants in the three PA groups had high DBP or ApoB/A1 after adjustment. Table 3.25 summarizes the crude and adjusted prevalence of participants with 3 or more risk factors by PA level. The highest crude prevalence of risk factors was in the low PA group (41.4%) and this was significantly different from the moderate and high PA groups (P = 0.00). When the prevalence was adjusted for by sex, ethnicity and age, the trend persisted. 60.6% of participants in the low PA group had 3 or more risk factors compared to 53.7% in the moderate PA group and 57.2% in the high PA group, but these differences were not statistically significant (P = 0.12).

Risk Factors	Low	Moderate	High	P value
SBP	22.4	10.0	9.3	0.00*
DBP	10.8	7.1	5.4	0.01*
WHR	58.6	52.3	47.7	0.00*
BMI	63.2	62.5	61.2	0.81
ApoB/A1	16.0	21.6	19.2	0.17
Fasting Glucose	53.8	49.2	43.5	0.01*

Table 3.23: Crude prevalence (%) of continuous risk factors by physical activity level

 Table 3.24: Adjusted prevalence (%) of continuous risk factors by physical activity

 level†

Risk Factors	Low	Moderate	High	P value
SBP	1.4	0.5	0.2	0.10
DBP	0.0	0.0	0.0	÷
WHR	68.1	59.1	59.1	0.01*
BMI	79.2	79.4	78.6	0.96
ApoB/A1	0.0	0.0	0.0	
Fasting Glucose	39.0	35.5	40.9	0.26

*Statistically significant at the P < 0.05 level †Adjusted for age, sex and ethnicity

Table 3.25: Crude and adjusted prevalence of participants with 3 or more risk factors

(%)

	Low	Moderate	High	P value
Crude Prevalence				
≥3 Risk Factors	41.4	30.2	29.4	0.00*
Adjusted Prevalence	e†			
≥3 Risk Factors	60.6	53.7	57.2	0.12

†Age, sex and ethnicity adjusted

Trends by Age

Tables 3.26 - 3.34 show the means for SBP, DBP, WHR, BMI, fasting glucose, HbA1c and ApoB/A1 adjusted by sex, age and ethnicity and Table 3.35 shows the adjusted OR for smoking. For all categories of PA, mean SBP, WHR, fasting glucose (excluding diabetic participants) and odds of smoking were significantly higher in the older age group (> 45 years old) compared to the younger age group (\leq 45 years old). Mean DBP was significantly higher in the older group than in the younger group (74.78 mmHg vs. 70.96 mmHg) for low levels of PA (Table 3.27).

When including diabetic participants in the analysis, fasting glucose was significantly higher in the older age group for both low (6.31 mmol/L vs. 5.60 mmol/L) and moderate PA (5.75 mmol/L vs. 5.19 mmol/L) (Table 3.29). As well, mean HbA1c was significantly higher in the older group for moderate PA when diabetic participants were included (5.96% vs. 5.68%) (Table 3.32). However, when diabetics were excluded from analysis, this difference was no longer significant (5.87% vs. 5.66%, P = 0.05) (Table 3.33).

Trends by Sex

Men had significantly higher mean SBP, DBP, WHR, BMI and fasting glucose (excluding diabetic participants) than females for low, moderate and high PA levels. Males also had higher odds of smoking than females across all PA groups (Low: 4.39; Moderate: 2.34; High: 3.18) (Table 3.35). When diabetics were included in analysis, mean fasting glucose was significantly higher for males than females for low (6.43 mmol/L vs. 5.74 mmol/L) and moderate PA groups (5.72 mmol/L vs. 5.40 mmol/L) (Table 3.30). No significant differences between males and females were found for HbA1c (including and excluding diabetic participants) and ApoB/A1 across all PA groups.

Trends by Ethnicity

No significant differences in mean SBP between ethnic groups were found for low, moderate and high PA levels (Table 3.26). However, Aboriginals had significantly lower mean DBP compared to European, Chinese and South Asian ethnic groups for all levels of PA (Table 3.27).

Mean WHR was significantly higher in Aboriginal compared to other ethnic groups for all levels of PA (Table 3.28). For the low PA group, mean WHR in Chinese participants was significantly lower than other ethnic groups. For moderate and high PA, Chinese and South Asian groups differed significantly (Moderate: 0.86 vs. 0.89; High: 0.85 vs. 0.87). When WHR was adjusted for sex, age and BMI, WHR was lower for low PA European and Aboriginal groups, for moderate PA Aboriginals and high PA Chinese compared to means adjusted only for sex and age. Mean WHR was slightly higher in Europeans and South Asians after adjustment that included BMI.

Across all PA groups, mean BMI was also significantly higher in Aboriginal participants compared to the other ethnic groups (Table 3.29). As well, Chinese

participants had significantly lower mean BMI in all pair-wise ethnic comparisons at all PA levels except against South Asians in the high PA group (24.02 kg/m² vs. 25.68 kg/m², P = 0.13).

When including diabetic participants in the analysis, mean fasting glucose was higher in Aboriginal participants in all pair-wise comparisons by ethnicity and PA level (Table 3.30). For low PA, fasting glucose was significantly lower in the Chinese ethnic group compared to the South Asian group (5.35 mmol/L vs. 6.08 mmol/L). When diabetics were excluded from analysis, mean fasting glucose in Aboriginals was still significantly higher than in other groups for moderate PA (Table 3.31). Aboriginals also had significantly higher mean glucose than Europeans in the high PA group (5.46 mmol/L vs. 5.05 mmol/L). In the low PA group, Europeans (5.13 mmol/L) had significantly lower glucose levels than Aboriginals (5.63 mmol/L) and South Asians (5.60 mmol/L) while South Asians had significantly higher fasting glucose compared to Chinese participants (5.60 mmol/L vs. 5.23 mmol/L).

Mean HbA1c was also significantly higher in Aboriginals compared to the other ethnic groups for all levels of PA when diabetic participants were included (Table 3.32). When diabetics were excluded from the analysis, Aboriginals had significantly higher mean HbA1c compared to the other ethnic groups only for the low PA group (Table 3.33). As well, mean HbA1c was significantly higher in Aboriginals compared to South Asians in the high PA group (5.89% vs. 5.58%). Aboriginal participants were not included in the comparison of adjusted ApoB/A1 means because ApoB was not measured in SHARE-AP (Table 3.34). No significant differences were found between groups for all levels of PA.

Aboriginals and Europeans had significantly higher adjusted odds of smoking for all levels of PA compared to the referent ethnic group, South Asians (Table 3.35). Chinese participants were also at significantly increased odds of smoking for low PA (OR = 2.07) but not at moderate or high levels (Moderate: OR = 0.74, P = 0.45; High: OR = 1.03, P = 0.94). Table 3.26: Age, sex and ethnicity adjusted means for systolic blood pressure (in

mmHg)

	Low		N	Ioderate	High		
	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	
Sex [‡]							
Male	211	125.50 (1.27)	204	120.57 (1.11)	212	119.49 (1.07)	
Female	220	121.71 (1.23)	215	113.16 (1.05)	216	113.49 (1.07)	
Ethnicity ^{‡‡}							
European	81	123.67 (2.01)	119	116.92 (1.42)	123	116.98 (1.40)	
South Asian	125	123.91 (1.62)	109	117.99 (1.48)	106	116.71 (1.51)	
Chinese	123	125.02 (1.65)	115	115.60 (1.46)	77	116.70 (1.77)	
Aboriginal	102	121.84 (1.84)	76	116.96 (1.80)	122	115.46 (1.41)	
Age ^{‡‡‡}							
<u><</u> 45	111	110.32 (1.35)	168	111.66 (1.39)	207	112.07 (1.00)	
>45	320	128.38 (1.12)	251	120.53 (1.04)	221	120.56 (1.20)	

Table 3.27: Age, sex and ethnicity adjusted means for diastolic blood pressure (in mmHg)

	Low		N	Ioderate	High		
	Ν	Mean (SE)	N	Mean (SE)	N	Mean (SE)	
Sex [‡]							
Male	211	77.27 (0.81)	204	75.90 (0.72)	212	75.89 (0.65)	
Female	220	70.37 (0.78)	215	68.34 (0.68)	216	68.72 (0.65)	
Ethnicity ^{‡‡}							
European	81	73.93 (1.28)	119	72.99 (0.92)	123	73.05 (0.85)	
South Asian	125	77.14 (1.03)	109	75.20 (0.96)	106	74.58 (0.92)	
Chinese	123	77.56 (1.05)	115	73.77 (0.95)	77	73.92 (1.08)	
Aboriginal	102	66.64 (1.17)	76	66.58 (1.17)	122	67.67 (0.86)	
Age ¹¹¹							
≤45	111	70.96 (1.37)	168	71.68 (1.05)	207	71.67 (0.71)	
>45	320	74.78 (0.60)	151	72.80 (0.60)	221	72.82 (0.63)	

	Low		M	loderate	High		
	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	
Sex [‡]							
Male	210	0.95 (0.01)	202	0.94 (0.01)	212	0.93 (0.01)	
Female	222	0.85 (0.01)	214	0.84 (0.01)	216	0.82 (0.01)	
Ethnicity ^{‡‡}							
European	81	0.89 (0.01)	118	0.87 (0.01)	123	0.85 (0.01)	
South Asian	125	0.89 (0.01)	108	0.89 (0.01)	106	0.87 (0.01)	
Chinese	123	0.86 (0.01)	114	0.86 (0.01)	77	0.85 (0.01)	
Aboriginal	101	0.96 (0.01)	76	0.94 (0.01)	122	0.92 (0.01)	
Ethnicity ^{‡‡‡}							
European	81	0.88 (0.01)	118	0.87 (0.01)	123	0.86 (0.01)	
South Asian	125	0.89 (0.01)	108	0.89 (0.01)	106	0.89 (0.01)	
Chinese	123	0.86 (0.01)	114	0.86 (0.01)	77	0.84 (0.01)	
Aboriginal	101	0.95 (0.01)	76	0.93 (0.01)	122	0.92 (0.01)	
Age ¹¹¹¹							
<u></u> ≤45	111	0.88 (0.01)	165	0.88 (0.01)	207	0.87 (0.01)	
>45	319	0.91 (0.01)	251	0.89 (0.01)	121	0.88 (0.01)	

Table 3.28: Age, sex and ethnicity adjusted means for waist to hip ratio

[‡] Ethnicity and age adjusted ^{‡‡} Sex and age adjusted ^{‡‡‡}Sex, age and BMI adjusted ^{‡‡‡‡} Ethnicity and sex adjusted

	Low		N	Ioderate	High		
	N	Mean (SE)	N Mean (SE)		N	Mean (SE)	
Sex [‡]							
Male	210	28.29 (0.34)	204	28.05 (0.33)	212	27.75 (0.33)	
Female	222	27.23 (0.33)	216	27.15 (0.31)	216	26.17 (0.33)	
Ethnicity ^{‡‡}							
European	81	28.43 (0.54)	120	27.57 (0.42)	123	26.85 (0.43)	
South Asian	125	26.36 (0.44)	109	26.46 (0.44)	106	25.68 (0.47)	
Chinese	123	24.07 (0.44)	115	23.65 (0.43)	77	24.02 (0.55)	
Aboriginal	103	32.17 (0.49)	76	32.72 (0.53)	122	31.29 (0.44)	
Age ^{‡‡‡}							
≤45	111	27.43 (0.45)	168	26.86 (0.42)	207	27.50 (0.38)	
>45	321	28.02 (0.28)	252	27.64 (0.30)	221	26.54 (0.30)	

Table 3.29:	Age, sex	and et	hnicity	adjusted	means	for b	body	mass	index	(in l	kg/m	1 ²)
--------------------	----------	--------	---------	----------	-------	-------	------	------	-------	-------	------	------------------

	Low		N	Ioderate	High		
	N	Mean (SE)	N Mean (SE)		Ν	Mean (SE)	
Sex [‡]							
Male	210	6.43 (0.15)	204	5.72 (0.10)	212	5.75 (0.13)	
Female	222	5.74 (0.15)	215	5.40 (0.09)	215	5.44 (0.13)	
Ethnicity ^{‡‡}							
European	81	5.55 (0.24)	120	5.24 (0.13)	123	5.19 (0.17)	
South Asian	126	6.08 (0.19)	109	5.61 (0.13)	105	5.47 (0.18)	
Chinese	123	5.35 (0.20)	114	5.20 (0.13)	77	5.26 (0.21)	
Aboriginal	102	7.37 (0.22)	76	6.20 (0.16)	122	6.48 (0.17)	
Age ¹¹¹							
≤45	111	5.60 (0.17)	168	5.19 (0.09)	206	5.41 (0.13)	
>45	321	6.31 (0.13)	251	5.75 (0.10)	221	5.74 (0.14)	

Table 3.30: Age, sex and ethnicity adjusted means for fasting glucose including diabetic participants (in mmol/L)

	Low		N	loderate	High		
	N	N Mean (SE)		N Mean (SE)		Mean (SE)	
Sex [‡]							
Male	185	5.69 (0.09)	192	5.42 (0.06)	198	5.50 (0.08)	
Female	178	5.11 (0.08)	192	5.15 (0.05)	197	5.06 (0.08)	
Ethnicity ^{‡‡}							
European	74	5.13 (0.12)	115	5.08 (0.07)	120	5.05 (0.10)	
South Asian	112	5.60 (0.10)	102	5.30 (0.07)	102	5.36 (0.11)	
Chinese	117	5.23 (0.10)	109	5.08 (0.07)	77	5.25 (0.12)	
Aboriginal	60	5.63 (0.15)	58	5.67 (0.10)	96	5.46 (0.11)	
Age ^{‡‡‡}							
≤45	109	5.18 (0.08)	163	5.14 (0.06)	197	5.11 (0.05)	
>45	154	5.48 (0.08)	221	5.39 (0.06)	198	5.42 (0.10)	

Table 3.31: Age, sex and ethnicity adjusted means for Fasting Glucose excluding diabetic participants (in mmol/L)

Table 3.32: Age, sex and ethnicity adjusted means for HbA1c including diabetic

participants (in %)

	Low		N	Ioderate	High		
	Ν	Mean (SE)	N	Mean (SE)	N	Mean (SE)	
Sex [‡]							
Male	210	6.04 (0.08)	200	5.83 (0.07)	212	5.83 (0.07)	
Female	222	5.92 (0.08)	214	5.90 (0.07)	215	5.83 (0.07)	
Ethnicity ^{‡‡}							
European	81	5.47 (0.13)	118	5.64 (0.09)	122	5.60 (0.09)	
South Asian	126	5.68 (0.10)	108	5.74 (0.10)	106	5.64 (0.10)	
Chinese	122	5.65 (0.11)	113	5.83 (0.10)	77	5.66 (0.12)	
Aboriginal	103	7.13 (0.12)	75	6.24 (0.12)	122	6.42 (0.09)	
Age ^{‡‡‡}							
≤45	111	5.97 (0.13)	166	5.68 (0.08)	206	5.74 (0.07)	
>45	321	5.98 (0.06)	148	5.96 (0.07)	221	5.88 (0.08)	
		Low	N	Ioderate		High	
-------------------------	-----	-------------	-----	-------------	-----	-------------	
	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	
Sex [‡]							
Male	185	5.80 (0.07)	189	5.72 (0.07)	198	5.71 (0.06)	
Female	178	5.68 (0.07)	192	5.83 (0.07)	197	5.65 (0.06)	
Ethnicity ^{‡‡}							
European	74	5.50 (0.10)	114	5.66 (0.09)	119	5.60 (0.07)	
South Asian	112	5.67 (0.08)	101	5.68 (0.09)	103	5.58 (0.08)	
Chinese	116	5.65 (0.08)	108	5.84 (0.09)	77	5.65 (0.09)	
Aboriginal	61	6.15 (0.12)	58	5.92 (0.13)	96	5.89 (0.08)	
Age ^{‡‡‡}							
≤45	109	5.78 (0.11)	161	5.66 (0.08)	197	5.63 (0.06)	
>45	154	5.71 (0.05)	220	5.87 (0.07)	198	5.71 (0.06)	

Table 3.33: Age, sex and ethnicity adjusted means for HbA1c excluding diabetic

participants (in %)

[‡] Ethnicity and age adjusted ^{‡‡} Sex and age adjusted ^{‡‡‡} Ethnicity and sex adjusted

	Low		M	Moderate		High
	N	Mean (SE)	Ν	Mean (SE)	N	Mean (SE)
Sex [‡]						
Male	174	0.77 (0.02)	172	0.81 (0.03)	149	0.79 (0.02)
Female	150	0.74 (0.02)	167	0.78 (0.03)	155	0.79 (0.02)
Ethnicity ^{‡‡}						
European	79	0.72 (0.03)	118	0.79 (0.04)	121	0.77 (0.03)
South Asian	126	0.79 (0.02)	108	0.86 (0.04)	106	0.82 (0.03)
Chinese	119	0.77 (0.02)	113	0.75 (0.04)	77	0.79 (0.03)
Aboriginal	N/A	N/A	N/A	N/A	N/A	N/A
Age ¹¹¹						
≤45	92	0.77 (0.03)	144	0.83 (0.05)	155	0.76 (0.02)
>45	232	0.75 (0.02)	195	0.79 (0.02)	149	0.82 (0.03)

Table 3.34:	Age, sex	and ethnicity	adjusted	means fo	or ApoB/A1
-------------	----------	---------------	----------	----------	------------

*No ApoB data was collected in SHARE AP [‡] Ethnicity and age adjusted ^{‡‡} Sex and age adjusted ^{‡‡} Ethnicity and sex adjusted

	L	ow	Mod	lerate	High	
	Odds Ratio	P value	Odds Ratio	P value	Odds Ratio	P value
Sex ^{‡**}						
Female	1.00	-	1.00	-	1.00	-
Male	4.39	0.00*	2.34	0.00*	3.18	0.00*
Ethnicity ^{‡‡}						
South Asian	1.00	-	1.00	-	1.00	-
European	5.95	0.00*	6.35	0.00*	4.25	0.00*
Chinese	2.07	0.02*	0.74	0.45	1.03	0.94
Aboriginal	28.26	0.00*	19.81	0.00*	12.80	0.00*
Age ^{‡‡‡}						
<u><</u> 45	1.00	=	1.00	-	1.00	-
>45	1.46	0.16	1.10	0.70	1.08	0.73

Table 3.35: Odds ratio of smoking adjusted by age, sex and ethnicity

*Statistically significant at the P < 0.05 level

**Odds of men smoking relative to women

[‡] Ethnicity and age adjusted ^{‡‡} Sex and age adjusted ^{‡‡} Ethnicity and sex adjusted

3.3 – Risk factor clustering and physical activity categorization in SHARE-AP Action

This analysis determines whether an association exists between risk factor score and physical activity level when assessing physical activity with a questionnaire and the RT3 accelerometer. As well, the agreement between physical activity questionnaires and the RT3 accelerometer in categorizing participants by physical activity is determined.

3.3.1 – Methods

Fisher's exact test was used to determine whether the association between PA level and risk factor score was stronger when measuring PA by the PAQ or when using the RT3, a criterion measure. As well, Cohen's Kappa coefficient was used to determine the agreement between PA levels as measured by the PAQ and RT3. The risk factors used to determine the presence of risk factor clustering were SBP, DBP, smoking, WHR and %BF. The definitions used for each risk factor are summarized in Table 3.36. Cut points for %BF were identical across the three age groups. Clinical definitions of systolic and diastolic hypertension were used for blood pressure cut offs (130). Smoking risk was denoted by being a current smoker in the child and adolescent groups and by being a current or former smoker in the adult group. Former smokers were not included in the child and adolescent groups because the questionnaire did not allow those individuals to indicate that they were former smokers. Presence of each risk factor was given 1 point to a maximum risk factor score of 5. For analysis purposes, individuals with no, one or two risk factors were considered not to have risk factor clustering while individuals with three or more risk factor s were considered to have risk factor clustering.

Risk Factor	Children (5 - 10	Adolescents (11 – 18	Adults (18 years >)
	years)	years)	
%BF	Females: ≥30%	Females: ≥30%	Females: ≥30%
	Males: $\geq 25\%$	Males: $\geq 25\%$	Males: ≥25%
DBP	90 mmHg	90 mmHg	90 mmHg
SBP	140 mmHg	140 mmHg	140 mmHg
Smoking	Current Smoker	Current Smoker	Current or Former
			Smoker
WHR	Females: ≥ 0.89	Females: ≥ 0.86	Females: ≥ 0.95
	Males: ≥ 0.92	Males: ≥ 0.96	Males: ≥ 1.00

Table 3.36: Risk factor definitions

WHR for adults was determined according to guidelines used by Yusuf et al. (32). For children and adolescents, the 75th percentile of WHR in each respective group was used to determine risk factor cut offs. One of the concerns with this approach was that the cut off used to denote risk was too high. However, comparisons of the mean WHR in the Aboriginal child and adolescent groups in SHARE-AP Action were compared to the means of European Caucasian child and adolescent populations in two studies (see Table 3.37) (131, 132). The means of the groups were comparable to those in the other study populations, which suggests that the cut offs for WHR that were used were appropriate.

Age Group	Mean WHR – Females	Mean WHR – Males (SD)
	(SD)	
SHARE-AP Action		
5-10	0.86 (0.07)	0.86 (0.08)
11-17	0.84 (0.07)	0.88 (0.10)
Gillum, 1999		
4-5	0.93	0.94
6-11	0.86	0.89
12-15	0.81	0.87
16-19	0.81	0.86
Taylor et al., 2000		
3-19	0.79 (0.01)	0.83 (0.01)

Table 3.37: Comparisons of mean waist to hip ratios in children and adolescents

Participants were classified by PA level into low, moderate and high groups for both the PAQ and accelerometer. Participants were first divided by age group (adults, adolescents and children) and then by tertiles (see Tables 3.38 and 3.39). Children used the 24-hour recall PAQ, adolescents completed a modified IPAQ and adults filled out the long form IPAQ. The study design chapter outlines how PA was calculated for each PAQ. **Table 3.38:** Definitions for low, moderate and high physical activity as assessed by

 physical activity questionnaires (in MET-minutes/week)

Age Groups				
Adults	Adolescents	Children*		
76	43	46		
0.00 - 2753.00	0.00 - 2318.33	0.00 - 61.33		
2753.01 - 7862.00	2318.34 - 3995.83	61.34 - 120.42		
7862.01 and above	3995.84 and above	120.43 and above		
	Adults 76 0.00 – 2753.00 2753.01 – 7862.00 7862.01 and above	Age Groups Adults Adolescents 76 43 0.00 - 2753.00 0.00 - 2318.33 2753.01 - 7862.00 2318.34 - 3995.83 7862.01 and above 3995.84 and above		

*Measured as METs/day

 Table 3.39: Definitions for low, moderate and high physical activity as assessed by the

RT3 accelerometer at baseline (in MET-min/day)

	Age Groups			
	Adults	Adolescents	Children	
N	26	13	18	
Low PA	0.00 - 152.66	0.00 - 221.05	0.00 - 220.60	
Moderate PA	152.67 - 230.84	221.06 - 301.72	220.61 - 437.11	
High PA	230.85 and above	301.73 and above	437.12 and above	

3.3.2 – Analysis

The distribution of risk factors is shown in Figure 3.20. A Pearson's Chi-square was originally going to be performed to determine whether an association between PA level and risk factor score existed because both variables were categorical. However, multiple cells contained fewer than 5 counts and Fisher's exact test was used instead.

Mean values for PAQ scores and RT3 data are presented in Table 3.40. Table 3.41 summarizes the results of the Fisher's exact test for the PAQ and RT3 data. The associations between PAQ or RT3 PA level and risk factor score were not significant (P = 0.24 for PAQ; P = 0.92 for RT3, two-tailed), suggesting that PA level and risk factor score are independent of one another. Cohen's Kappa coefficient was used to determine the agreement in PA level classification of participants by the PAQ and the RT3. There was weak agreement between the measures and this was not statistically significant ($\kappa = -0.18$, P = 0.07).

	Overall	Adults	Adolescents	Children
	N = 56	N = 25	N = 13	N = 18
Mean PAQ Score	N/A	5603.02	4188.27	134.66
in MET-min/week		(6530.76)	(3574.83)	(107.27)*
(SD)				
Low PA		10	4	4
Moderate PA		11	5	4
High PA		4	4	8
Mean RT3 in	283.06 (162.81)	208.73 (97.65)	335.23	348.62
MET-min/day			(204.01)	(167.82)
(SD)				
Low PA		9	4	6
Moderate PA		8	5	6
High PA		8	4	6
*Measured in METs	/day			

 Table 3.40: Descriptive statistics





Table 3.41: Fisher's Exact Test values

Method of PA	od of PA Value	
Measurement		
PAQ	12.00	0.24
RT3	5.56	0.92

3.4 - Physical activity trends in SHARE and SHARE-AP

This analysis explores trends in total, work related and leisure time physical activity by sex, ethnicity and socioeconomic status. As well, variables associated with low total physical activity are determined using logistic regression.

3.4.1 – Methods

Trends by sex, ethnicity and SES were explored in different contexts of PA. The Chi-square goodness of fit test was used to compare groups. SES was measured indirectly using the SDI. The SDI score was derived using the method outlined by Anand et al. (8). Participants with an income between \$20 000-\$60 000 were given 1 point, those with an income less than \$20 000 were given 2 points, unemployed participants were given 2 points and unmarried individuals were given 1 point. The maximum SDI score was 5. Participants were categorized into low, moderate and high PA based on their Baecke score for each type of PA. For TPA, WRPA and LTPA, participants were divided into three equal tertiles which represented low, moderate and high categories of PA (Table 3.42).

Categories of Physical Activity					
Low	Moderate	High			
0.00 - 6.90	6.91 - 8.00	8.01 - 15.00			
0.00 - 1.28	1.29 - 2.57	2.58 - 5.00			
0.00 - 3.00	3.01 - 3.33	3.34 - 5.00			
	Cate Low 0.00 - 6.90 0.00 - 1.28 0.00 - 3.00	Low Moderate 0.00 - 6.90 6.91 - 8.00 0.00 - 1.28 1.29 - 2.57 0.00 - 3.00 3.01 - 3.33			

Table 3.42: Definitions of low, moderate and high physical activity categories by

 context of physical activity

Multiple binary logistic regression was used to determine the predictors of low PA, the dependent variable used in modeling. PA was coded as a binary variable for the regression analyses as low PA or moderate/high PA. The variables assessed as possible predictors were age, sex, ethnicity and SDI score. Age was measured continuously while sex, ethnicity and SDI score were categorically measured. In the analyses, Europeans were chosen as the reference category for ethnicity, men were the reference for sex and a SDI score of 0 was the reference level for the other SDI scores. Variables were all entered into the model simultaneously. Main effects of the variables were initially determined and interactions were tested if the main effects were statistically significant (P < 0.05). The Hosmer-Lemeshow test was used to ensure that the model was valid. The test is a measure of goodness of fit and assesses whether the model prediction deviates significantly from the observed values. A significant test indicates that the model has poor fit (134). The Wald statistic was used to determine whether the predictor variables

168

were significantly associated with the dependent variable. The -2 log likelihood (-2LL) test was used to compare models.

3.4.2 – Results

The prevalence of total, work related and leisure time PA is presented by SDI score, ethnicity and sex in tables 2 through 10. The Chi-square goodness of fit test was used to compare groups within PA levels.

When looking at TPA, more individuals with an SDI score of 0 - 1 were in the high PA group compared to the other two levels of PA (P = 0.00). As well, 47.0% of the individuals with a score of 4 - 5 were in the low PA group compared to 20.1% in the high PA group (P = 0.00) (Table 3.43). It was interesting to note that in the analysis of leisure time PA, 48.6% of individuals with a score of 0 - 1 were in the low PA group compared to 28.7% that were in the high PA group (P = 0.00). There were similar numbers of individuals with scores of 4 - 5 in the low and high PA groups (P = 0.12) (Table 3.44). When exploring WRPA, 38.9% of participants with an SDI score of 0 - 1 were in the high PA group compared to 20.0% in the low group and this was a statistically significant difference (P = 0.00) (Table 3.45).

Physical Activity Level	Social Disadvantage Index Score						
	0 - 1 $N = 648$		2 -	2-3		4 – 5	
			N = 489		N = 149		
	%	N	%	N	%	N	
Low	28.7	186	37.0	181	47.0	70	
Moderate	33.6	218	31.5	154	32.9	49	
High	37.7	244	31.5	154	20.1	30	
X ²	7.	82	2.9	98	16.	12	
P Value	0.0)2*	0.2	23	0.0	0*	

Table 3.43: Prevalence of total physical activity by social disadvantage index score

*Statistically significant at the P < 0.05 level

 Table 3.44: Prevalence of leisure time physical activity by social disadvantage index

 score

score

	Social Disadvantage Index Score								
Physical Activity Level	0 -	- 1	2 -	2-3		5			
	N = 648		N = 489		N = 149				
	%	Ν	%	N	%	N			
Low	48.6	315	47.4	232	38.3	57			
Moderate	22.7	147	21.9	107	25.5	38			
High	28.7	186	30.7	150	36.2	54			
X ²	71.	.58	49	.49	4.2	20			
P Value	0.00*		0.00*		0.12				

		Social Disadvantage Index Score								
Physical Activity Level	0 ·	-1	2 -	. 3	4 -	5				
	N =	640	N =	149	N =	- 9				
	%	N	%	N	%	Ν				
Low	20.0	128	34.6	75	22.2.	2				
Moderate	41.1	263	25.3	55	11.1	1				
High	38.9	249	40.1	87	66.7	6				
X^2	52.	11	7.2	23	137	.57				
P Value	0.0	0*	0.0	3*	0.0	0*				

Table 3.45: Prevalence of work related physical activity by social disadvantage index

 score (excluding people indicating "I do not work" on question 34)

*Statistically significant at the P < 0.05 level

When TPA was explored by ethnicity, Europeans comprised the smallest proportion of participants in the low PA group (19.0%) while South Asians (29.1%) and Chinese (28.4%) made up the largest and this difference was statistically significant (P = 0.01). In the high PA group, there were significantly more European and Aboriginal individuals than Chinese individuals (P = 0.01) (Table 3.46). When ethnicity trends were explored in leisure time PA, Chinese participants comprised 32.0% of the low PA group while Aboriginals only comprised 20.4% and Europeans comprised 22.0% of the group (P = 0.00). 28.5% of participants in the high PA group were South Asian and 27.7% were European. In contrast, Chinese individuals only comprised 17.9% of the high PA group (P = 0.01) (Table 3.47). When looking at work related PA, 38.0% of the participants in the low PA group were European while only 1.0% were Aboriginal (P = 0.00). Conversely, only 19.0% of the participants in the high PA group were European, while 28.4% were South Asian and 31.6% were Aboriginal (P = 0.00) (Table 3.48).

Physical										
Activity Level	Europeans		South Asians		Chinese		Aboriginals		X ²	P value
	%	N	%	N	%	N	%	N		
Low	19.0	83	29.1	127	28.4	124	23.6	103	11.54	0.01*
(N = 437)										
Moderate (N = 421)	28.5	120	25.9	109	27.6	116	18.1	76	11.43	0.01*
High	28.7	123	24.8	106	18.0	77	28.5	122	12.92	0.01*
(N = 428)										

 Table 3.46: Prevalence of total physical activity by ethnicity

Physical	Ethnicity									
Activity Level	Euroj	peans	Sou Asi	1th ans	Chi	nese	Abori	ginals	X ²	P value
	%	N	%	N	%	N	%	N		
Low	22.0	133	25.7	155	32.0	193	20.4	123	19.13	0.00*
(N = 604)										
Moderate (N = 292)	29.1	85	26.0	76	18.5	54	26.4	77	7.26	0.06
High	27.7	108	28.5	111	17.9	70	25.9	101	10.88	0.01*
(N = 390)										

 Table 3.47: Prevalence of leisure time physical activity by ethnicity

*Statistically significant at the P < 0.05 level

 Table 3.48: Prevalence of work related physical activity by ethnicity (excluding people indicating "I do not work" on question 34)

Physical										
Activity Level	Europ	oeans	So Asi	uth ans	Chir	lese	Abori	ginals	X ²	P value
	%	N	%	N	%	N	%	N		
Low	38.0	78	28.3	58	32.7	67	1.0	2	67.02	0.00*
(N = 205)										
Moderate (N = 319)	27.0	86	32.3	103	26.3	84	14.4	46	21.78	0.00*
High	19.0	65	28.4	97	21.1	72	31.6	108	14.52	0.00*
(N = 342)										

When exploring sex trends in TPA, there were slightly more women for all 3 levels of PA but this comparison was not significant for low, moderate or high PA (P: 0.53 - 0.85) (Table 3.49). There were significantly more men in the low leisure time PA group than women (P = 0.00). In both the moderate and high PA groups, there were more women than men and this difference was significant (P = 0.00 for both PA levels) (Table 3.50). When looking at WRPA, there were significantly more women in the low PA group compared to men (P = 0.01). Conversely, there were more men in the high PA group but this difference was not significant (Table 3.51).

Physical Activity Level	Ma	lles	Females X ²		\mathbf{X}^2	P Value	
	%	N	%	N			
Low	48.5	212	51.5	225	0.39	0.53	
N = 437							
Moderate	48.7	205	51.3	216	0.29	0.59	
N = 421							
High	49.5	212	50.5	216	0.04	0.85	
N = 428							

Table 3.49: Prevalence of total physical activity by sex

Physical Activity Level	Ma	les	Females		X^2	P Value
	%	N	%	N		
Low	62.6	378	37.4	226	172.32	0.00*
N = 604						
Moderate	44.2	129	55.8	163	62.47	0.00*
N = 292						
High	31.3	122	68.7	268	71.63	0.00*
N = 390						

 Table 3.50:
 Prevalence of leisure time physical activity by sex

*Statistically significant at the P < 0.05 level

 Table 3.51: Prevalence of work related physical activity by sex (Excluding people indicating "I do not work" on question 34)

Physical Activity Level	Ma	les	es Females		X^2	P Value
	%	N	%	N		
Low	41.0	84	59.0	121	6.68	0.01*
N = 205						
Moderate	46.7	170	53.3	149	1.38	0.24
N = 319						
High	55.0	188	45.0	154	3.38	0.07
N = 342						

Logistic Modeling

Univariate binary regression was first performed to determine the main effects of age, sex, ethnicity and SDI using the Wald statistic. SDI scores were collapsed into two categories, low (SDI score = 0-2) and high (SDI score = 3-5). Each predictor variable was entered separately and the results are summarized in Table 3.52. European ethnicity, male sex and low SDI score served as referent categories. Chinese, South Asian and Aboriginal ethnicity, increasing age and high SDI score were significantly associated with low PA (P < 0.05).

Variable	Coefficient	Standard Error	Wald X ²	P value	Odds Ratio	95% CI
Ethnicity: Chinese vs. European	0.63	0.17	13.57	0.00*	1.88	1.34 to 2.63
Ethnicity: South Asian vs. European	0.55	0.17	10.46	0.00*	1.73	1.24 to 2.41
Ethnicity: Aboriginal vs. European	0.42	0.18	5.72	0.02*	1.52	1.08 to 2.15
Sex: Female vs. Male	0.02	0.12	0.04	0.84	1.02	0.81 to 1.29
Age	0.05	0.01	76.61	0.00*	1.05	1.04 to 1.07
SDI Score: High vs. Low	0.56	0.13	18.80	0.00*	1.75	1.36 to 2.25

Table 3.52: Univariate binary regression results

The variables were then fit into a multivariate binary logistic regression model (Table 3.53). Europeans were chosen as the reference category by ethnicity, males were the reference category for sex and low SDI score was the reference category for SDI. While most of the significant associations persisted, Aboriginal ethnicity, compared to European ethnicity, was no longer significantly associated with low PA (P = 0.23). The - 2 log likelihood value was 1518.78.

Interaction terms were added to the main effects model one at a time to determine whether they were significantly associated with low PA. Their improvement on model fit was assessed using the Wald statistic and by comparing the model to the main effects model using the likelihood ratio test. Table 3.54 summarizes the likelihood ratio tests for all models tested. Although SDI-ethnicity and sex-ethnicity interaction terms both improved model fit, only SDI-ethnicity interaction terms were significantly associated with low PA. As a result, only the SDI-ethnicity term was added to the main effects model.

Variable	Coefficient	Standard Error	Wald X ²	P value	Odds Ratio	95% CI
Intercept	-4.18	0.39	122.12	0.00*	0.02	-
Age	0.06	0.01	77.28	0.00*	1.06	1.05 to 1.07
Sex: Male [†]		-	-	-	1.00	-
Sex: Female vs. Male	0.04	0.13	0.11	0.74	1.04	0.82 to 1.33
Ethnicity: European†	-	-	-	-	1.00	-
Ethnicity: South Asian vs. European	0.74	0.18	16.98	0.00*	2.10	1.48 to 2.99
Ethnicity: Chinese vs. European	0.92	0.18	25.05	0.00*	2.51	1.75 to 3.60
Ethnicity: Aboriginal vs. European	0.23	0.19	1.44	0.23	1.26	0.86 to 1.84
SDI: Low†	5 4 1		1	-	1.00	-
SDI: High vs. Low	0.40	0.15	7.66	0.01*	1.50	1.12 to 1.99

 Table 3.53: Model assessing main effects of age, sex, ethnicity and social disadvantage

 index score as predictors of low physical activity

*Statistically significant at the P < 0.05 level

† Reference category

Interaction Term Added to Main Effects Model	-2 Log Likelihood	Likelihood Ratio Test†	P value
SDI-Ethnicity	1506.31	12.47	0.00*
SDI-Sex	1517.83	0.95	0.33
SDI-Age	1517.43	1.35	0.25
Sex-Ethnicity	1514.32	4.46	0.03*
Sex-Age	1517.51	1.27	0.26
Ethnicity-Age	1517.50	1.28	0.26
10 1 1 11 10			

Table 3.54: Assessment of model fit using likelihood ratio test

*Statistically significant at the P < 0.05 level df = 1

The final model is summarized in Table 3.55. The Hosmer-Lemeshow test statistic was 10.89 (P = 0.21), suggesting that the final model was valid and fitted the data well. Age was significantly associated with low PA (OR = 1.06, P = 0.00). As well, compared to those of European ethnicity, South Asian (OR = 2.17, P = 0.00) and Chinese participants (OR = 2.40, P = 0.00) were more likely to be classified as engaging in low levels of PA yes. Although a high SDI score was no longer significantly associated with low PA (OR = 1.12, P = 0.72), the interaction of Aboriginal ethnicity and a high SDI score was predictive of low PA (OR = 2.94, P = 0.01), suggesting that Aboriginals of low SES are sedentary compared to those of high SES.

Variable	Coefficient	Standard Error	Wald X ²	P value	Odds Ratio	95% CI
Age	0.05	0.01	70.60	0.00*	1.06	1.04 to 1.07
Sex: Male	-	-	-	-	1.00	-37
Sex: Female	0.04	0.13	0.09	0.76	1.04	0.81 to 1.33
Ethnicity: European	-	-	-	-	1.00	-
Ethnicity: South Asian	0.77	0.20	15.04	0.00*	2.17	1.47 to 3.20
Ethnicity: Chinese	0.88	0.20	18.42	0.00*	2.40	1.61 to 3.59
Ethnicity: Aboriginal	-0.32	0.27	1.38	0.24	0.73	0.42 to 1.24
SDI: Low†	-	-	-	-	1.00	
SDI: High vs. Low	0.12	0.33	0.13	0.72	1.12	0.59 to 2.13
European x SDI	-	-	11.95	0.01*	1.00	-
Chinese x SDI	0.15	0.43	0.12	0.73	1.16	0.50 to 2.72
South Asian x SDI	-0.23	0.44	0.28	0.60	0.79	0.34 to 1.87
Aboriginal x SDI	1.08	0.43	6.26	0.01*	2.94	1.26 to 6.83

Table 3.55: Summary of the final logistic regression model

CHAPTER FOUR: DISCUSSION AND CONCLUSION

This chapter reviews the major findings of the analyses and compares these results with the existing literature. As well, the strengths, limitations and implications of this thesis are described and future directions for this research are identified.

4.1 – Validity and reliability of physical activity assessment tools

4.1.1 – SHARE/SHARE-AP

The overall modified ARIC/Baecke PAQ and its subsections (WRPA, LTPA and sports related PA) demonstrated good reliability when baseline and 1-month administrations were compared. Spearman rank correlations ranged from 0.72 for LTPA to 0.80 for WRPA and all correlations were significant (P < 0.05). This is comparable to other assessments of reliability of the Baecke and the ARIC/Baecke PAQs. Other studies reported retest periods spanning from one to eleven months and found that test-retest correlations ranged from 0.71 to 0.93 (55, 112, 135, 136). This suggests that the modified ARIC/Baecke has reasonable reliability in the multiethnic population in which it was used.

In contrast, the criterion validity of the PAQ was quite poor. When the total PAQ score and subsection scores were correlated with SBP and DBP, which served as criterion measures, the associations were weak. While these results were statistically significant, this relationship was not clinically significant. It was anticipated that the PAQ would have a negative relationship with SBP and DBP because it would be expected that active

individuals would have lower BP and conversely that sedentary individuals would have higher BP. Correlation coefficients ranged from -0.17 to -0.07 for validity against SBP and from -0.16 to -0.07 against DBP as the criterion in the SHARE/SHARE-AP study population. Assessments of criterion validity of the ARIC/Baecke PAQ in other studies have varied depending on the measure used as the criterion. These measures have included the Caltrac accelerometer, PA histories or diaries, %BF and VO₂ max. In studies using the Caltrac accelerometer, validity ranged from 0.19 to 0.32 (112, 135, 137) while the validity in studies using PA records or diaries ranged from 0.33 to 0.59 (112, 136). Correlation with %BF ranged from -0.17 to -0.51 (112, 135) and correlation with VO₂ max ranged from 0.23 to 0.57 (112, 135).

One reason that may explain the poor construct validity of the ARIC/Baecke PAQ in this study is that SBP and DBP may not be ideal criterion measures for PA. SBP and DBP were utilised as construct measures because it was believed that they would be influenced by an individual's PA. However, the relationship between BP and PA may be affected by other variables such as age, diet and smoking that may act as effect modifiers (126). Exploring the association between PA and BP while stratifying by these factors may reveal if they are indeed effect modifiers.

As well, exploring the criterion validity of the PAQ against other constructs may help determine whether the poor association was a result of a poor choice for a construct measure or is indicative of shortcomings in the questionnaire. Some studies have found that the validity of different sections of the ARIC/Baecke can vary depending on the construct measure used. In Richardson et al.'s evaluation of the PAQ, LTPA correlated poorly with the Caltrac accelerometer (men: -0.05; women: 0.06) but was moderately correlated with the PA record (men: 0.37; women: 0.42) (112). It may be useful to assess the criterion validity of the ARIC/Baecke PAQ in this particular population with other measures. Because no measurements were taken in SHARE and SHARE-AP of variables considered to be gold standards for PA such as accelerometers or DLW, other measures including SBP and DBP believed to be associated with PA were used instead. Using a measure that attempts to directly measure PA, instead of one that may be influenced by PA, such as BP, may help determine the validity of the PAQ.

4.1.2 – SHARE-AP Action

The IPAQ used in adults and the 24-hour PA recall used in children had moderate correlations between baseline and 3-month administrations. The Spearman rank correlations were 0.43 and 0.42 for adults and children, respectively (P < 0.05 for both correlations). However, there was nearly no correlation for the modified IPAQ used in adolescents (Spearman rank correlation = 0.06, P = 0.75). Other assessments of the IPAQ's reliability have ranged from 0.46 to 0.93 when retest administrations took place within 10 days of initial administration (45). No reliability assessments evaluations of the PAQ used in the PATHWAYS study, which provided the basis for the PAQ used in children in SHARE-AP Action, could be found for comparative purposes. Sallis et al. evaluated the reliability of another 24-hour PA recall questionnaire, the Yesterday Activity Checklist, in a group of 4th grade students. The test-retest period was 4 days and

the correlation between the two administrations was 0.60, which is comparable to the reliability of the 24-hour recall used in SHARE-AP Action (138).

The weak reliability, particularly in adolescents, may be attributed to a few aspects of the study design. SHARE-AP Action was a RCT where one of the primary outcomes was change in PA. As a result, it is possible that there may have been real changes in PA that affected the assessment of reliability. It would be ideal to reassess the reliability of the PAQs in an observational or cross sectional study. As well, improved PAQ reliability may have been achieved with a shorter retest period. This study assessed the association between the initial and 3-month administrations of the PAOs, over which time there may have been actual PA changes. Selecting a shorter time frame such as a month may provide a more accurate measure of reliability. As well, the instructions in the IPAQ and 24-hour recall ask the participant to recall recent PA and not their average PA. In the case of the IPAQ and modified IPAQ, this was a period of 7 days. If a participant had been uncharacteristically active or inactive while completing either the initial or 3-month PAQs, the relationship between their PAQ measurements would be weakened. As well, participants may not have read the instructions correctly and instead recalled their average PA and not their most recent PA.

The criterion validity of all PAQs against the RT3 accelerometer was very poor. Spearman rank correlations ranged from -0.10 in adults to -0.30 in children and none of these associations were significant. It was surprising to find that the PAQs were negatively correlated with the RT3 because a positive correlation was expected. Participants that indicated they performed more PA would have been expected to also have higher RT3 measurements. A 12-country assessment of the IPAQ's criterion validity against the Computer Science and Application's Inc. (CSA) accelerometer found correlations ranging from -0.27 to 0.61 (45). A negative correlation may be the result of overestimation of PA by sedentary individuals, as was discussed with respect to the modified ARIC/Baecke PAQ used in SHARE and SHARE-AP.

In the Pathways study, the Tritrac accelerometer was used to assess the criterion validity of the 24 recall. The average correlation between the accelerometer and the PAQ for activities before and during school was 0.15 while the correlation for activities performed after school was 0.41. The authors of the study stated that the low correlation was not surprising since the recall PAQ was designed only to identify activities that children participated in and not to provide a comprehensive measurement of PA (124). Although accelerometers are frequently used in assessing validity, using other criterion measures such as DLW may be useful in ascertaining whether the IPAQ and the 24-hour recall can provide valid measurements of PA.

4.2 – Association between physical activity levels and risk factor clustering

When comparing mean values of CVD risk factors between different levels of PA in the SHARE and SHARE-AP data, SBP, DBP, WHR, fasting glucose and smoking were significantly different between low and high PA groups. The differences in SBP, DBP, WHR and smoking persisted even after diabetic participants were removed from analysis. BMI, ApoB/A1 and HbA1c levels were not significantly different between groups. While some studies also found significant differences in SBP between PA groups (139-141), others did not find the groups differed significantly (141-143). Similarly, there were mixed findings with respect to DBP between PA groups (139-144).

WHR differed significantly between groups but mean BMI was not found to be different between groups. While BMI was explored as a risk factor in multiple studies, relatively few included WHR in their analyses. In contrast to this study, many studies found that BMI was significantly different between low and high PA groups (140-142, 144, 145) and only two studies found there was no difference between groups (143, 146). Of the studies including WHR in their analysis, one found that there was a real difference between groups (144) and the other found there was no significant difference (142).

Other studies did not find a significant difference in fasting glucose between PA groups (141, 146), similar to this study's findings. However, Fransson et al. found there was a significant difference in the proportion of diabetics between individuals regularly performing PA and those that indicated they seldom or never engaged in PA (139). No studies could be found that used HbA1c as a risk factor in their analyses.

Although only one study could be found that explicitly looked at differences in ApoB/A1, many studies did look at mean LDL and HDL levels. Raitakari et al. did not find ApoB/A1 levels between groups differed significantly (146). Panagiotakos et al. found that ApoB and Apo A1 levels differed significantly in women but not in men (145). Multiple studies found that there were significant differences in HDL between PA groups (140-142, 144). LDL differences were not significant between PA groups in Aadahl et al. and Sofi et al.'s studies (142, 144), but Fransson et al. did find significant differences in mean LDL levels (139). Other studies supported the finding that the proportion of smokers in low and high PA groups differed significantly (139, 143, 145, 146), although this difference was not seen in men in Panagiotakos et al.'s study (145) and it was also not seen in women in Raitakari et al.'s study (146).

The variability in associations reported in the literature may be partially explained by different methods of assessing PA. An individual may be categorized differently by 2 PAQs, which would affect potential associations with risk factors. As well, there was variability in the composition of the populations in the studies that were found. None of the other studies included all of the ethnic groups that were looked at in this thesis and it is possible that trends may vary by ethnicity. For example, some ethnic groups may be more predisposed to CVD risk factors than other groups, even at higher levels of PA, which would affect the association between PA and risk factors.

When comparing risk factor clustering between PA groups, the analysis performed on SHARE and SHARE-AP data found that there was a significantly higher proportion of clustering in the low PA group compared to the high PA group. This trend was still apparent after adjustment by sex, age and ethnicity but it was no longer significant. The SHARE-AP Action analysis exploring the relationship between different PA measures and risk factor score found that the associations between the PAQ and the RT3 and the risk factor score were not significant. Other studies support the findings of the SHARE and SHARE-AP analyses. Ribeiro et al.'s study in a pediatric population found that in the lowest quartile of PA, 59.1% of male participants had biological risk factors while 40.9% did not. Comparatively in the highest quartile, 41.3% had risk factors and 58.7% had no risk factors. In females, 58.6% in the lowest quartile had risk factors while 41.4% did not, and in the highest quartile, there were equal numbers of participants with and without biological risk factors. The differences between males were significant but the differences between female participants were not (147). Similarly, Anderson et al. found a negative association between PA quintile and risk factor clustering using multinomial logistic regression. The study was performed in a pediatric population and the risk factors included were hypertension, high percent fat mass and high TC (148). As well, Raitakari et al. and Twisk et al. both found that physical inactivity, along with other lifestyle risk factors, was associated with clustering of biological risk factors (149, 150).

These findings contrast with the results of the SHARE-AP Action analysis. However, some variability in results is expected because different studies have compared different risk factors. Some risk factors may be more strongly associated with PA than others and hence, more likely to demonstrate clustering (147). A significant association between either or both of the PA measures used in SHARE-AP Action and risk factor score may have been found if different risk factors were selected. As well, the differences found in the SHARE and SHARE-AP analysis were not statistically

188

significant after adjustment, but they may have remained significant if different risk factors had been included or omitted.

As well, much of the existing literature has focused on adolescent and young adult populations. The SHARE and SHARE-AP populations were older while the SHARE-AP Action population included a broader age distribution that ranged from children to adults. It is feasible that clustering related to PA may be less obvious in an older population when the acquisition of CVD risk factors may also be related to normal aging. This may explain why the differences in SHARE and SHARE-AP were not significant and account for the lack of association between PA measures and risk factors in SHARE-AP Action.

4.3 – Physical activity patterns by sex, ethnicity and socioeconomic status

The results of this study showed that for TPA, there were more individuals with a low SDI score in the high PA group than the low PA group. For participants with a high SDI score, there were more participants classified as having low PA compared to those in the high PA group. The SDI findings have been supported by other studies that used education, family income and neighbourhood as surrogate measures of SES (151). South Asian and Chinese participants comprised the largest proportion of the low PA group while there were more Aboriginals and Europeans in the high PA group compared to the other ethnicities. The high TPA in Aboriginals may be attributed to their higher WRPA while high TPA in Europeans may be attributed to their high LTPA. The trend of higher PA in individuals of European ethnicity has also been seen in other studies (103, 151).

There were more females than males in each level of PA, but this difference was not significant. Contrary to the findings of this study, Gordon-Larsen et al. found higher levels of PA in men than women (103). When performing a logistic regression analysis, South Asian and Chinese ethnicity, age and high SDI score were found to be associated with low PA. As well, the interaction between Aboriginal ethnicity and high SDI score was significantly associated with low TPA.

When looking at WRPA, more individuals with a low SDI score were classified as high PA than low PA. The findings of this study with respect to SDI go against the findings of other studies, which have shown that those that are less educated are more likely to be active in their work (91, 92). Europeans formed the largest proportion of individuals performing low PA while Aboriginals were the largest group in the high PA group. Other studies have shown that Europeans and whites are less likely to be active in their work compared to other ethnic and racial groups (91, 92). This may be partially attributed to SES differences between ethnicities. More women than men were classified as low PA while more men were in the high PA group compared to women. He and Baker's study also found that men were more likely to be active in their work compared to women (91). One explanation for this is that men may be more likely than women to engage in physically demanding work, which would account for their higher WRPA. Fields of work such as construction and carpentry that are physically demanding have traditionally been male-dominated. In the combined SHARE and SHARE-AP data, 19.6% of women compared to 23.5% of men were in occupations that were considered to be moderately or highly active.

For LTPA, a greater proportion of individuals with a low SDI score were classified as low PA compared to high PA. Contrary to this study, other studies have found that those of higher SES are more likely to be active in their leisure time (85, 91, 104). The proportion of Chinese individuals was markedly higher in the low PA group compared to the other three ethnic groups. As well, there was a smaller proportion of Chinese individuals in the high PA group. Other studies have shown that whites and individuals of European ethnicity are more active in their leisure time compared to other racial and ethnic groups (91, 92, 104). Crespo et al. found this difference persisted even when SES was taken into account (86). Bryan et al. found similar findings of high PA among whites and Aboriginals and low PA in Asians (85). This study also found that there were more men than women in the low PA group while there were more women than men in the high PA group. Other studies have found that men are more active in their leisure time compared to women, which was not seen in this thesis (85, 104).

While many of the findings of this study were supported by existing literature, some trends were contradicted by other studies and went against hypothesized findings. One explanation to account for these differences is over-reporting of PA by different groups. This may be attributable to social desirability bias where participants may wish to exhibit traits that are seen as desirable, such as being physically active (85). As well, the PAQs used to measure PA may not have been comprehensive enough to fully capture the range of PA performed by participants. Consequently, some activities that individuals participate in may not have been accounted for in their PA scores, resulting in under-reporting of PA. In addition, while this study collectively looked at income, marital status and education when considering the relationship between SES and PA, other studies used different variables to classify participants by SES.

4.4 – Implications

This study showed that the PAQs used in the SHARE, SHARE-AP and SHARE-AP Action studies had good reliability, similar to what existing literature has already shown. However, the criterion validity of the PAQs was found to be poor against the criterion measures employed in this study. This suggests that the PAQs are not adequately capturing PA, that the criterion measures selected are not appropriate, or that the PAQ is not sensitive enough to pick up small changes in these outcomes (i.e. it is an indirect and rather crude measurement tool). Other studies have found that the validity of the PAQs to be moderately better using other criterion measures. This is significant because the PAQs may not be measuring PA in the selected populations. To determine whether the PAQs are able to accurately measure PA in this study population, further evaluations of criterion validity using other measures, such as PA diaries and VO₂ max, should be performed. Once this is done, the PAQs can then be used with greater confidence in this study. The results of the CVD risk factor analysis in SHARE and SHARE-AP suggest that there is an inverse relationship between increased PA and decreased risk factor clustering. In particular, there were significant differences in SBP, DBP, WHR, fasting glucose and smoking between individuals in the low and high groups while BMI, ApoB/A1 and HbA1c did not differ significantly. The finding that WHR differed significantly while BMI did not was interesting because they are both considered surrogate measures of body fat. Yusuf et al. found that WHR was significantly associated with risk of myocardial infarction and recommended that WHR be used as a measure of obesity instead of BMI (32). Although there is conflicting literature regarding which risk factors differ by PA level, the overall finding of risk factor clustering in low PA individuals has been supported by other studies. This finding emphasizes that PA does play a role in the development of risk factors for CVD and highlights the importance of encouraging individuals to be more active to prevent the development of disease.

Differences in PA were evident in all 3 domains of PA by sex, ethnicity and SES. Many of these findings were in accordance with other studies, although much of the existing research was conducted in an American population. Identifying differences in PA with respect to ethnicity, sex and SES is important in the development of effective health promotion strategies that are able to target groups of people that tend to be more sedentary. Identifying these at-risk groups is an important step in improving health at the population level. It was interesting to find that sex was not significantly associated with low TPA, yet increasing age, South Asian and Chinese ethnicity and low SES were
significantly associated. The reasons for these associations with PA need to be further explored and elucidated. While some research has already been conducted to explore these differences, such as understanding the causes of low PA in minority groups (90, 92, 94, 95), there is a need to look at the issues specific to each group within their unique cultural setting (i.e. findings among Black Africans in the US will likely not be the same explanatory factors among South Asians in Canada). As well, while this study looked at 4 ethnic groups from 4 communities in Canada, there is a need to expand beyond these and explore trends in other ethnic groups in Canada.

4.5 - Strengths

Multiethnic Population and Physical Activity

While other studies have explored patterns of PA, most of them have sampled a multiethnic American population. One of the strengths of this thesis was that it was conducted in a multiethnic Canadian population, focusing on European, South Asian, Chinese and Aboriginal ethnic groups. As well, the method of recruiting participants by ethnicity ensured that participants were not misclassified. Participants were first identified as potential participants by surname or location (in the case of people living on the Six Nations reserve) and ethnicity was then confirmed by self-report, and face-to-face clinic visit.

As well, other studies that have looked at Canadian populations have limited their scope of PA to only one context, such as LTPA or WRPA (85). In this study, patterns in

LTPA and WRPA were explored in addition to overall PA. As this study found, there are variations by sex, ethnicity and SES, and these trends vary depending on the type of PA being explored. Information about these differences in PA may have been obscured by only looking at PA broadly.

Interplay of Ethnicity, Sex and Socioeconomic Status

This thesis investigated PA variations by sex, ethnicity and SES. Differences in PA by these factors have been identified in the literature and it was important to look at which differences persisted when accounted for these factors. For example, the logistic models revealed that female sex was not significantly associated with low PA, which other studies have found, while both ethnicity and SES were associated (103). In particular, looking at both ethnicity and SES together was important because some studies have noted that differences in PA by ethnicity can sometimes be explained by differences in SES (92, 99, 100).

Furthermore, while other studies have often only employed one measure to serve as a surrogate measure of SES, this study used a composite score that took three aspects of SES into account: income, marital status and employment status. As well, the SDI index used was developed within the SHARE and SHARE-AP population combined social and economic measures, which are associated with prevalent CVD (101).

Risk Factors

Another strength of the study was the inclusion of multiple risk factors when exploring risk factor clustering by PA level. Many risk factors, such as smoking, hypertension and hyperlipidemia have been identified as being associated with an increased risk of developing CVD. While other studies have typically only included three to four risk factors in their analysis, this thesis included eight risk factors. Among these were WHR and HbA1c which have not been as extensively studied as other more conventional risk factors such as SBP, DBP, BMI and smoking. Understanding the relationship between these relatively novel risk factors and PA may help improve our understanding of PA's protective effect against the development of CVD.

4.6 – Limitations

Study Design

One of the limitations of this study is that no causal relationships can be definitely ascertained because of the study design. SHARE and SHARE-AP were both cross sectional studies and as a result, only associations can be determined since temporality cannot be established. The development of risk factors for CVD may not necessarily precede PA and vice versa. As well, SHARE-AP Action was a RCT and this design may not have been conducive to evaluating the criterion validity of the PAQs used because the interventions in the study were aiming to increase PA in participants, real changes may have occurred over the test-retest period that would affect the reliability analysis.

Issues in Self-Reporting

The PAQs in all study populations were self-administered, which may have affected PA measurement. If participants did not understand the wording of a question, they may have incorrectly filled out the question or left it blank because they were not able to receive clarification of the question. This would have resulted in under-reporting of PA. In addition, the PAQs required participants to recall PA from periods ranging from 24 hours to a week. Some studies have found that participant recall is poor, which may result in inaccurate measurement of PA. Furthermore, the issues associated with bias in self-reporting need to be acknowledged. Social desirability bias may have led to overreporting of PA. Some participants may perceive being physically active as a desirable attribute and as a result, they may have overestimated their PA. This would have resulted in the misclassification of individuals, which could affect the results of the analyses. Overestimation of self-reported PA may also explain some of the poor associations that were found in this study regarding reliability and validity. Other studies have addressed the issue of over-reporting of PA in sedentary individuals, which would attenuate a potential association (44). One approach to determining whether this occurred in this thesis is to re-evaluate criterion validity by stratifying by PA level and then comparing validity assessments in low PA groups compared to high PA groups (44).

Small Sample Size and Missing Data

One of the issues that may have affected the criterion validity analysis of the SHARE-AP Action PAQs and the associations between risk factor clustering and 197

different PA measures is the small sample of participants with RT3 data. Only 57 participants out of a possible 174 had baseline RT3 data recorded that met the data cleaning criteria. There were varied reasons for participants not having RT3 data. Some of the missing RT3 data was due to the accelerometer failing to record data or malfunctioning. However, other participants did not have RT3 data because they refused to wear the accelerometer.

Missing data can be categorized as missing completely at random, missing at random and missing not at random (152). Data is considered missing completely at random when the reasons for the data being absent are not related to the value of the missing variable or any other variables. Data missing at random is unrelated to the value of the missing variable, but may be related to other variables and data missing not at random is likely related to the value of the missing value (153). The data missing because of accelerometer issues may be considered missing at random because malfunctioning is likely not related to activity level. Some inactive participants may have refused to wear the RT3 because they did not want to appear sedentary. Conversely, active individuals may have chosen not to wear the accelerometer because they may have thought that wearing the accelerometer would be cumbersome and interfere with their daily activities.

Assumption of Independent Data

Participants from the SHARE-AP Action study were considered independent in analyses. However, because participants were enrolled and randomized as a family unit 198 in the trial, this assumption may not be valid because participants from the same family are not necessarily independent. Two participants from the same family may be more similar in terms of PA, for example, than two participants from different families. One alternate approach to analyzing this data is mixed or multilevel modelling. Mixed modelling takes variables at both the group and individual level into account and determines what their respective effects are on individual outcomes. As well, it allows variation at the group and individual level to be explored (154).

Multiple Statistical Testing

Although one of the strengths of this thesis was including multiple risk factors in the analysis of risk factor clustering by PA level, one of the issues associated with this approach is multiple statistical testing. As the number of statistical tests performed increases, the occurrence of a spurious significant finding also increases. For example, if the level of significance is set at P = 0.05 (as it was for the analyses in this thesis), it can be expected that one out of twenty analyses would be spuriously significant and not indicate a truly significant difference (155). One approach to minimize the chances of spuriously significant findings is employing a correction method, such as the Bonferroni criterion. According to the Bonferroni correction, alpha is divided by the number of analyses being performed and that number becomes the new alpha value. This requires P values to be lower to be considered statistically significant because the threshold for significant results is below the conventional 0.05 level (155).

4.7 - Future directions

Based on the findings of this thesis, there are several areas where further research is needed. Although the PAQs used in this study showed good reliability, their validity was poor against the selected criterion standards. The poor validity may be a consequence of the PAQ not accounting for all PA. Conversely, the criterion methods selected may not be the most ideal. The criterion validity of the PAQs needs to be reassessed using other measures including gold standards so that they can be used in future studies confidently. In addition, the reliability of the PAQs used in SHARE-AP Action should be re-evaluated in an observational study. It would be expected that the reliability would improve because change in PA was not being targeted by the intervention being administered.

As well, for the purposes of this study, cutpoints to divide participants into low, moderate and high levels of PA were based on tertiles of PA. Similar methods of dividing participants into tertiles, quartiles and quintiles of PA have been performed by other studies. However, it may be beneficial to develop cutpoints for each of the PAQ scales that are based on risk of developing CVD.

This thesis established that they are variations in TPA, LTPA and WRPA with respect to sex, ethnicity and SES. Further research needs to be done to determine the reasons for these variations and whether there are barriers to PA that may exist for some groups. Conducting qualitative studies may be the best avenue to address these questions. Such investigations may also reveal whether the PAQs are sufficiently comprehensive and include all relevant activities.

As well, having identified differences in PA among the ethnicities included in this study, it would be worthwhile to determine the existence of such patterns among other ethnic groups in Canada. This study could be expanded to include off-reserve Aboriginals, African, South American and Middle Eastern ethnic groups. As well, this study did not account for urban and rural differences and whether these may be associated with variations in PA. Exploring the influence of geographical location on PA in conjunction with sex, ethnicity and SES may show that it is associated with PA.

4.8 – Conclusions

The modified ARIC/Baecke PAQ and its sub scales were reliable instruments in the SHARE and SHARE-AP populations using a 1-month retest period. In SHARE-AP Action, the IPAQ used in adults and the 24-hour PA recall used in children both had good reliability. However, the modified IPAQ used in adolescents had poor reliability. The criterion validity of the ARIC/Baecke against SBP and DBP and of the PAQs used in SHARE-AP Action against the RT3 was poor, and re-evaluating these instruments against other measures may show better validity.

In the SHARE and SHARE-AP populations, mean levels of SBP, DBP, WHR and the prevalence of smoking differed significantly between low and high PA groups while no significant differences were found in mean BMI, ApoB/A1, fasting glucose and HbA1c. When comparing the crude prevalence of 3 or more risk factors, the prevalence was significantly higher in the low PA group compared to the high PA group, but this difference was no longer significant when adjusted for age, sex and ethnicity. Trends by sex, age and ethnicity were explored for each risk factor. Mean levels of SBP, DBP, WHR, fasting glucose, HbA1c and prevalence of smoking were higher in older individuals. SBP, DBP, WHR, BMI, fasting glucose, ApoB/A1 and smoking rates were higher in men. WHR, BMI, fasting glucose, HbA1c and smoking rates were higher in Aboriginals while DBP was lowest in Aboriginals. Mean ApoB/A1 was highest in South Asians. In SHARE-AP Action, there was no association between risk factor score and PA as measured by either the RT3 or the PAQ.

In SHARE and SHARE-AP, TPA was highest in Aboriginal, European, low SES and female participants, and it was lowest in South Asians, Chinese, high SES and male participants. Variables significantly associated with low TPA were increasing age, South Asian and Chinese ethnicity and the interaction between Aboriginal ethnicity and high SES. Looking at WRPA, low SES, Aboriginal and male individuals had higher PA while European and female participants had lower PA. Low SES, Chinese and male participants had lower LTPA while European, Aboriginal, South Asian and female individuals had higher PA.

4.9 - Main Findings

The main findings of this thesis were:

- The modified ARIC/Baecke had good reliability but poor validity in the SHARE study population. In the SHARE-AP Action population, the IPAQ in adults and 24-hour physical activity recall in children had moderate reliability, while the reliability of the modified IPAQ in adolescents was poor. In addition, all three questionnaires demonstrated poor validity against the RT3.
- Decreased physical activity was associated with increased risk factor clustering in the SHARE population. No such association was seen in the SHARE-AP Action study when measuring physical activity using a PAQ or the RT3 accelerometer. As well, there was no agreement in the classification of individuals by PA level between PAQs and the RT3.
- 3. High TPA was seen in Aboriginals, Europeans, individuals of low SES. Low TPA was associated with increasing age, South Asian and Chinese ethnicity and with Aboriginals of low SES. High WRPA was seen in Aboriginals, males and individuals of low SES. High LTPA was most prevalent in Europeans, Aboriginals, South Asians and females.

4.10 – Summary

This thesis explored PA and the use of various PAQs in a multiethnic Canadian population. The modified ARIC/Baecke was reliable when used in a population composed of European, South Asian, Chinese and Aboriginal Canadians, but had poor validity. In Aboriginals, the IPAQ in adults and the 24 hour PA recall in children were proven to be reliable questionnaires, but the validity of both against the RT3 accelerometer was poor. The reliability and validity of the modified IPAQ in adolescents needs to be reassessed in future research if it is to be used confidently in other studies. As well, this thesis identified variations in PA by sex, ethnicity and SES and future research should elucidate the reasons for these differences. Finally, having established that differences in PA exist between the four ethnic groups included in this thesis, potential differences in other ethnic groups should also be explored.

References

(1) Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep. 1985 Mar-Apr;100(2):126-131.

(2) Kriska AM, Caspersen CJ. Introduction to a Collection of Physical Activity Questionnaires. Medicine & Science in Sports & Medicine 1997;29(6):5-9.

(3) Tai MM, Castillo P, Pi-Sunyer FX. Meal size and frequency: effect on the thermic effect of food. Am.J.Clin.Nutr. 1991 Nov;54(5):783-787.

(4) Froberg K, Andersen LB. Mini review: physical activity and fitness and its relations to cardiovascular disease risk factors in children. Int.J.Obes.(Lond) 2005 Sep;29(Suppl 2):S34-9.

(5) Kohl HW,3rd. Physical activity and cardiovascular disease: evidence for a dose response. Med.Sci.Sports Exerc. 2001 Jun;33(6 Suppl):S472-83; discussion S493-4.

(6) Morris JN, Heady JA, Raffle PA, Roberts CG, Parks JW. Coronary heart-disease and physical activity of work. Lancet 1953 Nov 28;265(6796):1111-20; concl.

(7) Kannel WB, Belanger A, D'Agostino R, Israel I. Physical activity and physical demand on the job and risk of cardiovascular disease and death: the Framingham Study. Am.Heart J. 1986 Oct;112(4):820-825.

(8) Donahue RP, Abbott RD, Reed DM, Yano K. Physical activity and coronary heart disease in middle-aged and elderly men: the Honolulu Heart Program. Am.J.Public Health 1988 Jun;78(6):683-685.

(9) Manson JE, Nathan DM, Krolewski AS, Stampfer MJ, Willett WC, Hennekens CH. A prospective study of exercise and incidence of diabetes among US male physicians. JAMA 1992 Jul 1;268(1):63-67.

(10) Sesso HD, Paffenbarger RS, Jr, Lee IM. Physical activity and coronary heart disease in men: The Harvard Alumni Health Study. Circulation 2000 Aug 29;102(9):975-980.

(11) Lee IM, Rexrode KM, Cook NR, Manson JE, Buring JE. Physical activity and coronary heart disease in women: is "no pain, no gain" passe? JAMA 2001 Mar 21;285(11):1447-1454.

(12) Rockhill B, Willett WC, Manson JE, Leitzmann MF, Stampfer MJ, Hunter DJ, et al. Physical activity and mortality: a prospective study among women. Am.J.Public Health 2001 Apr;91(4):578-583.

(13) Berlin JA, Colditz GA. A meta-analysis of physical activity in the prevention of coronary heart disease. Am.J.Epidemiol. 1990 Oct;132(4):612-628.

(14) Sacco RL, Gan R, Boden-Albala B, Lin IF, Kargman DE, Hauser WA, et al. Leisure-time physical activity and ischemic stroke risk: the Northern Manhattan Stroke Study. Stroke 1998 Feb;29(2):380-387.

(15) Rastogi T, Vaz M, Spiegelman D, Reddy KS, Bharathi AV, Stampfer MJ, et al. Physical activity and risk of coronary heart disease in India. Int.J.Epidemiol. 2004 Aug;33(4):759-767.

(16) Rothenbacher D, Hoffmeister A, Brenner H, Koenig W. Physical activity, coronary heart disease, and inflammatory response. Arch.Intern.Med. 2003 May 26;163(10):1200-1205.

(17) Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 2004 Sep 11-17;364(9438):937-952.

(18) Tudor-Locke C, Bell RC, Myers AM, Harris SB, Ecclestone NA, Lauzon N, et al. Controlled outcome evaluation of the First Step Program: a daily physical activity intervention for individuals with type II diabetes. Int.J.Obes.Relat.Metab.Disord. 2004 Jan;28(1):113-119.

(19) Margareta Eriksson K, Westborg CJ, Eliasson MC. A randomized trial of lifestyle intervention in primary healthcare for the modification of cardiovascular risk factors. Scand.J.Public Health 2006;34(5):453-461.

(20) Yin Z, Moore JB, Johnson MH, Barbeau P, Cavnar M, Thornburg J, et al. The Medical College of Georgia Fitkid project: the relations between program attendance and changes in outcomes in year 1. Int.J.Obes.(Lond) 2005 Sep;29 Suppl 2:S40-5.

(21) Araiza P, Hewes H, Gashetewa C, Vella CA, Burge MR. Efficacy of a pedometerbased physical activity program on parameters of diabetes control in type 2 diabetes mellitus. Metabolism 2006 Oct;55(10):1382-1387.

(22) Murphy MH, Murtagh EM, Boreham CA, Hare LG, Nevill AM. The effect of a worksite based walking programme on cardiovascular risk in previously sedentary civil servants [NCT00284479]. BMC Public Health 2006 May 22;6:136.

(23) Bassuk SS, Manson JE. Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease. J.Appl.Physiol. 2005 Sep;99(3):1193-1204.

(24) Pan XR, Li GW, Hu YH, Wang JX, Yang WY, An ZX, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study. Diabetes Care 1997 Apr;20(4):537-544.

(25) Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N.Engl.J.Med. 2001 May 3;344(18):1343-1350.

(26) Diabetes Prevention Program Research Group. Reduction in the Incidence of Type 2 Diabetes with Lifestyle Intervention or Metformin. NEJM 2002;346(6):393-403.

(27) Hill JO, Wyatt HR. Role of physical activity in preventing and treating obesity. J.Appl.Physiol. 2005 Aug;99(2):765-770.

(28) Wareham NJ, van Sluijs EM, Ekelund U. Physical activity and obesity prevention: a review of the current evidence. Proc.Nutr.Soc. 2005 May;64(2):229-247.

(29) Petersen L, Schnohr P, Sorensen TI. Longitudinal study of the long-term relation between physical activity and obesity in adults. Int.J.Obes.Relat.Metab.Disord. 2004 Jan;28(1):105-112.

(30) Donnelly JE, Kirk EP, Jacobsen DJ, Hill JO, Sullivan DK, Johnson SL. Effects of 16 mo of verified, supervised aerobic exercise on macronutrient intake in overweight men and women: the Midwest Exercise Trial. Am.J.Clin.Nutr. 2003 Nov;78(5):950-956.

(31) Reilly JJ, Kelly L, Montgomery C, Williamson A, Fisher A, McColl JH, et al. Physical activity to prevent obesity in young children: cluster randomised controlled trial. BMJ 2006 Nov 18;333(7577):1041.

(32) Yusuf S, Hawken S, Ounpuu S, Bautista L, Franzosi MG, Commerford P, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. Lancet 2005 Nov 5;366(9497):1640-1649.

(33) Price GM, Uauy R, Breeze E, Bulpitt CJ, Fletcher AE. Weight, shape, and mortality risk in older persons: elevated waist-hip ratio, not high body mass index, is associated with a greater risk of death. Am.J.Clin.Nutr. 2006 Aug;84(2):449-460.

(34) Ross R, Janssen I. Is abdominal fat preferentially reduced in response to exerciseinduced weight loss? Med.Sci.Sports Exerc. 1999 Nov;31(11 Suppl):S568-72. (35) Kay SJ, Fiatarone Singh MA. The influence of physical activity on abdominal fat: a systematic review of the literature. Obes.Rev. 2006 May;7(2):183-200.

(36) Spanier PA, Marshall SJ, Faulkner GE. Tackling the obesity pandemic: a call for sedentary behaviour research. Canadian Journal of Public Health.Revue Canadienne de Sante Publique 2006;97(3):255-257.

(37) Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. Br.J.Sports Med. 2003 discussion 206; Jun;37(3):197-206.

(38) Janz KF. Physical activity in epidemiology: moving from questionnaire to objective measurement. Br.J.Sports Med. 2006 Mar;40(3):191-192.

(39) Tudor-Locke CE, Myers AM. Challenges and opportunities for measuring physical activity in sedentary adults. Sports Med. 2001 Feb;31(2):91-100.

(40) Baecke Questionnaire of Habitual Physical Activity. Med.Sci.Sports Exerc. 1997;29(6 (Supplement)):15-18.

(41) Modifiable Activity Questionnaire. Med.Sci.Sports Exerc. 1997;29(6 (Supplement)):73-78.

(42) Framingham Physical Activity Index. Med.Sci.Sports Exerc. 1997;29(6 (Supplement)):33-35.

(43) CARDIA Physical Activity History. Med.Sci.Sports Exerc. 1997;29(6 (Supplement)):25-32.

(44) Fogelholm M, Malmberg J, Suni J, Santtila M, Kyrolainen H, Mantysaari M, et al. International Physical Activity Questionnaire: Validity against fitness. Med.Sci.Sports Exerc. 2006 Apr;38(4):753-760.

(45) Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med.Sci.Sports Exerc. 2003 Aug;35(8):1381-1395.

(46) Streiner DL, Norman GR. Health Measurement Scales: a practical guide to their development and use. Third ed. New York: Oxford University Press; 2003.

(47) Florindo AA, Latorre Mdo R, Santos EC, Negrao CE, Azevedo LF, Segurado AA. Validity and reliability of the Baecke questionnaire for the evaluation of habitual physical activity among people living with HIV/AIDS. Cad.Saude Publica 2006 Mar;22(3):535-541.

(48) Boreham CA, Ferreira I, Twisk JW, Gallagher AM, Savage MJ, Murray LJ. Cardiorespiratory fitness, physical activity, and arterial stiffness: the Northern Ireland Young Hearts Project. Hypertension 2004 Nov;44(5):721-726.

(49) Warnecke RB, Johnson TP, Chavez N, Sudman S, O'Rourke DP, Lacey L, et al. Improving question wording in surveys of culturally diverse populations. Ann.Epidemiol. 1997 Jul;7(5):334-342.

(50) Motl RW, McAuley E, DiStefano C. Is social desirability associated with self-reported physical activity? Prev.Med. 2005 Jun;40(6):735-739.

(51) Thomas JR, Nelson JK. Research Methods in Physical Activity. Second Edition ed. Champaign, Illinois: Human Kinetics Books; 1990.

(52) Trochim WMK. The Research Methods Knowledge Base. 2nd ed. Cincinnati, OH: Atomic Dog Publishing; 2001.

(53) Bowles HR, FitzGerald SJ, Morrow JR, Jr, Jackson AW, Blair SN. Construct validity of self-reported historical physical activity. Am.J.Epidemiol. 2004 Aug 1;160(3):279-286.

(54) Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med. 2001;31(6):439-454.

(55) Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am.J.Clin.Nutr. 1982 Nov;36(5):936-942.

(56) Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychol.Bull. 1955 Jul;52(4):281-302.

(57) Matthews CE, Freedson PS, Hebert JR, Stanek EJ,3rd, Merriam PA, Rosal MC, et al. Seasonal variation in household, occupational, and leisure time physical activity: longitudinal analyses from the seasonal variation of blood cholesterol study. Am.J.Epidemiol. 2001 Jan 15;153(2):172-183.

(58) Uitenbroek DG. Seasonal variation in leisure time physical activity. Med.Sci.Sports Exerc. 1993 Jun;25(6):755-760.

(59) Rifas-Shiman SL, Gillman MW, Field AE, Frazier AL, Berkey CS, Tomeo CA, et al. Comparing physical activity questionnaires for youth: seasonal vs annual format. Am.J.Prev.Med. 2001 May;20(4):282-285.

(60) Kriska AM, Knowler WC, LaPorte RE, Drash AL, Wing RR, Blair SN, et al. Development of questionnaire to examine relationship of physical activity and diabetes in Pima Indians. Diabetes Care 1990 Apr;13(4):401-411.

(61) Mahabir S, Baer DJ, Giffen C, Clevidence BA, Campbell WS, Taylor PR, et al. Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women. Am.J.Clin.Nutr. 2006 Jul;84(1):230-236.

(62) Conway JM, Seale JL, Jacobs DR, Jr, Irwin ML, Ainsworth BE. Comparison of energy expenditure estimates from doubly labeled water, a physical activity questionnaire, and physical activity records. Am.J.Clin.Nutr. 2002 Mar;75(3):519-525.

(63) Arvidsson D, Slinde F, Hulthen L. Physical activity questionnaire for adolescents validated against doubly labelled water. Eur.J.Clin.Nutr. 2005 Mar;59(3):376-383.

(64) Westerterp KR, Plasqui G. Physical activity and human energy expenditure. Curr.Opin.Clin.Nutr.Metab.Care 2004 Nov Nov;7(6):607-613.

(65) Schoeller DA. Recent advances from application of doubly labeled water to measurement of human energy expenditure. J.Nutr. 1999 Oct;129(10):1765-1768.

(66) Speakman JR. The history and theory of the doubly labeled water technique. Am.J.Clin.Nutr. 1998 Oct;68(4):932S-938S.

(67) Stellaard F. Use of dual isotope tracers in biomedical research. Isotopes Environ.Health Stud. 2005 Sep;41(3):275-286.

(68) Tudor-Locke C. Taking Steps Toward Increased Physical Activity: Using Pedometers to Measure and Motivate. President's Council on Physical Fitness and Sports Research Digest 2002;3(17):1-8.

(69) Tudor-Locke C, Williams JE, Reis JP, Pluto D. Utility of pedometers for assessing physical activity: convergent validity. Sports Med. 2002;32(12):795-808.

(70) Ainslie P, Reilly T, Westerterp K. Estimating human energy expenditure: a review of techniques with particular reference to doubly labelled water. Sports Med. 2003;33(9):683-698.

(71) Dishman RK, Washburn RA, Heath GW. Physical Activity Epidemiology. Windsor, Ontario, Canada: Human Kinetics; 2004.

(72) Friedenreich CM, Courneya KS, Neilson HK, Matthews CE, Willis G, Irwin M, et al. Reliability and validity of the Past Year Total Physical Activity Questionnaire. Am.J.Epidemiol. 2006 May 15;163(10):959-970.

(73) Mathie MJ, Coster AC, Lovell NH, Celler BG. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. Physiol.Meas. 2004 Apr;25(2):R1-20.

(74) King GA, Torres N, Potter C, Brooks TJ, Coleman KJ. Comparison of activity monitors to estimate energy cost of treadmill exercise. Med.Sci.Sports Exerc. 2004 Jul;36(7):1244-1251.

(75) Vanhees L, Lefevre J, Philippaerts R, Martens M, Huygens W, Troosters T, et al. How to assess physical activity? How to assess physical fitness? Eur.J.Cardiovasc.Prev.Rehabil. 2005 Apr;12(2):102-114.

(76) Nichols JF, Morgan CG, Sarkin JA, Sallis JF, Calfas KJ. Validity, reliability, and calibration of the Tritrac accelerometer as a measure of physical activity. Med.Sci.Sports Exerc. 1999 Jun;31(6):908-912.

(77) Welk GJ, Blair SN, Wood K, Jones S, Thompson RW. A comparative evaluation of three accelerometry-based physical activity monitors. Med.Sci.Sports Exerc. 2000 Sep;32(9 Suppl):S489-97.

(78) Bailey RC, Olson J, Pepper SL, Porszasz J, Barstow TJ, Cooper DM. The level and tempo of children's physical activities: an observational study. Med.Sci.Sports Exerc. 1995 Jul;27(7):1033-1041.

(79) Pfeiffer KA, McIver KL, Dowda M, Almeida MJ, Pate RR. Validation and calibration of the Actical accelerometer in preschool children. Med.Sci.Sports Exerc. 2006 Jan;38(1):152-157.

(80) DeVoe D, Gotshall R, McArthur T. Comparison of the RT3 Research Tracker and Tritrac R3D accelerometers. Percept.Mot.Skills 2003 Oct;97(2):510-518.

(81) Hill C, Carnes B. RT3 Tri-axial Research Tracker Kit: Comparing the Stayhealthy RT3 to the Tritrac-R3D. 2004-2007; Available at: http://www.stayhealthy.com/products/rt3-compare.php. Accessed April 5, 2007.

(82) Rowlands AV, Thomas PW, Eston RG, Topping R. Validation of the RT3 triaxial accelerometer for the assessment of physical activity. Med.Sci.Sports Exerc. 2004 Mar;36(3):518-524.

(83) Powell SM, Rowlands AV. Intermonitor variability of the RT3 accelerometer during typical physical activities. Med.Sci.Sports Exerc. 2004 Feb;36(2):324-330.

(84) Kozub FM, Oh HK, Rider RA. RT3 Accelerometer Accuracy in Estimating Short Term Physical Activity in Individuals With Visual Impairments. Adapted Physical Activity Quarterly 2005;22:265-276.

(85) Bryan SN, Tremblay MS, Perez CE, Ardern CI, Katzmarzyk PT. Physical activity and ethnicity: evidence from the Canadian Community Health Survey. Can.J.Public Health 2006 Jul-Aug;97(4):271-276.

(86) Crespo CJ, Smit E, Andersen RE, Carter-Pokras O, Ainsworth BE. Race/ethnicity, social class and their relation to physical inactivity during leisure time: results from the Third National Health and Nutrition Examination Survey, 1988-1994. Am.J.Prev.Med. 2000 Jan;18(1):46-53.

(87) Kriska AM, Rexroad AR. The role of physical activity in minority populations. Womens Health Issues 1998 Mar-Apr;8(2):98-103.

(88) Kriska A. Ethnic and cultural issues in assessing physical activity. Res.Q.Exerc.Sport 2000 Jun;71(2 Suppl):S47-53.

(89) Johnson TP. Approaches to equivalence in cross-cultural and cross-national survey research. ZUMA Nachrichten Spezial 1998;3:1-40.

(90) Henderson KA, Ainsworth BE. A synthesis of perceptions about physical activity among older African American and American Indian women. Am.J.Public Health 2003 Feb;93(2):313-317.

(91) He XZ, Baker DW. Differences in leisure-time, household, and work-related physical activity by race, ethnicity, and education. J.Gen.Intern.Med. 2005 Mar;20(3):259-266.

(92) Sternfeld B, Ainsworth BE, Quesenberry CP. Physical activity patterns in a diverse population of women. Prev.Med. 1999 Mar;28(3):313-323.

(93) Ford ES, Merritt RK, Heath GW, Powell KE, Washburn RA, Kriska A, et al. Physical activity behaviors in lower and higher socioeconomic status populations. Am.J.Epidemiol. 1991 Jun 15;133(12):1246-1256.

(94) Johnson TP, Cho YI, Holbrook AL, O'Rourke D, Warnecke RB, Chavez N. Cultural variability in the effects of question design features on respondent comprehension of health surveys. Ann.Epidemiol. 2006 Sep;16(9):661-668.

(95) Tortolero SR, Masse LC, Fulton JE, Torres I, Kohl HW,3rd. Assessing physical activity among minority women: focus group results. Womens Health Issues 1999 May-Jun;9(3):135-142.

(96) Torgrimson BN, Minson CT. Sex and gender: what is the difference? J.Appl.Physiol. 2005 Sep;99(3):785-787.

(97) Ford ME, Kelly PA. Conceptualizing and categorizing race and ethnicity in health services research. Health Services Research 2005;40(5, Part II):1658-1675.

(98) Oppenheimer GM. Paradigm lost: race, ethnicity, and the search for a new population taxonomy. Am.J.Public Health 2001 Jul;91(7):1049-1055.

(99) LaVeist TA. Disentangling race and socioeconomic status: a key to understanding health inequalities. J.Urban Health 2005 Jun;82(2 Suppl 3):iii26-34.

(100) Sallis JF, Zakarian JM, Hovell MF, Hofstetter CR. Ethnic, socioeconomic, and sex differences in physical activity among adolescents. J.Clin.Epidemiol. 1996 Feb;49(2):125-134.

(101) Anand SS, Razak F, Davis AD, Jacobs R, Vuksan V, Teo K, et al. Social disadvantage and cardiovascular disease: development of an index and analysis of age, sex, and ethnicity effects. Int.J.Epidemiol. 2006 Oct;35(5):1239-1245.

(102) Kriska AM, Knowler WC, LaPorte RE, Drash AL, Wing RR, Blair SN, et al. Development of questionnaire to examine relationship of physical activity and diabetes in Pima Indians. Diabetes Care 1990 Apr;13(4):401-411.

(103) Gordon-Larsen P, McMurray RG, Popkin BM. Adolescent physical activity and inactivity vary by ethnicity: The National Longitudinal Study of Adolescent Health. J.Pediatr. 1999 Sep;135(3):301-306.

(104) Marshall SJ, Jones DA, Ainsworth BE, Reis JP, Levy SS, Macera CA. Race/Ethnicity, Social Class, and Leisure-Time Physical Inactivity. Med.Sci.Sports Exerc. 2007 Jan;39(1):44-51.

(105) Newbold KB. Problems in search of solutions: health and Canadian aboriginals. J.Community Health 1998 Feb;23(1):59-73.

(106) Sheth T, Nargundkar M, Chagani K, Anand S, Nair C, Yusuf S. Classifying ethnicity utilizing the Canadian Mortality Data Base. Ethn.Health 1997 Nov;2(4):287-295.

(107) ProPhone Canada. Available at: http://www.cd-rom-directories.com/.

(108) Allain CC, Poon LS, Chan CS, Richmond W, Fu PC. Enzymatic determination of total serum cholesterol. Clin.Chem. 1974 Apr;20(4):470-475.

(109) McGowan MW, Artiss JD, Strandbergh DR, Zak B. A peroxidase-coupled method for the colorimetric determination of serum triglycerides. Clin.Chem. 1983 Mar;29(3):538-542.

(110) Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of lowdensity lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin.Chem. 1972 Jun;18(6):499-502.

(111) Warnick GR, Nguyen T, Albers AA. Comparison of improved precipitation methods for quantification of high-density lipoprotein cholesterol. Clin.Chem. 1985 Feb;31(2):217-222.

(112) Richardson MT, Ainsworth BE, Wu HC, Jacobs DR, Jr, Leon AS. Ability of the Atherosclerosis Risk in Communities (ARIC)/Baecke Questionnaire to assess leisure-time physical activity. Int.J.Epidemiol. 1995 Aug;24(4):685-693.

(113) Folsom AR, Arnett DK, Hutchinson RG, Liao F, Clegg LX, Cooper LS. Physical activity and incidence of coronary heart disease in middle-aged women and men. Med.Sci.Sports Exerc. 1997 Jul;29(7):901-909.

(114) Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Medicine & Science in Sports & Exercise 2000 Sep;32(9 Suppl):S498-504.

(115) Institute of Medicine. Dietary reference intakes for energy, carbohydrates, fibre, fat, fatty acids, protein and amino acids. Washington, D.C.: National Academy Press; 2002.

(116) Stevens J, Cornell CE, Story M, French SA, Levin S, Becenti A, et al. Development of a questionnaire to assess knowledge, attitudes, and behaviors in American Indian children. Am.J.Clin.Nutr. 1999 Apr;69(4 Suppl):773S-781S.

(117) Parcel GS. Social learning theory and health education. Health Education 1981;12(3):14-18.

(118) Perry CL, Baranowski T, Parcel G. How individuals, environments, and health behavior interact: Social learning theory. In: Glantz K, Lewis FM, Rimer B, editors. Health behavior and health education San Francisco: Jossey-Bass; 1990. p. 161-186.

(119) Canada Health Active Living. Available at: http://www.hc-sc.gc.ca/hppb/paguide.

(120) Canada Food Guide to Healthy Eating. Available at: www.hc-sc.gc.ca/hppb/nutrition/pube/foodguid/index.html.

(121) Willett WC. Nutritional Epidemiology. Cary, North Carolina, U.S.A.: Oxford University Press; 1990.

(122) Anand SS, Yusuf S, Jacobs R, Davis AD, Yi Q, Gerstein H, et al. Risk factors, atherosclerosis, and cardiovascular disease among Aboriginal people in Canada: the Study of Health Assessment and Risk Evaluation in Aboriginal Peoples (SHARE-AP). Lancet 2001 Oct 6;358(9288):1147-1153.

(123) Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). 2005 November 2005.

(124) Going SB, Levin S, Harrell J, Stewart D, Kushi L, Cornell CE, et al. Physical activity assessment in American Indian schoolchildren in the Pathways study. Am.J.Clin.Nutr. 1999 Apr;69(4 Suppl):788S-795S.

(125) Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet 2002 Jan 19;359(9302):248-252.

(126) Kelsey JL, Whittemore AS, Evans AS, Thompson WD. Methods in Observational Epidemiology. Second ed. New York: Oxford University Press; 1996.

(127) Anand SS, Yusuf S, Vuksan V, Devanesen S, Teo KK, Montague PA, et al. Differences in risk factors, atherosclerosis, and cardiovascular disease between ethnic groups in Canada: the Study of Health Assessment and Risk in Ethnic groups (SHARE). Lancet 2000 Jul 22;356(9226):279-284.

(128) Maclure M, Schneeweiss S. Causation of bias: the episcope. Epidemiology 2001 Jan;12(1):114-122.

(129) Adams SA, Matthews CE, Ebbeling CB, Moore CG, Cunningham JE, Fulton J, et al. The effect of social desirability and social approval on self-reports of physical activity. Am.J.Epidemiol. 2005 Feb 15;161(4):389-398.

(130) The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. 2004 August 2004.

(131) Taylor RW, Jones IE, Williams SM, Goulding A. Evaluation of waist circumference, waist-to-hip ratio, and the conicity index as screening tools for high trunk fat mass, as measured by dual-energy X-ray absorptiometry, in children aged 3-19 y. Am.J.Clin.Nutr. 2000 Aug;72(2):490-495.

(132) Gillum RF. Distribution of waist-to-hip ratio, other indices of body fat distribution and obesity and associations with HDL cholesterol in children and young adults aged 4-19 years: The Third National Health and Nutrition Examination Survey. Int.J.Obes.Relat.Metab.Disord. 1999 Jun;23(6):556-563.

(133) RT3 User Manual Version 1.2. Elkader, IA: Stayhealthy, Inc.

(134) Hosmer DW, Lemeshow S. Applied Logistic Regression. Second ed. Toronto, Canada: Wiley Inter-Science; 2000.

(135) Jacobs DR, Jr, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med.Sci.Sports Exerc. 1993 Jan;25(1):81-91.

(136) Pols MA, Peeters PH, Bueno-De-Mesquita HB, Ocke MC, Wentink CA, Kemper HC, et al. Validity and repeatability of a modified Baecke questionnaire on physical activity. Int.J.Epidemiol. 1995 Apr;24(2):381-388.

(137) Miller DJ, Freedson PS, Kline GM. Comparison of activity levels using the Caltrac accelerometer and five questionnaires. Med.Sci.Sports Exerc. 1994 Mar;26(3):376-382.

(138) Sallis JF, Condon SA, Goggin KJ, Roby JJ, Kolody B, Alcaraz JE. The development of self-administered physical activity surveys for 4th grade students. Res.Q.Exerc.Sport 1993 Mar;64(1):25-31.

(139) Fransson EI, Alfredsson LS, de Faire UH, Knutsson A, Westerholm PJ, WOLF Study. Leisure time, occupational and household physical activity, and risk factors for cardiovascular disease in working men and women: the WOLF study. Scand.J.Public Health 2003;31(5):324-333.

(140) Mensink GB, Deketh M, Mul MD, Schuit AJ, Hoffmeister H. Physical activity and its association with cardiovascular risk factors and mortality. Epidemiology 1996 Jul;7(4):391-397.

(141) Oppert JM, Thomas F, Charles MA, Benetos A, Basdevant A, Simon C. Leisuretime and occupational physical activity in relation to cardiovascular risk factors and eating habits in French adults. Public Health Nutr. 2006 Sep;9(6):746-754.

(142) Sofi F, Capalbo A, Marcucci R, Gori AM, Fedi S, Macchi C, et al. Leisure time but not occupational physical activity significantly affects cardiovascular risk factors in an adult population. Eur.J.Clin.Invest. 2007 Dec;37(12):947-953.

(143) Hu G, Pekkarinen H, Hanninen O, Yu Z, Guo Z, Tian H. Commuting, leisure-time physical activity, and cardiovascular risk factors in China. Med.Sci.Sports Exerc. 2002 Feb;34(2):234-238.

(144) Aadahl M, Kjaer M, Jorgensen T. Associations between overall physical activity level and cardiovascular risk factors in an adult population. Eur.J.Epidemiol. 2007;22(6):369-378.

(145) Panagiotakos DB, Pitsavos C, Chrysohoou C, Skoumas J, Zeimbekis A, Papaioannou I, et al. Effect of leisure time physical activity on blood lipid levels: the ATTICA study. Coron.Artery Dis. 2003 Dec;14(8):533-539.

(146) Raitakari OT, Porkka KV, Taimela S, Telama R, Rasanen L, Viikari JS. Effects of persistent physical activity and inactivity on coronary risk factors in children and young adults. The Cardiovascular Risk in Young Finns Study. Am.J.Epidemiol. 1994 Aug 1;140(3):195-205.

(147) Ribeiro JC, Guerra S, Oliveira J, Teixeira-Pinto A, Twisk JW, Duarte JA, et al. Physical activity and biological risk factors clustering in pediatric population. Prev.Med. 2004 Sep;39(3):596-601.

(148) Andersen LB, Harro M, Sardinha LB, Froberg K, Ekelund U, Brage S, et al. Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). Lancet 2006 Jul 22;368(9532):299-304.

(149) Raitakari OT, Leino M, Rakkonen K, Porkka KV, Taimela S, Rasanen L, et al. Clustering of risk habits in young adults. The Cardiovascular Risk in Young Finns Study. Am.J.Epidemiol. 1995 Jul 1;142(1):36-44.

(150) Twisk JW, Kemper HC, Van Mechelen W, Post GB. Clustering of risk factors for coronary heart disease. the longitudinal relationship with lifestyle. Ann.Epidemiol. 2001 Apr;11(3):157-165.

(151) Gordon-Larsen P, McMurray RG, Popkin BM. Determinants of adolescent physical activity and inactivity patterns. Pediatrics 2000 Jun;105(6):E83.

(152) Howell DC. Treatment of Missing Data. 2007; Available at: http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/Missing.html. Accessed Aug 12, 2008, 2008.

(153) Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol.Methods 2002 Jun;7(2):147-177.

(154) Diez-Roux AV. Multilevel analysis in public health research. Annu.Rev.Public Health 2000;21:171-192.

(155) Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ 1995 Jan 21;310(6973):170.