ASSESSING PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR IN INDIVIDUALS WITH SCHIZOPHRENIA

by

Markus Joseph Duncan

M.Sc., University of Toronto, 2014

B.Sc., University of Toronto, 2012

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Kinesiology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

ASSESSING PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR IN INDIVIDUALS WITH SCHIZOPHRENIA

submitted by Markus Joseph Duncan in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Kinesiology

Examining Committee:

Dr. Guy Faulkner, School of Kinesiology, UBC
Supervisor

Dr. Alasdair Barr, Faculty of Medicine (Anesthesiology, Pharmacology & Therapeutics), UBC
Supervisory Committee Member

Dr. Mark Beauchamp, School of Kinesiology, UBC
Supervisory Committee Member

Dr. Jean-Sebastien Blouin, School of Kinesiology, UBC
University Examiner

Dr. Donna Lang, Faculty of Medicine (Radiology), UBC
University Examiner

Abstract

Individuals with schizophrenia have a greatly reduced life expectancy compared to the general population due in part to higher rates of diabetes, obesity, and cardiovascular disease. Individuals with schizophrenia also engage in lower volumes of physical activity and engage in more sedentary behaviour, which contributes to poor overall health. Evaluating time spent in such movement behaviours accurately is necessary to engage in all phases of behavioural epidemiology and for developing interventions to induce behaviour change.

This dissertation evaluates whether the most commonly used self-report questionnaire, the International Physical Activity Questionnaire (IPAQ), for measuring time spent in physical activity (Chapter 2) and sedentary behaviour (Chapter 3) among people with schizophrenia is an accurate representation of movement behaviours in people with schizophrenia compared to accelerometry derived scores.

After quantifying the discrepancy between measurement methods, Chapter 4 explores individual level correlates of this discrepancy to understand what factors may contribute to less accurate IPAQ scores, and regression-calibration adjustments to IPAQ scores were evaluated using a 5-fold cross-validation approach as a possible method to improve score accuracy.

Subsequently, Chapter 5 assesses whether accelerometry protocols were adhered to consistently in the sample and accelerometry data were used to evaluate the composition of movement behaviours in this sample across the week; an isotemporal substitution approach is used to determine whether replacing one movement behaviour with another (e.g. sedentary behaviour

with moderate intensity physical activity) is associated with health and demographic factors while controlling for all other movement behaviours. The results of these studies suggests that IPAQ scores should no longer be used as a measure of time spent in movement behaviours in individuals with schizophrenia; researchers and medical professionals are recommended to use more direct methods of movement measurement such as accelerometry or leverage the already widespread adoption of consumer smart devices to collect movement behaviour data that better represents how individuals with schizophrenia spend their waking day.

Lay Summary

Increasing physical activity and reducing sedentary behaviour (e.g. sitting) is an important lifestyle component of improving health in individuals with schizophrenia, who in addition to psychosis experience higher rates of serious physical health conditions than the general population. Accurate forms of measurement for behaviours are necessary for researchers and clinicians trying to address this issue. The first two studies evaluate whether a commonly used self-report questionnaire is indeed accurate for measuring physical activity and sedentary behaviour, which it is not. The third study explores reasons why it may not be accurate and attempts to rescale the questionnaire to improve accuracy. The final study quantifies the amount of time spent by this sample in various intensities of movement behaviour. Overall, this research recommends moving away from self-reported measures of physical activity in this population.

Preface

The studies from chapters 2-5 are written in a 'manuscript format'. These studies have either already been published, are currently under review for publication, or are being prepared for submission to a peer-reviewed journal. Ethical approval of data collection was performed by research ethics boards at the Centre for Addiction and Mental Health, Toronto (Protocol # 099/2013) and the University of Toronto (Protocol # #29489). Data were collected by myself (Markus Duncan), Mehala Subramaniapillai, and Carol Borlido as part of a study on determinants of moderate-to-vigorous physical activity in individuals with schizophrenia [citation: Arbour-Nicitopoulos, K. P., Duncan, M. J., Remington, G., Cairney, J., & Faulkner, G. E. (2017). The Utility of the Health Action Process Approach Model for Predicting Physical Activity Intentions and Behavior in Schizophrenia. *Frontiers in Psychiatry*, *8*(August), 1–8. https://doi.org/10.3389/fpsyt.2017.00135]. I contributed to the study design to include self-reported physical activity data collection to be used for this dissertation and performed accelerometer post-processing.

Chapter 2. This study has been published in a peer-reviewed journal, the citation is Duncan, M. J., Arbour-Nicitopoulos, K. P., Subramanieapillai, M., Remington, G., & Faulkner, G. E. (2017). Revisiting the International Physical Activity Questionnaire (IPAQ): Assessing physical activity among individuals with schizophrenia. *Schizophrenia Research*, *179*, 2–7. https://doi.org/10.1016/j.schres.2016.09.010. My contribution involved the formulation of the research question, data collection, data analysis, and manuscript preparation. Guy Faulkner provided guidance on data analysis methodology; all co-authors provided revisions on my drafts of the manuscript.

Chapter 3. This study has been published in a peer-reviewed journal, the citation is: Duncan, M. J., Arbour-Nicitopoulos, K. P., Subramaniapillai, M., Remington, G., & Faulkner, G. E. (2019). Revisiting the International Physical Activity Questionnaire (IPAQ): Assessing sitting time among individuals with schizophrenia. *Psychiatry Research*, *271*(November 2018), 311–318. https://doi.org/10.1016/j.psychres.2018.11.063. My contribution involved the formulation of the research question, data collection, data analysis, and manuscript preparation. Dr. Guy Faulkner provided guidance on data analysis methodology; all co-authors provided revisions on my drafts of the manuscript.

Chapter 4. A manuscript of this study is being prepared for submission to a peer-reviewed journal. My contribution to this study involved formulation of the research question, data collection, data analysis, and manuscript preparation. Dr. Guy Faulkner, Dr. Mark Beauchamp and Dr. Alasdair Barr contributed guidance on relevant literature and provided revisions on my drafts of the manuscript.

Chapter 5. A manuscript of this study is being prepared for submission to a peer-reviewed journal. My contribution to this study involved formulation of the research question, data collection, data analysis, and manuscript preparation. Dr. Guy Faulkner, Dr. Mark Beauchamp and Dr. Alasdair Barr contributed guidance on relevant literature and provided revisions on my drafts of the manuscript.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AES – Apathy Evaluation Scale

ANOVA – Analysis of Variance

MANOVA – Multivariate ANOVA

BMI – Body Mass Index

BPRS – Brief Psychiatric Research Scale 18-item Anchored version

CGI-S – Clinical Global Impression Severity scale

CI – Confidence Interval

CPZ – Chlorpromazine equivalents

cpm – counts per minute

CVD – cardiovascular disease

ICC – Intraclass Correlation Coefficients

IDF-WC – International Diabetes Federation Waist Circumference classification

ILR – Isometric Log Ratio

IPAQ – International Physical Activity Questionnaire

IPAQ-SF – IPAQ short form survey

LNS – Letter Number Span

LoA – Limits of Agreement

MAE – Mean absolute error

MB – Movement Behaviour

MET(s) – Metabolic Equivalent(s)

PA – Physical Activity

LPA – Light PA

MPA – Moderate PA

VPA – Vigorous PA

MVPA – Moderate to vigorous PA

SB – Sedentary Behaviour

SC – Symbol Coding

SF12 – 12-item Short Form Health Survey version 2

SF12-MCS – SF12 Mental Composite Score

SF12-PCS – SF12 Physical Composite Score

Acknowledgements

Foremost I want to express my deep gratitude to my supervisor Dr. Guy Faulkner for being my mentor over the course of my Doctorate. It has truly been an honour working with someone who is both a preeminent mind in the field of mental health and physical activity and genuinely invested in the personal and professional development of his students. I would not have come this far without your guidance and am genuinely grateful to have been a part of your laboratory over these many years. I have said it before, and I am sure I will say it in the future: I would always choose to follow you to the University of British Columbia if I was faced with the choice again.

I also owe a debt of gratitude to my advisory committee, who have provided thoughtful and valuable feedback along the way which has helped me think critically and deeply about my work. Dr. Alasdair Barr has been a fount of knowledge in the realm of schizophrenia literature and was particularly helpful in connecting me with his students while writing the fourth chapter of this dissertation. Dr. Mark Beauchamp has been a continuing source of support, your guidance in the field of validity research has been essential and invaluable to my academic development.

Thank you as well to the support I received through Dr. Gary Remington and Dr. Arbour-Nicitopoulos while collecting data at the Centre for Addiction and Mental Health and the University of Toronto. Thank you to Mehala Subramaniapillai and Carol Borlido for your assistance in participant recruitment and data collection and thank you to the those who participated as well.

I would also like to acknowledge the Canadian Institutes of Health Research for providing funding to pursue my Doctorate and the University of British Columbia for the various

scholarships, professional development opportunities, and institutional support I have received along the way.

To my family, thank you for your belief in the value of education and the value of my work. I would not be standing on the shoulders of giants were it not for the step up you have provided for me. To my lab-mates at the POP-PA lab, Dr. Negin Riazi and Dr. Krista Glowacki, thank you for being incredible people to work alongside and for being there when it's time to call it a day.

Finally, to my wife, Alexis Davis, for providing every kind of support imaginable along the way. Thank you for being part of this journey even when it meant being half a world away. Your unwavering belief in me has been more important than you could ever know.

Dedication

For Alexis

Chapter 1: General Introduction

1.1 Introduction

The role of physical activity (PA) for enhancing and maintaining physical health has been well documented. Canadian Physical Activity Guidelines for adults recommend accruing at least 150 minutes of moderate to vigorous PA (MVPA) every week in uninterrupted bouts of at least 10 minutes (Canadian Society of Exercise Physiology, 2012). These guidelines are based on extensive meta-analysis (Warburton et al., 2010) which demonstrate that meeting this threshold greatly reduces the risk of several chronic diseases such as diabetes, cardiovascular disease (CVD), several forms of cancer, and all-cause mortality. Additionally, PA is a well-accepted behavioural strategy for weight management. Thus, there is a strong impetus to promote PA in all populations to enhance long-term health and reduce the burden of disease, an impetus which is even stronger in sub-populations at high risk for preventable illness and who tend not to engage in sufficient amounts of PA.

One sub-population in need of the health enhancing effects of PA are individuals with schizophrenia and schizophrenia-like illnesses (e.g. schizoaffective disorder). Individuals with schizophrenia are at greater risk of CVD, diabetes and obesity than the general population (Annamalai et al., 2017; Bresee et al., 2010; Carney et al., 2006; Coodin, 2001; Dixon et al., 2000; Gurpegui et al., 2012; Hennekens et al., 2005; Manu et al., 2015; Silverstone et al., 1988; Stubbs et al., 2015) which contribute to a greatly reduced life expectancy (Hjorthøj et al., 2017; Laursen et al., 2012, 2014; Olfson et al., 2015). However, a growing body of evidence suggests that behavioural strategies involving diet and exercise can be used to manage weight in individuals with schizophrenia (Caemmerer et al., 2012; Faulkner, Cohn, & Remington, 2007; Hjorth et al., 2014; Naslund et al., 2017).

Sallis and colleagues (2000) have developed a behavioural epidemiological framework that outlines a systematic approach for research in behavioural approaches to health management such as promoting PA. The purpose of Phase 1 of the framework is to "Establish Links Between Behaviors and Health". This phase consists of basic epidemiological research to establish whether a relationship exists between a behaviour and certain health outcomes – including whether greater exposure to the behaviour has a greater impact on health (i.e. a dose-response relationship). Phase 2 is to "Develop Methods for Measuring the Behavior". In the words of Sallis et al. (2000): "high-quality measures are essential for all stages of research." Research in this phase assesses existing measurement tools for evidence of validity and reliability or novel tools are developed and similarly assessed. Phase 3 involves "Identify[ing] Factors That Influence the Behavior" such as demographic correlates, moderating factors, and modifiable determinants which are related to the behaviour. Doing so informs the targets of interventions to change the behaviour, which are then evaluated in Phase 4 ("Evaluate Interventions to Change the Behavior"). Finally, Phase 5 "Translate Research into Practice" uses the knowledge gained from the previous phases to implement evidence-based policy, guidelines, and practice to enhance the PA. While this approach follows a logical, staged, progression, building on the evidence accrued from the previous phase, Sallis et al. (2000) note that the phases may have non-linear or reciprocal feedback on each other. In particular, they highlight that Phase 1 may be revisited and improved upon through evidence generated at Phase 2, and that Phase 1 and 2 may have direct impacts on Phases 4 and 5. Thus, revisiting earlier phases of this framework are warranted to improve the quality of evidence generated at later stages. The overall purpose of this dissertation therefore, is to work backwards along this framework starting by adding to evidence of validity for a common method of assessing PA and sedentary behaviour (SB) in

people with schizophrenia (Phase 2) and assessing whether a more holistic approach to assessing daily movement behaviours can improve the understanding of the relationships between PA, SB and various health outcomes (Phase 1).

1.1.1 Cardio-Metabolic Risk in Schizophrenia

Schizophrenia is a severe and chronic form of mental illness, characterized primarily by the presence of psychosis, specifically hallucinations, delusions, or disorganized speech (American Psychiatric Association, 2013). The presentation of schizophrenia is, however, largely heterogeneous. Different individuals may display a wide variety of symptoms characterized as both positive (symptoms that are an addition or an excessive distortion of typical or healthy functioning, e.g., hallucinations, delusions, and bizarre behaviour) and negative (deficits or loss of functioning, e.g., affective flattening, amotivation, and anhedonia). Cognitive deficits (such as impaired memory, attention, and reasoning) may also occur in 75%-85% of the patient population (McCleery & Nuechterlein, 2019; Palmer et al., 1997; Reichenberg et al., 2006, 2009), and are considered distinct from the negative symptoms (Harvey et al., 2006). In addition to having a heterogenous presentation, schizophrenia is part of a spectrum of schizophrenia-like illnesses which share psychosis as a core symptom, but may have significant mood disturbances during episodes of psychosis (in the case of schizoaffective disorder) or psychotic symptoms may not be as pervasive (<6 months in the case of schizophreniform disorder) (American Psychiatric Association, 2013). In particular, differential diagnosis between schizophrenia and schizoaffective disorders has been notoriously unreliable (although prognostically useful) (Malaspina et al., 2013). Medical treatment of these schizophrenia spectrum disorders often relies on antipsychotic medications to help control psychosis, and many of the physical and

mental health concerns for individuals with schizophrenia spectrum disorders are similar (Eskelinen et al., 2017; Partti et al., 2015; Suvisaari et al., 2008).

Although severe, schizophrenia spectrum disorders are relatively uncommon. The most recent epidemiological reviews concluded that the median worldwide lifetime prevalence of schizophrenia is 0.4-0.5%, and the median lifetime morbidity rate is 0.72% (McGrath et al., 2008; Saha et al., 2005; Simeone et al., 2015). Incidence and onset tend to be different in males when compared to females. Males are about 1.4 times as likely to develop schizophrenia as females (McGrath et al., 2008; Saha et al., 2005). Schizophrenia tends to onset during late adolescence and early adulthood, though males tend to experience a slightly earlier onset than females (Castle et al., 1993; McGrath et al., 2008). Despite being a relatively small proportion of the population, management of schizophrenia carries with it a significant societal burden due to lifelong disability caused by both psychological impairment and physical comorbidity.

Individuals with schizophrenia have a 10-25 year reduced life expectancy compared to the general population, and, while suicide is still a major contributor, this increased mortality is due primarily to natural causes such as physical illness (Chesney et al., 2014; Hjorthøj et al., 2017; Laursen et al., 2012, 2014; Olfson et al., 2015). In particular, individuals with schizophrenia tend to have high rates of CVD (Carney et al., 2006; Newcomer & Hennekens, 2007; Tandon et al., 2008), which accounts for 12%-46% of all-cause mortality within the population (Bushe et al., 2010). Meta-analysis of cross sectional studies estimates the worldwide prevalence of CVD for individuals with schizophrenia is 0.9-2.1 times higher than the general population, while pooled data from longitudinal studies estimates the hazard ratio for CVD at 1.4-2.7 and CVD-related death at 1.3-2.1 times the general population (Correll et al., 2017). Notably, a cohort study of over 1 million individuals with schizophrenia in the United States of

America found that CVD related deaths were the largest contributor to all-cause mortality, and occurred at a rate 3.5-3.6 times that of the general adult population (Olfson et al., 2015). Much like in the general population, the causes of CVD in patients with schizophrenia are varied and complex. However, it appears that unhealthy lifestyle factors and metabolic side effects of antipsychotic medication are likely major contributors to increased risk of CVD in this population.

With regards to lifestyle behaviours, individuals with schizophrenia are more likely to smoke (Cather et al., 2017; D. L. Kelly et al., 2011; McCreadie, 2003; Sagud et al., 2018), consume less healthful diets (Jakobsen et al., 2018; McCreadie et al., 1998; Ryan et al., 2003; Strassnig et al., 2003) and tend to be less physically active (Cohn et al., 2004; Daumit et al., 2005; Jakobsen et al., 2018; Lindamer et al., 2008; Soundy et al., 2013; Stubbs, Firth, et al., 2016; Vancampfort, Firth, et al., 2017; Yamamoto et al., 2011) than the general adult population. This tendency towards an unhealthy lifestyle is compounded by weight gain associated with the commencement of antipsychotic treatment (Allison et al., 1999; Bak et al., 2014; Homel et al., 2002; Meyer et al., 2008). Evidence suggests that even medications that are considered to be of low metabolic risk may instead have a delayed pattern of weight gain compared to other antipsychotics (Correll et al., 2009; Perez-Iglesias et al., 2008). The exact cause of this phenomenon is not well understood though several possible biological mechanisms have been posited, such as antagonism of histamine and dopamine receptors, leading to increased appetite, and changes in weight homeostasis as a result of hyperprolactinemia or drug-induced leptin resistance (Jin et al., 2008; Reynolds & Kirk, 2010). As a result of both an unhealthy lifestyle and medication side effects, individuals with schizophrenia are not only at greater risk of CVD, but also diabetes and obesity (Dixon et al., 2000; Hennekens et al., 2005; Manu et al., 2015).

Specifically, individuals with schizophrenia have a 1.6-3 times greater odds of having diabetes (Annamalai et al., 2017; Bresee et al., 2010; Carney et al., 2006; Dixon et al., 2000; Stubbs et al., 2015), 4.0-4.3 times more likely to die of diabetes related complications (Olfson et al., 2015), and are 1.5-4 times more likely to be obese (Annamalai et al., 2017; Carney et al., 2006; Coodin, 2001; Gurpegui et al., 2012; Silverstone et al., 1988) than the general population.

Obesity has additional implications for individuals living with schizophrenia. In particular, central obesity has been shown to be associated with decreased quality of life within the schizophrenia population (Dayabandara et al., 2017; Faulkner, Cohn, Remington, et al., 2007; Sugawara et al., 2013) Moreover, distress regarding weight gain is a contributor to medication non-compliance in people with schizophrenia (Dayabandara et al., 2017). Thus, developing non-pharmacological strategies to mitigate weight gain will likely increase quality of life and simultaneously reduce the risk of CVD, diabetes, and all-cause mortality within this population.

1.1.2 Movement Behaviours in Schizophrenia

Movement behaviours (MB) of the waking day can be broadly categorized as PA or SB. PA is “any bodily movement produced by skeletal muscles that requires energy expenditure” (World Health Organisation, 2004). PA can be subdivided into light, moderate, and vigorous intensities which are defined based on Metabolic Equivalents (METs). One MET is the resting state of humans and is defined as 3.5 ml of O_2 consumed per kg body weight per minute of activity (approximately 1kcal per kg of bodyweight per hour). Based on this, extensive testing has resulted in a list of activities and their MET values compared to this standard (Ainsworth et al., 2011) with more intense activity resulting in higher MET numbers. Activities considered

light PA (LPA) has a MET value <3, moderate PA (MPA) has a MET value between 3-5.9, and vigorous PA (VPA) has a MET value ≥ 6 (World Health Organisation, 2004).

PA is contrasted with SB – any *waking* activity characterized by an energy expenditure less than 1.5 metabolic equivalents in a sitting, reclining or lying posture (Tremblay, Aubert, et al., 2017). Thus an individual's day can be described as a combination of the time spent in sleep, SB, and varying intensities of PA. Based on extensive meta-analysis (Warburton et al., 2010), current Canadian Physical Activity Guidelines for adults recommend accruing at least 150 minutes of MVPA every week in uninterrupted bouts of at least 10 minutes to reduce the risk of several chronic diseases (Canadian Society of Exercise Physiology, 2012). Individuals who meet these guidelines are considered to be *physically active*, whereas those who do not are classified as *inactive*. However, the 2012 Canadian Physical Activity Guidelines for adults (Canadian Society of Exercise Physiology, 2012) do not have recommendations for LPA or SB, though updated guidelines will account for 24 hour time use in the near future (Ross et al, in press).

PA may provide multiple health benefits for persons with schizophrenia, such as reducing excess weight, improving glycemic control, and reducing the risk of CVD which would help address the high mortality and reduce the lower life expectancy observed in individuals with schizophrenia (Suetani et al., 2016). Several systematic reviews of randomized controlled trials have demonstrated that non-pharmacological interventions for managing weight – including PA – in individuals who have schizophrenia (Faulkner, Cohn, & Remington, 2007; Gurusamy et al., 2018), severe mental illness (Naslund et al., 2017) or who are using antipsychotics (Caemmerer et al., 2012) are plausible and modestly efficacious. In addition to the physical health benefits of PA, systematic reviews of randomized control trials reveal small but significant improvements in the mental health of patients with schizophrenia who participate in the assigned PA intervention

(Gorczynski & Faulkner, 2010), suggesting that regular PA can not only help prevent and manage physical co-morbidities, but also may improve psychological health. A meta-analysis of 7 randomized control trials suggests that exercise interventions can also improve cognitive function in individuals with schizophrenia (Firth et al., 2016). Despite these health benefits, physical inactivity remains highly prevalent within this population (Cohn et al., 2004; Daumit et al., 2005; Lindamer et al., 2008; Soundy et al., 2013; Stubbs, Firth, et al., 2016; Vancampfort, Firth, et al., 2017; Yamamoto et al., 2011).

While MVPA is an important component of health promotion and weight management, meta-analysis has demonstrated that greater amounts of daily SB are associated with higher rates of cardiovascular disease, diabetes, and all-cause mortality (Biswas et al., 2015; Warren et al., 2010). This association appears to be independent of MVPA (Biswas et al., 2015). Even individuals who are currently meeting PA guidelines may receive cardio-metabolic health benefits from interrupting sitting time with short breaks (Owen, Healy, Matthews, & Dunstan, 2010; Thorp, Owen, Neuhaus, & Dunstan, 2011). As such, studies have begun to demonstrate an association between schizophrenia and health outcomes such as body mass index among individuals with schizophrenia as well (Bueno-Antequera et al., 2017). However, a common criticism of SB research is that *only* MVPA is adjusted for and not LPA (Pedišić, 2014; van der Ploeg & Hillsdon, 2017). Additionally, observation periods have a limited time budget available and therefore increasing any one MB inherently requires that another be decreased (Pedišić, 2014). As van der Ploeg and Hillsdon (2017) state: "If no adjustment is made for light physical activity then it is difficult to determine whether the association with poor health is due to more time sedentary or less time in light activity." Thus, the benefits of decreasing SB must be considered with regards to which PA behaviour replaces SB. Therefore, while continuing to test

interventions to increase MVPA among individuals with schizophrenia is still important (McNamee et al., 2013; Vancampfort, Rosenbaum, et al., 2016), it is possible that physically active individuals with schizophrenia may accrue additional health benefits by reducing prolonged sitting time. Furthermore, replacing SB with LPA may be easier than promoting more intense PA while still accruing some health benefit.

Despite the potential benefits of being more active, individuals with schizophrenia and other severe mental illnesses tend to engage in large amounts of SB and low levels of MVPA. One meta-analysis demonstrated that individuals with psychosis engage in high levels of SB (Stubbs, Williams, et al., 2016). Across 13 studies the authors reported a pooled effect size of 11h of SB per day in this population. Among the four studies that compared individuals with psychosis to healthy controls, individuals with psychosis were engaging in 2.8h more SB. These numbers increase to 12.6h of SB and 2.9h more than healthy controls when only objective measures of SB are used. The authors suggested that self-report questionnaires may be largely underestimating the amount of SB that individuals with psychosis are engaging in, though no studies included in the review used both objective and self-report methods simultaneously to evaluate this hypothesis. A subsequent meta-analysis (Vancampfort, Firth, et al., 2017) of sixty-nine studies found people with severe mental illness (major depressive disorder, bipolar or schizophrenia) engaged in an average of 476.0 min/day (95% CI: 407.3-545.4) SB and 38.4 min/day (95% CI: 32.0-44.8) of MVPA. Individuals with SMI engaged in significantly more SB and less MVPA than age- and gender-matched healthy controls and were less likely to meet MVPA guidelines.

1.1.3 Measuring Time Spent in Movement Behaviours

As per Sallis and colleagues' (2000) behavioural epidemiology framework, assessing tools and methods of measuring the target behaviours (i.e. PA and SB) is necessary to proceed with identifying factors that influence behaviour, developing interventions to change behaviours, and translating this research into practice (i.e. Phases 3-5) as well as refining the understanding of the relationship between behaviours and health outcomes (i.e. revisiting Phase 1). In the case of PA and SB, if one of the end-goals is to develop specific guidelines and policy initiatives regarding the intensity and amount of time that individuals with schizophrenia should engage in regular PA (or limit SB) then tools used in these research endeavours must be able to provide accurate representations of time and intensity.

Methods of assessing PA range from direct physical and biological methods to various subjective self-report methods. However, many of the most objective methods can only be conducted in a controlled laboratory environment or are otherwise not suitable for surveillance of typical behaviour in the field. Field methods of assessing PA are therefore the most relevant in the context of epidemiology and intervention research. Accelerometers have become the *de facto* direct measure in PA field research for a number of reasons. Compared to the gold standard method of assessing metabolic expenditure using doubly labelled water, accelerometers are easier to distribute, less costly, and less burdensome. Furthermore, the only data that doubly labelled water provides is energy expenditure in the time between water intake and body water sampling. As such, achieving a measurement of PA in terms of minutes of PA would require multiple samples and additional burden on participants. Given that PA guidelines use minutes of activity and distinguish between SB, LPA, MPA and/or VPA, measurement tools that provide

data in this format are more appropriate for epidemiological research that can be easily translated into guidelines and recommendations.

Computer based software for an accelerometer allows researchers to transform this data into estimates of energy expenditure, activity bout duration, and intensity data by using established cut points for intensity based on previous calibration studies in a controlled setting. For example Troiano and colleagues, (2008) have suggested that SB be set ≤99 counts per minute, 100-2019 counts for LPA, 2020-5998 for MPA, and ≥5999 VPA. However, other cut-points exist, and the software allows researchers to set and test cut points appropriate to the research question and population. Additionally, accelerometers provide a high level of time resolution in the data as the epochs are time stamped, so researchers can examine patterns of physical activities, such as when in the day activities were performed, and if activity was continuous or interspersed with breaks. As such, accelerometers are powerful tools for identifying patterns of activity throughout the day. However, like other objective measures, accelerometers require supplemental data collection to assess the contextual aspects of PA such as mode and location. Given the role of the researcher in making decisions about cut-points and data cleaning it must be acknowledged that there are subjective aspects of data collected through accelerometry

Self-report questionnaires are less costly to administer than more objective and direct measures, can generally be completed in a short amount of time, and do not require the use of monitoring devices that must be worn for extended periods. As a result questionnaires may help address the needs of physical therapists working in mental health settings who have asked for assessments of MBs appropriate for individuals with schizophrenia (Vancampfort, Rosenbaum, et al., 2016). Furthermore, depending on the items, questions, and prompts used, self-report

methods can assess the frequency, intensity, duration, and time of activity, which in turn can be used to estimate energy expenditure through MET conversion. Questionnaires that attempt to assess PA time and intensity differ from traditional psychometric questionnaires aimed at assessing psychological constructs. A psychological construct such as depression represents a latent (unobservable) variable scored by combining measured items to assess the construct, and scores are arbitrarily determined based on how the designer feels the score will best communicate the meaning (e.g. a score of 0 does not necessarily mean an absence of depression). Time spent in PA and SB are however observable and non-arbitrary, thus PA questionnaire scores are intended to estimate time spent in a given activity at a specified intensity. Therefore, while Messick (1995) suggests that all sources of evidence for validity should be integratively considered when considering the confidence in a measurement tool, accuracy of the PA measurement – how much the tool differs from the actual behaviour is of particular relevance (Freedson et al., 2012; Prince et al., 2008; Soundy et al., 2014). As a result, validation studies in PA focus on the relationship between the questionaire scores and device-based methods of assessing PA time or metabolic expenditure (Craig et al., 2003; Morrow, 2002; Sallis & Saelens, 2000; Soundy et al., 2014).

Common criticisms of self-report methods are that they may not cover important dimensions of PA and can be misinterpreted (Bishop, 2008; Dale et al., 2002). Additionally, self-report methods require participants to be able to recall their activities. This may be especially challenging for incidental activity such as walking to locations and among populations with memory impairment. Participants may also incorrectly classify the intensity of activities. For example, Ekkekakis, Parfitt, and Petruzzello's review (2011) suggests that individuals experience different affective responses to the same intensity of exercise, which may be

attributable in part to age, gender, physical fitness, and weight. Additionally, fitness level of participants affects ratings of perceived exertion in sub-ventilatory threshold exercise (Garcin et al., 2004). Thus, individual differences in cognitive ability, physical health and fitness, and demographic factors may affect how accurately individuals are able to respond to self-report questionnaires.

While many tools for measuring PA have undergone validation studies in the general population, evidence of validity is still lacking for measurements of PA in the schizophrenia population. A search of MEDLINE[1] for articles on the psychometrics of PA measurement tools returned 45 unique records, of which 8 appeared to be relevant based on their title. Three of these were omitted for assessing fitness as an external criterion variable rather than time spent in or energy expended due to PA (Gomes et al., 2016; Vancampfort, Guelinckx, et al., 2014; Vancampfort, Probst, Sweers, et al., 2012), another two were omitted for not reporting any novel validity or reliability data (Beebe & Harris, 2012; Janney et al., 2015). The remaining 3 studies (Faulkner, Cohn, & Remington, 2006; Lindamer et al., 2008; Sharpe, Stedman, Byrne, & Hills, 2006) were supplemented with an additional known study (Soundy et al., 2007) examining reliability and validity of PA among samples predominantly composed of people with schizophrenia, and a validation study of the IPAQ-SF MPA and VPA scores in individuals with first episode psychosis (Vancampfort, De Hert, et al., 2017). Only Vancampfort and colleagues' study (2017) reflects an update in the literature since Soundy and colleagues' (2014) review of validation studies in severe mental illness which concluded that such studies have overlooked evidence of agreement (i.e. the use of Bland-Altman analysis (Bland & Altman, 1986)) between

[1] Search strategy: [Psychometrics OR *psychometric** OR *valid** OR *reliab** OR Surveys and Questionnaires OR Reproducibility of Results] AND [Exercise OR *physical activity*] AND [Schizophrenia].

the self-report measurements and objective measurements as an indication of accuracy. Instead studies tend to use correlations between measurements. While correlations may be appropriate for indicating that two scales measuring the same construct are related (that is individuals who score highly on one scale tend to also score highly on another) common statistical methods used such as Pearson's correlation coefficient, Spearman's rho and Kendall's tau do not account for discrepancies in absolute values. That is, perfect correlations are still possible with these methods even if the values differ widely, so long as a straight line can be drawn between data points regardless of the slope or intercept. However, some correlation variants such as the concordance correlation coefficient, absolute agreement intraclass correlation coefficients, and Cohen's kappa do take agreement into account such that greater disagreement between measurements results in lower correlations.

Three of the five studies identified (Faulkner et al., 2006; Lindamer et al., 2008; Soundy et al., 2007) compared self-report based assessments of PA to an accelerometer as evidence of validity as well as test-retest reliability of these tools. Vancampfort and colleagues (2017) did not measure retest reliability, but did compare a self-report assessment to accelerometry. Sharpe et al. (2006) compared two direct measures. One study examined the Yale Physical Activity Questionnaire (Lindamer et al., 2008), another examined the 7-day Physical Activity Recall (Soundy et al., 2007), while the remaining examined the International Physical Activity Questionnaire (IPAQ) (Faulkner et al., 2006; Vancampfort, De Hert, et al., 2017) all of which are designed to quantify moderate and vigorous intensity PA over a 7-day period. Lindamer et al. (2008) used the Actigraph model 7164 – a uniaxial device, while others (Faulkner et al., 2006; Soundy et al., 2007; Vancampfort, De Hert, et al., 2017) used a triaxial accelerometer.

Correlations between questionnaires and accelerometer derived energy expenditure was 0.37 for total minutes of self-reported MVPA (Faulkner et al., 2006) and ranged from 0.33 to 0.43 for energy expenditure (Faulkner et al., 2006; Soundy et al., 2007). Lindamer et al. (2008) reported that correlations were not significant, but did not report specific values. Vancampfort and colleagues (2017) reported correlations separately for MPA ($r_{Pearson} = 0.29$, $p = 0.23$) and VPA ($\rho_{Spearman} = -0.11$, $p = .66$) and that the IPAQ overestimated MVPA by 54%. Only (Soundy et al., 2007) produced or reported the results of Bland-Altman plots to compare data for agreement between objective measurements and questionnaires and found an overestimation of 606.5 kcal/day (Limits of agreement: -604.5 to 1817.5 kcal/day) by the 7-day recall.

These validation studies were small: Soundy et al. (2007) included 14 outpatients, 9 of which had schizophrenia; Lindamer et al. (2008) included 54 people with schizophrenia in the broader study, but only a subsample of 16 received accelerometers; Faulkner and colleagues (Faulkner et al., 2006) included 35 outpatients with schizophrenia; Vancampfort and colleagues (Vancampfort, De Hert, et al., 2017) included 19 outpatients who had experienced first episode psychosis. Overall results demonstrate similar correlations between self-report and accelerometry as observed in the general population (Craig et al., 2003; Sallis & Saelens, 2000), suggesting that well established PA questionnaires may be just as (in)accurate among people with schizophrenia as the general population. However, when reported, bias can be large and vary greatly between and within studies.

In response to Faulkner et al. (2006), Sharpe and colleagues (2006) compared accelerometry to doubly labeled water using RT3 accelerometers in a small sample of 8 individuals with schizophrenia. The correlation between the vector magnitude of accelerometry data and doubly labeled water energy expenditure was reported as not significant, but no value

was reported. The accelerometer over-predicted energy expenditure by an average of 148 kcal/day (Range: -614 to 582 kcal/day). Inactivity (counts less than 20/min) was significantly correlated with doubly labeled water energy expenditure ($r = -0.83$, $p = 0.001$). While a disagreement between the accelerometer and doubly labelled water exists, it is small compared to what Soundy et al. (2007) reported for the discrepancy between a self-report tool and accelerometer. Given the advantages of an accelerometer over doubly-labelled water, it is reasonable to use these tools as an estimate of PA related energy expenditure in this population.

These self-report tools which have been examined in samples of individuals with schizophrenia demonstrate comparable correlations as they do in the general population, suggesting that these measurement tools may generalize well to psychiatric populations. However, as pointed out by Soundy and colleagues (2014), agreement between the self-report measurements and more objective measurements has gone under-examined. The one study that did assess agreement found an average discrepancy of 606.5 kcal/day with relatively wide limits of agreement. Various symptoms, comorbidities, and psychosocial demographic factors may contribute to poor accuracy of self-report measurements when used with people with schizophrenia. This is especially relevant for self-report methods which are already at greater risk for misinterpretation than more objective measures. However, to my knowledge no work has been done to identify whether any of these factors may explain poor evidence of validity (e.g. whether higher symptom severity is associated with greater differences in tests of convergent evidence), rather literature searches to identify such records return studies evaluating the psychometrics of scales used to assess symptom severity and other outcomes.

With the rising interest in SB and LPA as a determinant of health, SB has begun to be assessed in populations with severe mental illness such as schizophrenia. The most common self-

report instrument in the studies of individuals with psychosis examining SB summarized by Stubbs et al. (2016) was the single item sitting scale of the IPAQ-SF (Craig et al., 2003). The question asks, "During the last 7 days, how much time did you spend sitting on a week day?" specifying to "Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television" in the preamble. However, only the MPA and VPA scales of the IPAQ-SF among samples of individuals with schizophrenia have been examined for evidence of validity (Faulkner et al., 2006). However, to my knowledge, and despite its apparent widespread use, a similar assessment of the sitting scale among individuals with schizophrenia has yet to be reported.

Contrary to Stubbs and colleagues' (2016) criticism of self-report questionnaires underestimating SB in individuals with psychosis, existing agreement data in the general adult population between the IPAQ-SF sitting scale and accelerometer-defined SB (≤99 counts/min) indicates that the IPAQ-SF may overestimate the amount of sedentary time by +130 min/d with large 95% limits of agreement (LoA) -275 to 536 min/d (Hagstromer et al., 2010). Additionally, bias appeared to be proportional, with greater overestimation associated with greater reported sitting time. A systematic review of validation and reliability studies with the IPAQ-SF found Spearman's correlations (ρ) of 0.07 to 0.61 with accelerometer-defined SB and test-retest reliability ranging from Spearman's ρ of 0.18 to 0.95, and intraclass correlation coefficients (ICC) ranging from 0.80 to 0.97 (Healy et al., 2011). Evaluating the sitting scale of the IPAQ-SF as a measure of SB will evaluate whether agreement with direct measure of SB follows similar patterns of error as scores from the general population or aligns with Stubbs and colleagues' (2016) findings.

An alternative to assessing existing behavioural measurement tools for measuring behaviour is to develop new tools. Concurrent to the work done on this dissertation, an effort to develop a new questionnaire to measure MBs in people with mental illness was commenced (Rosenbaum & Ward, 2016): the Simple Physical Activity Questionnaire (SIMPAQ). The SIMPAQ takes approximately 15 minutes for interviewers to administer (Schilling et al., 2018). Recently the primary validation paper for SIMPAQ scores in individuals with mental illness has been published (Rosenbaum et al., 2020). Among the subgroup of individuals with schizophrenia, SIMPAQ MVPA correlated with accelerometry at $\rho_{spearman}$ 0.04, $p = 0.66$, n =130 and SB $\rho_{spearman} = 0.26$, $p < 0.01$, n =140. For the full sample, mean bias of Bland-Altman analyses showed small differences with the SIMPAQ overestimating MVPA (~20min/week) and underestimating SB (~-3h/week; estimated from figures); LoA were wide with MVPA ranging from approximately -2.5h to 3h/week and SB from -11h to 5h/week (estimated from figures). Agreement was not evaluated by patient subgroup. Further examination of the measurement properties of existing tools such as the IPAQ and new tools such as the SIMPAQ is required in considering the value of self-report measures of physical activity in this clinical population.

1.1.4 Purpose

This dissertation will undertake four studies using a data set I collected comprised of 130 individuals with schizophrenia and schizoaffective disorder (See Table 1.1) who wore accelerometers over a 7-day observation period and subsequently responded to the IPAQ-SF that assessed the same 7-day time frame. Table 1.2 summarizes the purpose and methods used for each of the four studies. Study 1 (Chapter 2) measures the agreement between accelerometers and the IPAQ as a tool for assessing MVPA in individuals with schizophrenia and Study 2

(Chapter 3) will do the same for SB. Study 1 and 2 have already been published (Duncan et al., 2019; Duncan, Arbour-Nicitopoulos, et al., 2017). Having found significant disparity between the accelerometer and IPAQ, Study 3 (Chapter 4) examines whether individual differences in demographic data, symptom severity, health status, and cognition are associated with differences between measurement tools in order to identify potential sources of error. Study 3 also attempts to adjust IPAQ-SF scores using regression calibration to better reflect time spent in MVPA and SB. Having evaluated whether the IPAQ-SF MVPA and SB scores can be used as a measure of MB time use in individuals schizophrenia, Sallis and colleagues' (2000) framework would suggest revisiting what health outcomes are associated with MVPA and SB behaviours. However, given the evidence presented in Study 1-3 that the IPAQ-SF cannot be used as a measure of time use, device-based measures such as accelerometry are likely necessary to characterize the MBs of daily life in individuals with schizophrenia. In light of this, Study 4 (Chapter 5) assesses whether the accelerometry protocols were well adhered to and whether individual factors may contribute poor adherence, leading to possible bias during data collection; subsequently the composition of waking day MB were evaluated for differences across the observation period and whether the exchanging of time spent in one MB for another is associated with participant descriptors, such as health outcomes. Together these four studies contribute to MB measurement research with: (a) quantifiable evidence indicating that IPAQ scores are a problematic measure of time use in individuals with schizophrenia, and that alternative measurement tools are necessary for many applications, (b) insight in to plausible sources of error when using a self-report method of measurement, (c) whether a commonly used accelerometry protocol leads to systematic bias in MB measurements in individuals with

schizophrenia, and (d) an exploratory assessment of how the spectrum of waking day MBs, from SB to VPA, are associated with participant descriptors and health outcomes.

1.1.5 Tables

Table 1.1 Summary of Participant Characteristics

	Enrolled (n = 130)	Received Accelerometer (n = 113)	Accelerometer Protocol Adherence (n = 101)
Participant Descriptors			
Male : Female	80 (61%):50 (38.5%)	68 (60.2%):45 (39.8%)	60 (59.4%):41 (40.6%)
Current smokers	63 (48.5%)	55 (48.7%)	49 (48.5%)
Age [years]	40.1 (11.5)	41.0 (11.7)	41.5 (11.7)
Schizophrenia : Schizoaffective	89 (68.5%):41 (31.5%)	76 (67.3%):37 (32.7%)	68 (67.3%):33 (32.7%)
Ethnicity			
African origin/Black	21 (16.2%)	19 (16.8%)	19 (18.8%)
Asian/South Asian	20 (15.4%)	16 (14.2%)	14 (13.9%)
Caucasian/White	74 (56.9%)	67 (59.3%)	57 (56.4%)
Other (Including Multiple Ethnicities)	15 (11.5%)	11 (9.7%)	11 (10.9%)
Education			
Some High School (no diploma)	28 (21.5%)	21 (18.6%)	17 (16.8%)
High School Diploma	36 (27.7%)	31 (27.4%)	24 (23.8%)
At least some Postsecondary	62 (47.7%)	59 (52.2%)	58 (57.4%)
Trade School	4 (5.4%)	2 (1.8%)	2 (2.0%)
Employment			
Full-Time	3 (2.3%)	2 (1.8%)	2 (2.0%)
Part-Time	36 (27.7%)	35 (31.0%)	31 (30.7%)
Student	7 (5.4%)	4 (3.5%)	3 (3.0%)
Unemployed	77 (59.2%)	70 (61.9%)	63 (62.4%)
Other (e.g. retired, volunteer)	7 (5.4%)	2 (1.8%)	2 (2.0%)
Living arrangements			
Independent	62 (47.7%)	55 (48.7%)	51 (50.5%)
Family/Spouse	8 (6.2%)	7 (6.2%)	6 (5.9%)
Group (meals provided)	18 (13.8%)	14 (12.4%)	11 (10.9%)
Group (no meals provided)	42 (32.3%)	37 (32.7%)	33 (32.7%)

Health & Well-Being Descriptors			
BMI (kg/m^2)*			
Mean	31.7 (8.1)	31.5 (8.4)	31.2 (7.8)
Underweight (BMI<18.5)	1 (.8%)	1 (0.9%)	1 (1.0%)
Healthy Weight (18.5≤BMI<25)	21 (16.2%)	19 (16.8%)	18 (17.8%)
Overweight (25≤BMI<30)	35 (26.9%)	32 (28.3%)	29 (28.7%)
Obese (30≤BMI<40)	72 (55.4%)	60 (53.1%)	52 (51.5%)
Waist Circumference Category**			
Low:High	29 (22.3%):97 (74.6%)	25 (22.1%):84 (74.3%)	24 (23.8%) :73 (72.3%)
SF-12 Health Survey			
Physical Composite Score	29.6 (5.0)	29.6 (5.1)	29.9 (4.9)
Mental Composite Score	51.1 (11.0)	52.0 (11.1)	51.9 (11.1)
Symptom Severity			
BPRS-A total	34.3 (8.5)	34.2 (8.5)	33.5 (7.3)
CGI-S	3.5 (1.1)	3.5 (1.1)	3.4 (1.1)
AES	32.0 (7.9)	31.7 (7.9)	31.2 (7.9)

Note: Values represent mean (sd)/n(%) where appropriate; * One participant opted out of being weighed, **4 opted out of waist measurement, low vs high designation based on International Diabetes Foundation sex and ethnicity-based cut points. BPRS-A = Brief Psychiatric Rating Scale 18-item Anchored version (Woerner et al., 1988), CGI-S = Clinical Global Impression Severity Scale (Guy, 1976b), AES = Apathy Evaluation Scale (Marin et al., 1991), higher scores represent more severe symptoms for all scales. CPZ = Chlorpromazine Equivalents (Gardner et al., 2010), BMI = Body Mass Index. Higher cognitive scores represent better functioning. Higher cognitive scores and SF-12 scores represent better functioning and wellbeing respectively; SF-12 composite scores are norm based with 50 representing the mean in the general population.

Table 1.2 *Summary of Study Purposes and Methods by Chapter*

	Chapter 2	Chapter 3	Chapter 4	Chapter 5
Purpose	Compare IPAQ-SF MVPA scores to accelerometer criterion for the same observation period[1,2] Assess reliability of IPAQ-SF MVPA 4-weeks apart[1,2]	Compare IPAQ-SF SB scores to accelerometer criterion for the same observation period[1,2] Assess reliability of IPAQ-SF sitting 4-weeks apart[1,2]	Identify individual factors associated with difference between IPAQ-SF and accelerometer for MVPA & SB[3] Test adjustments to IPAQ scores to better align with accelerometry derived estimates[4]	Assess factors associated with accelerometer wear adherence[5] Compare daily MB time use derived from accelerometry across 7-day observation period[5,6] Assess associations between time use exchange (e.g. replacing SB for LPA) and participant descriptors and health factors[6]
Methods	[1] Bland-Altman plots to assess mean difference & LoA [2] Correlation between measurements	[1] Bland-Altman plots to assess mean difference & LoA [2] Correlation between measurements	[3] Regression analysis of difference between measurement tools and [4] 5-fold repeated resampling to test regression calibration equations against holdout sample	[5] Mixed model regression [6] Isotemporal substitution analysis
Novel Contributions	IPAQ-SF MVPA scores as a measure of time use have not been evaluated for agreement (previous data correlational)	IPAQ-SF sitting item as an indicator of time spent in SB in individuals with	Evaluates whether IPAQ-SF is more accurate for some individuals	Factors associated with accelerometer compliance have not been reported in

	schizophrenia had not been validated	Adjustments based on these factors would improve IPAQ-SF scores as an estimate of time use	individuals with schizophrenia Assess daily MB behaviour for influences of bias Effects of MB exchange has also not been performed

Note: Superscripts correspond between purpose and methods

Chapter 2: (In)Accuracy of Self-Report Physical Activity Assessment Compared to Accelerometry

2.1 Introduction

In addition to the psychological symptoms of schizophrenia, individuals with the disorder suffer from high rates of obesity, diabetes, and cardiovascular disease compared to the general population (Dixon et al., 2000; Hennekens et al., 2005; Manu et al., 2015). These physical comorbidities contribute, in part, to the 15-25 year reduced life expectancy for people with schizophrenia (Laursen et al., 2012) as well as reduced quality of life (Faulkner, Cohn, Remington, et al., 2007; Foldemo et al., 2014; Guo et al., 2013; Sugawara et al., 2013). Physical activity (PA) is well established as an efficacious method of preventing and managing these physical illnesses in the general population (Orozco et al., 2008; Shaw et al., 2007; Thomas et al., 2009; Warburton et al., 2010), and greater PA and less SB have been associated with better quality of life (Costa et al., 2018). Testing methods for increasing PA among individuals with schizophrenia is therefore warranted (McNamee et al., 2013; Vancampfort, Rosenbaum, et al., 2016). Essential to this is the need to accurately measure PA within the schizophrenia population in order to identify the prevalence of physical (in)activity, assess the effectiveness of PA interventions, and examine relationships between physical (in)activity and other outcomes of interest to researchers and clinicians (Vancampfort, Rosenbaum, et al., 2016).

With the advent of wearable accelerometers it is now possible to objectively measure PA in the field, with good reliability and validity (L. A. Kelly et al., 2013), without direct observation of participants (LaPorte et al., 1985). However, to obtain

accurate results, accelerometers present some burden to participants such as having to adhere to wearing the device for the duration of an assessment period (e.g., 1 week). Additionally, accelerometers are costly relative to self-report measures of PA. When objective measures such as accelerometry are not feasible, a common alternative is to use a subjective, self-report measure of PA (Soundy et al., 2014).

The International Physical Activity Questionnaire (IPAQ) (Craig et al., 2003) is one such subjective measure of PA. The IPAQ asks participants to self-report on the frequency, intensity (moderate, vigorous, walking, sitting) and duration over the past 7 days they have engaged in PA. The IPAQ is available as both a long and short-form paper survey in multiple languages, which allows for cross-cultural comparisons. The short-form version of the IPAQ has been previously assessed for its test-retest reliability and criterion validity with accelerometry among 35 outpatients with schizophrenia over a 1-week period (Faulkner et al., 2006). The authors reported test-retest reliability in the sample of Spearman's $\rho = 0.68$ (95% CI: 0.45–0.83) for minutes of moderate to vigorous PA (MVPA). Reliability of moderate PA (MPA) was $\rho = 0.50$ (95% CI: 0.20–0.70) and $\rho = 0.69$ (95% CI: 0.46–0.83) for vigorous PA (VPA). Criterion validity for MVPA with accelerometry was $\rho = 0.37$ (95% CI: 0.04–0.63). Overall, the authors concluded the psychometric properties of the IPAQ observed in this sample of individuals with schizophrenia was comparable to the pooled values reported by Craig et al. (2003) in the general population (reliability: pooled $\rho = 0.76$, 95% CI 0.73–0.77; validity: pooled $\rho = 0.30$, 95% CI 0.23–0.36). After ten years of use in the field, now is a timely opportunity to revisit the psychometric properties of the IPAQ within the schizophrenia population. This is particularly pertinent given efforts to develop a new self-report, physical activity

measure for use among individuals with serious mental illness (Rosenbaum & Ward, 2016).

Understanding the reliability of the IPAQ over a longer period may be informative for exploring stability of physical activity in epidemiological research and estimating how much variation could be attributed to measurement error. Additionally, when initially evaluated by Faulkner et al. (2006), the IPAQ was aided by a structured recall lead by the experimenter, rather than being provided to the participant in questionnaire form, as is intended. An unstructured administration where the participant is simply presented with the questionnaire may result in different psychometric characteristics, especially in populations where cognitive impairment is prevalent.

Cognitive deficits among people with schizophrenia are common. It has been suggested that 75-85% of people with schizophrenia have some form of significant cognitive impairment (Reichenberg et al., 2006). Deficits in memory are common (Keefe & Fenton, 2007; Reichenberg et al., 2006), and may significantly impact participant's ability to accurately recall activities over an extended period without assistance. Furthermore, deficits in attention and executive function (Keefe & Fenton, 2007; Reichenberg et al., 2006) in conjunction with deficits in reading comprehension (Arnott et al., 2016; Hayes & O'Grady, 2003) may impact the reliability and validity of any self-report questionnaire.

An alternative to the IPAQ and other self-report questionnaires is having participants specifically recount their activities over the previous day through a detailed structured recall. Ostensibly having to remember only the most recent day in detail, rather than over a 7-day period, may be easier and thus provide a more accurate recall. In the

general population, interview based protocols have demonstrated higher criterion validity for measuring PA than self-report measures (Sallis & Saelens, 2000). However, one potential disadvantage is that a 24-hour recall may not be as sensitive to regular activity patterns; for example, the previous day may not be representative of an individual's typical week.

In order to identify and develop more effective subjective tools in the measurement of PA among people with schizophrenia, the current study was undertaken to assess the psychometric properties of the IPAQ and a 24-hour structured recall compared to accelerometry. Specifically, data from a pre-existing prospective study was available to assess 1) the test-retest reliability of the self-administered short-form IPAQ over a 4-week time period, 2) the criterion validity of a self-administered IPAQ compared to accelerometry, and 3) the criterion validity of a 24-hour PA recall compared to accelerometry, in a sample of people with schizophrenia. It was hypothesized that longer periods between assessments would reduce the reliability when using the IPAQ, and that the self-administration of the IPAQ would result in lower levels of criterion validity than previously examined in Faulkner et al.'s (2006) validation work. Finally, the 24-hour recall was expected to demonstrate stronger validity compared to the IPAQ.

2.2 Methods

2.2.1 Participants

Research ethics boards at the Centre for Addiction and Mental Health in Toronto and the University of Toronto approved the larger prospective study. To be included in the study, participants were required to: 1) be age 18-64 years (in line with the Canadian

Physical Activity Guidelines recommendations for adults (Canadian Society of Exercise Physiology, 2012)), and 2) have a diagnosis of schizophrenia or schizoaffective disorder. The Mini-International Neuropsychiatric Interview (Sheehan et al., 1998) was administered to confirm diagnosis. Participants were excluded if they had: 1) been hospitalized over the past 12 months for angina pectoris, myocardial infarction, or cardiac surgery of any kind; and/or 2) uncontrolled hypertension (defined as blood pressure > 140/90). If eligible, participants provided written consent prior to commencing the study. Capacity to consent was assessed immediately after using the MacArthur Competence Assessment Tool for Clinical Research (MacCAT-CR; Appelbaum and Grisso, 2001).

2.2.2 Data Collection

PA levels were assessed during the intake session with the IPAQ-SF (See: Appendix A.1) and again four weeks later. Participants received an Actigraph wGT3X+ accelerometer three weeks after intake and were instructed to wear the device for 7 days over their right hip and returned at the end of the study. If participants wore the accelerometer the day before the week 4 testing session, they were guided through a 24-hour recall of their PA from the previous day using a semi-structured interview, to recall the duration (lasting at least 10 minutes), type, and intensity of the physical activities they engaged in while awake and wearing the accelerometer, in blocks of 2 to 3 hours throughout the day. Individuals were prompted to consider the IPAQ definitions of moderate and vigorous physical activities throughout. Minutes spent in moderate and/or vigorous PA were then derived for the day.

2.2.3 Analysis

IPAQ-SF data were analyzed based on scoring guidelines: reported MPA and VPA time variables below 10 minutes are scored as 0 minutes and those exceeding 180 minutes are truncated 180 minutes (IPAQ, 2005). Days of activity per week were multiplied by time variables to calculate weekly minutes of PA at each intensity. Accelerometry data were analyzed using Actigraph's *Actilife* software (v6.12). To compensate for participants who did not remove the accelerometer before sleeping nor kept a record of their sleep and wake times, data from 0h00 to 5h59 from each day were not analyzed. Wear time was calculated using Choi and colleagues' (2011) algorithm. To be considered a valid day, the accelerometer needed to register 600 minutes of wear time (Troiano et al., 2008). A valid week was at least four valid days of the 7-day wear period (Troiano et al., 2008; Trost et al., 2005). Troiano and colleagues' (2008) adult PA cut-off points were used to determine time spent in moderate (≥2020 counts per minute) and vigorous (≥5999 counts per minute) intensity PA, and only bouts of PA lasting at least 10 minutes were included per Canadian PA guidelines (Canadian Society of Exercise Physiology, 2012), and in line with the IPAQ-SF operationalization of PA. Average daily PA was multiplied across a 7-day week to calculate weekly PA for comparison against the IPAQ-SF. If valid (n =101), accelerometry data from the day prior to the week 4 testing session was compared to the 24-hour recall.

As physical activity data were nonparametric, Spearman's rank correlation coefficient, with missing values excluded pairwise, was calculated using SPSS (v22) to assess IPAQ-SF test-retest reliability between week 1 and week 4, and criterion validity between the accelerometry data at week 3 and PA self-reported in the week 4 IPAQ

assessment and, where available, the 24-hour recall. A Bonferroni correction of α = 0.0167 was applied to the IPAQ and to the 24-hour recall criterion validity for MPA, VPA and MVPA. Confidence intervals were calculated by applying Fisher's z-transformation. These correlations were calculated again with inpatients and outliers (>+/- 3 SD) removed to assess reliability changes with these limitations. Bland-Altman plots (Bland & Altman, 1986, 2003) were generated for MVPA for the test-retest reliability and criterion validity assessments. Agreement between the week 1 and 4 administration of the IPAQ as well as week 4 IPAQ and accelerometry as to whether or not participants met guidelines of 150 minutes per week of MVPA was calculated using Cohen's κ.

2.3 Results

2.3.1 Demographics

Table 2.1 summarizes demographics of the study participants (See Appendix A.2 for data collection form).

2.3.2 IPAQ Retest Reliability

Table 2.2 summarizes the amount of PA reported in the week 1 and week 4 IPAQ-SF administrations. An initial 127 participants completed the week 1 IPAQ-SF, while 109 participants (86% of initial sample) completed the week 4 IPAQ-SF assessment. Table 2.3 reports the 4-week test-retest reliability of the IPAQ-SF. The Spearman correlation coefficient for minutes of MVPA per week was 0.47, $p < .001$, n = 107 and remained unaffected when inpatients were excluded, $\rho = 0.47$, $p < .001$, $n = 102$,

as well as when both inpatients and outliers were excluded, ρ = 0.46, $p < .001$, $n = 97$. The correlation coefficient was higher for VPA (ρ = 0.51, $p < .001$, $n = 104$) than either MPA (0.38, $p<.001$, $n = 91$) or combined MVPA. Figure 2.1 is a Bland-Altman plot that demonstrates the agreement between the week 1 and week 4 IPAQ-SF administration. Average disagreement across the week was -7.8 minutes of MVPA (-4%), with 95% limits of agreement ranging from -608.7 minutes (-270%) to 593.0 minutes (262%). Table 2.4 reports the number of participants meeting 150 minutes of MVPA per week across IPAQ-SF administrations. Proportionate agreement between IPAQ-SF administrations was .70 and Cohen's κ = .40, $p < .001$, n = 107.

2.3.3 IPAQ Validity

Table 2.2 summarizes the PA measured by accelerometry. Valid data were obtained from 79.5% of the 127 participants. Table 2.5 summarizes Spearman correlation coefficients between the week 4 IPAQ-SF and the week 3 accelerometry. Only the combined MVPA outcome of the two measures was significantly correlated ρ = .30, $p = .003$, $n = 97$. Removing inpatients, ρ = 0.31, $p = 0.003$, $n = 93$, or both inpatients and outliers ρ = .32, $p = .002$, $n = 91$, did not change this result. A Bland-Altman plot of agreement between the IPAQ-SF and accelerometry (Figure 2.2) indicated that the IPAQ-SF under reported by an average disagreement of -119.2 minutes (-52%), with 95% limits of agreement ranging from -1017.1 to 778.7 minutes (-301% to 198%) over the week. This represents an average disagreement of -9.4 minutes per day, with 95% limits of agreement ranging from -113.7 to 95.0. Table 2.6 reports the number of participants meeting 150 minutes of MVPA per week based on week 4 IPAQ-SF and accelerometry

data. Proportionate agreement between IPAQ at week 4 and accelerometry was .56 and Cohen's $\kappa = .123$, $p = .21$, n = 97.

2.3.4 24-hour Recall Criterion Validity

Table 2.7 details the PA reported in the 24-hour recall compared to accelerometry from the previous day. Ninety participants (70.8% of the initial sample) responded to the 24-hour recall, of which 87 were able to provide a recall of their PA over the previous day, and 101 had valid accelerometry data for the same period. MVPA was moderately correlated significantly between the two measures $\rho = .27$, $p = .012$, $n = 83$, and this remained largely unaffected by removing inpatients $\rho = .23$, $p = .04$, $n = 79$, and outliers $\rho = .27$, $p = .02$, $n = 74$. VPA was correlated significantly at $\rho = .30$, $p = .006$, $n = 80$, as was MPA at $\rho = .27$, $p = .014$, $n = 83$. The Bland-Altman plot of agreement between 24-hour recall IPAQ-SF and accelerometry (Figure 2.3) indicated an average disagreement of -20.4 minutes (-133%) per day, with 95% limits of agreement ranging from -128.9 to 88.2 minutes (-372% to 106%).

2.4 Discussion

Overall, results from the current study confirm and expand the psychometric results of the IPAQ-SF previously reported by Faulkner et al. (2006). Compared to this previous study, 4-week test-retest reliability of the IPAQ-SF was slightly lower than 1-week retest reliability in all categories of PA, but not outside of confidence interval estimates. These results indicate that, as expected, the IPAQ-SF may be less reliable over longer periods. The present study also expands previous knowledge by presenting the direction and range of bias between the test-retest reliability of the IPAQ-SF. While the

average bias is reasonably low (< 8 minutes of MVPA over a week), the range of differences is exceptionally large. In 95% of cases, absolute agreement between assessments may differ by roughly 600 minutes in either direction. Despite this the IPAQ-SF was consistent in classifying whether individuals were meeting weekly MVPA guidelines or not across administrations with a 70% agreement rate and a κ coefficient approaching Cohen's proposed .41 acceptable level (McHugh, 2012). Although PA tended to be relatively stable over time, it is not possible to conclude whether any differences reflect issues with recall or differences in PA over time. Future reliability assessments may seek to examine accelerometry data across weeks where the IPAQ-SF is administered to provide an objective assessment as to whether lower reliability can be explained by variation in objectively measured PA. The results obtained for reliability here are, however, much lower than the pooled reliability of the IPAQ-SF when used with the general population (Craig et al., 2003).

Contrary to the hypothesis, criterion validity for the IPAQ-SF when self-administered (MVPA $\rho = 0.30$) was similar to both the semi-structured interview method used by Faulkner et al. (2006) ($\rho = 0.33$) as well as pooled results from the general population ($\rho = 0.30$) (Craig et al., 2003), indicating that for a predominantly outpatient sample with schizophrenia, the IPAQ-SF may serve as a suitable PA surveillance tool. On the other hand, while this level of validity is comparable to other self-reported measures of PA (Sallis & Saelens, 2000; Skender et al., 2016), it is still relatively inaccurate compared to an objective PA measure. Furthermore, use of the IPAQ-SF with this population resulted in poor discrimination between MPA and VPA. Validity coefficients for MPA and VPA individually were both $\rho=0.10$. This provides further

evidence that while patient self-reports may able to account for total levels of MVPA, they may not be able to discriminate between moderate and vigorous intensity PA (Faulkner et al., 2006; Sallis & Saelens, 2000; Skender et al., 2016).

In regard to difference, the IPAQ-SF underestimated MVPA by an average of 119.2 minutes across the week compared to accelerometry. Additionally, there is a possibility of even greater inaccuracy as indicated by the 95% limits of agreement ranging by 1795.8 minutes of MVPA. The IPAQ-SF and accelerometer agreed 56% of the time as to whether participants met MVPA guidelines suggesting that the IPAQ may have some use for classifying individuals with schizophrenia as "active" or "inactive" in epidemiological research. Overall, while the IPAQ-SF may agree with accelerometry in regards to rank order, as indicated by the results of the Spearman's correlation coefficients, the absolute agreement between the two methods of PA measurement seems relatively poor.

Criterion validity of the 24-hour recall method ($\rho = 0.27$) was also comparable to the IPAQ-SF for overall MVPA level, but was stronger for MPA ($\rho = 0.27$ vs $\rho = 0.10$) and VPA ($\rho = 0.30$ vs $\rho=0.10$). As the semi-structured interview allows for the rater to evaluate the intensity of reported activities, this may compensate for patients' inability to differentiate MPA from VPA, which is indicated by the size of the correlation between MPA and VPA from the 24-hour recall approaching that measured by accelerometry over the same time span. This ability to differentiate between intensities of PA, as well as being able to record the types of PA engaged in may be advantageous for some research questions, but confers no added benefit for determining overall level of MVPA compared to the IPAQ-SF. Given these results, assessing the reliability of the 24-hour recall method

is warranted to determine any other psychometric advantages this method may have over the IPAQ-SF and other subjective methods of assessing PA.

With regards to agreement with accelerometry, the 24-hour recall and IPAQ-SF were relatively similar. On average the 24-hour recall underestimated MVPA by 20.4 minutes per day. Using the IPAQ-SF across 7 days, MVPA is underestimated by 9.4 minutes per day. However, the 95% agreement limits of the 24-hour recall range from 129 minutes of underestimation to 88 minutes of overestimation. By comparison, the 95% agreement limits of the IPAQ-SF range from 145 minutes of underestimation to 111 minutes of overestimation. Thus, the 24-hour recall method by-and-large suffers from similar issues of absolute agreement with objective measures as the IPAQ-SF, but may benefit from slightly narrower limits of agreement than the IPAQ-SF.

Overall, the present study indicates that while the self-reported IPAQ-SF may have reasonable rank-order criterion validity for overall MVPA, semi-structured interviews assisting a 24-hour recall may improve criterion validity for MPA from VPA. However, Bland-Altman plots indicate a wide agreement range for both the IPAQ-SF and 24-hour recall compared to accelerometry. Over the 4-week period, test-retest reliability was lower than previously reported 1-week test-retest reliability. Criterion validity for both tools tested were comparable to each other, as well as other tools used to assess PA in the general population (Craig et al., 2003; Sallis & Saelens, 2000; Skender et al., 2016), indicating that both may be used for surveillance particularly when minutes of MVPA is the outcome of interest. As criterion validity of the IPAQ-SF and 24-hour recall, and reliability of the IPAQ-SF over a 4-week period among people with schizophrenia remain relatively low, objective measures of PA, such as accelerometry,

remain recommended for intervention work. With the decreasing costs of accelerometry and the increasing availability of commercial trackable devices, it is likely that objective assessment of physical activity will become more accessible.

In building on the initial work by Faulkner and colleagues (2006), the present study provides expanded psychometric details of the IPAQ-SF questionnaire when used with individuals with schizophrenia. Notably, this study presents test-retest reliability over a longer (4 week) period, criterion validity for MPA and VPA subscales, and agreement data for MVPA test-retest reliability and criterion validity. This study is further strengthened by including a larger sample, and comparing criterion validity of the IPAQ-SF to a semi-structured 24-hour recall method as one alternative. It remains to be seen whether the development of new self-report measures will substantively demonstrate greater psychometric properties than existing tools such as the IPAQ-SF.

2.5 Tables

Table 2.1 *Summary of Participant Demographics*

Demographic	Enrolled (n = 130)	Completed (n = 113)
Male: Female	80:50	68:45
Current smokers	63	55
Outpatient: Inpatient	124:6	108:5
Mean Age (SD)[years]	40.1 (11.5)	41.0 (11.7)
Ethnicity		
African Descent	21	19
Asian/South Asian	20	16
White	74	67
Multi-ethnic	5	5
Other	10	6
Symptom Severity		
BPRS-A mean score (SD)	34.3 (8.5)	34.2 (8.5)
CGI-S mean score (SD)	3.5 (1.1)	3.5 (1.1)
AES mean score (SD)	32.0 (7.9)	31.7 (7.9)

CPZ equivalents mean (SD)[mg]	788.5 (1237.2)	748.1 (1216.6)
BMI (kg/m^2)*		
Mean (SD)	31.7 (8.1)	31.5 (8.4)
Underweight (BMI<18.5)	1	1
Normal Weight (18.5<BMI<25)	21	19
Overweight (25<BMI<30)	35	32
Obese (BMI>30)	72	60
Education		
Some High School (no diploma)	31	23
High School Diploma	37	31
At least some Postsecondary	62	59
Employment		
Full-Time	3	2
Part-Time	36	33
Student	7	6
Unemployed	77	65
Other (e.g. retired, volunteer)	7	7

Note: Higher score represent more severe symptoms. * One participant opted out of being weighed. BPRS-A = Brief Psychiatric Rating Scale 18-item Anchored version (Woerner et al., 1988), CGI-S = Clinical Global Impression Severity Scale (Guy, 1976b), AES = Apathy Evaluation Scale (Marin et al., 1991), CPZ = Chlorpromazine Equivalents (Gardner et al., 2010), BMI = Body Mass Index.

Table 2.2 *Physical Activity over Two 1-Week Periods*

	IPAQ						Accelerometer (Bouts >10min)					
	Mean	Median	SD	n	Min	Max	Mean	Median	SD	n	Min	Max
Week 1												
Minutes MPA	164.9	60	235.7	116	0	1260						
Minutes VPA	55.1	0	135.9	124	0	1080						
Minutes MVPA	204.5	90	277.4	127	0	1290						
Week 4												
Minutes MPA	180.3	85	256.6	102	0	1260	222.4	152.6	221.9	100	0	1118.8
Minutes VPA	56.4	0	136.7	106	0	1080	23.8	0	66.4	100	0	412
Minutes MVPA	223.6	125	315.9	109	0	2070	287.1	200.7	303.3	101	0	2060.3

Table 2.3 *IPAQ Test-Retest Reliability: Week 1 vs. Week 4*

	ρ	95% CI	p	n
Weekly Minutes MPA	0.377	.19-.54	<.001*	91
Weekly Minutes VPA	0.514	.36-.64	<.001*	104
Weekly Minutes MVPA	0.466	.30-.60	<.001*	107

Note: * indicates statistical significance at corrected α=0.0167

Table 2.4 *Agreement on Meeting ≥150 Minutes MVPA/week Between Week 1 and Week 4 IPAQ-SF*

	IPAQ Week 4		
IPAQ Week 1	Yes	No	Total
Yes	32 (30%)	13 (12%)	45 (42%)
No	19 (18%)	43 (40%)	62 (58%)
Total	51 (48%)	56 (52%)	107

Table 2.5 *IPAQ Criterion Validity: Week 4 vs. 7-day Accelerometry*

	ρ	95% CI	p	n
Weekly Minutes MPA	0.098	-.11-30	0.356	90
Weekly Minutes VPA	0.098	-.11-.30	0.351	93
Weekly Minutes MVPA	0.298	.11-.47	0.003*	97

Note: * indicates statistical significance at corrected α = 0.0167. All physical activity calculated based on bouts greater than 10min as per Table 2.2.

Table 2.6 *Agreement on Meeting ≥150 Minutes MVPA/week Between Week 4 IPAQ-SF and Accelerometry*

	IPAQ Week 4		
Accelerometer	Yes	No	Total
Yes	29 (30%)	27 (28%)	56 (58%)
No	16 (16%)	25 (26%)	41 (42%)
Total	45 (46%)	52 (54%)	97

Note: All physical activity calculated based on bouts greater than 10min as per Table 2.2.

Table 2.7 *24-hour Recall Criterion Validity*

	24h Recall						Accelerometer (Bouts >10min)						Correlation			
	Mean	SD	Median	n	Min	Max	Mean	SD	Median	n	Min	Max	ρ	95% CI	p	n
Minutes MPA	10.3	30.9	0	84	0	180	24.8	34.3	14	101	0	157	0.273	.06-.46	0.014*	80
Minutes VPA	3.3	22.0	0	87	0	195	3.3	11.6	0	101	0	92	0.298	.09-.48	0.006*	83
Minutes MVPA	13.2	37.8	0	87	0	195	33.7	46.7	20	101	0	312	0.274	.06-.46	0.012*	83

Note: * indicates statistical significance at corrected $\alpha = 0.0167$

2.6 Figures

Figure 2.1 *Bland-Altman Plot for the International Physical Activity Questionnaire (IPAQ) Administration at Week 1 and Week 4*

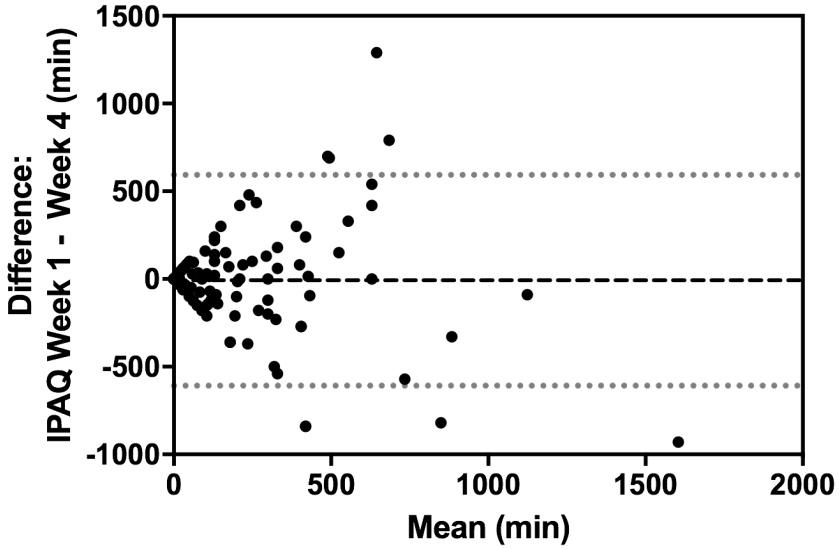


Figure 2.2 *Bland-Altman Plot for the International Physical Activity Questionnaire (IPAQ) Administration at Week 4 and Accelerometry Data over the Same 7-day Period*

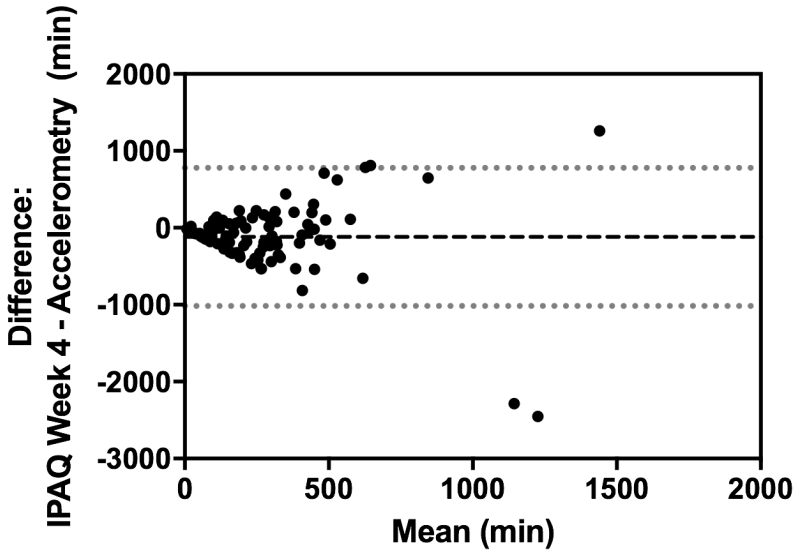
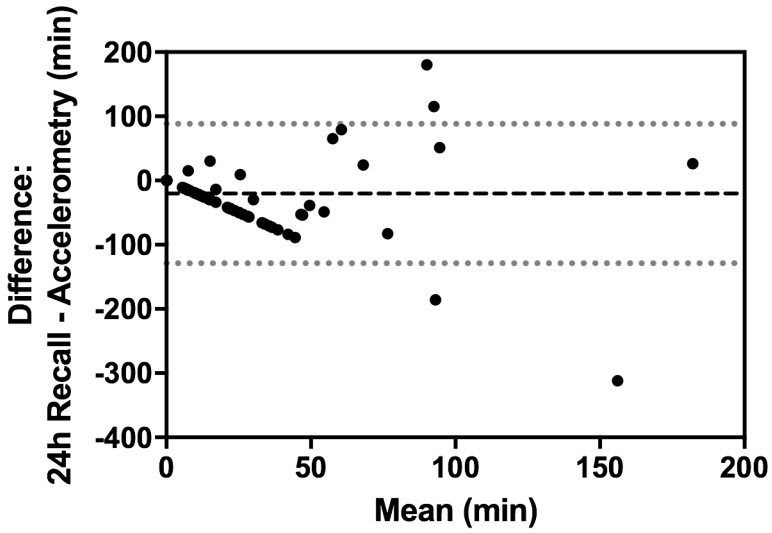


Figure 2.3 *Bland-Altman Plot for the 24-hour Recall Method and Accelerometry Data for the Previous Day*



Chapter 3: (In)Accuracy of Self Report Sedentary Behaviour Assessment Compared to Accelerometry

3.1 Introduction

Individuals with schizophrenia suffer from higher rates of obesity, diabetes, and cardiovascular disease compared to the general population (Correll et al., 2017; Dixon et al., 2000; Hennekens et al., 2005; Manu et al., 2015; Vancampfort et al., 2015; Vancampfort, Correll, et al., 2016) While moderate-to-vigorous physical activity (PA) is an important component of health promotion and weight management, meta-analysis has demonstrated that greater amounts of daily sedentary behaviour (SB) – any waking activity characterized by an energy expenditure less than 1.5 metabolic equivalents in a sitting, reclining or lying posture (Tremblay, Aubert, et al., 2017) – are associated with higher rates of cardiovascular disease, diabetes, and all-cause mortality (Biswas et al., 2015; Warren et al., 2010). This association appears to be independent of moderate-to-vigorous PA (MVPA) (Biswas et al., 2015). Even individuals who are currently meeting physical activity guidelines may receive cardio-metabolic health benefits from interrupting sitting time with short breaks (Owen et al., 2010; Thorp et al., 2011). As such, studies have begun to demonstrate an association between SB and health outcomes such as body mass index among individuals with schizophrenia as well (Bueno-Antequera et al., 2017; Vancampfort, De Hert, et al., 2014; Vancampfort, Probst, Knapen, et al., 2012). Therefore, while continuing to test interventions to increase moderate to vigorous PA among individuals with schizophrenia is still important (McNamee et al., 2013; Vancampfort, Rosenbaum, et al., 2016), physically active individuals with schizophrenia may accrue additional health benefits by reducing

prolonged sitting time. Furthermore, reducing SB among inactive individuals may be easier than promoting more vigorous PA while still accruing some health benefit.

A recent meta-analysis demonstrated that individuals with psychosis engage in high levels of SB (Stubbs, Williams, et al., 2016). Across 13 studies the authors reported a pooled effect size of 11h of SB per day in this population. Among the four studies that compared individuals with psychosis to healthy controls, individuals with psychosis were engaging in 2.8h of more SB. These numbers increase to 12.6h of SB and 2.9h more than healthy controls when only objective measures of SB are used. The authors suggested that self-report questionnaires may be largely underestimating the amount of SB that individuals with psychosis are engaging in, though no studies included in the review used both objective and self-report methods simultaneously to evaluate this hypothesis. A subsequent meta-analysis of SB in schizophrenia, bipolar disorder and major depressive disorder similarly found lower self-reported SB compared to when measured objectively (Vancampfort, Firth, et al., 2017). While the objective measurement of SB through accelerometry is the gold standard for assessment in the field, self-report will likely continue to be used for pragmatic reasons, particularly in epidemiological research.

The most common self-report instrument in the studies summarized by Stubbs et al. (2016) was the single item sitting scale of the International Physical Activity Questionnaire Short Form (IPAQ-SF) (Craig et al., 2003). The question asks, “During the last 7 days, how much time did you spend sitting on a weekday?” specifying to “Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television” in the preamble. Reliability and agreement have been previously

characterized between the IPAQ-SF and accelerometry among samples of individuals with schizophrenia (Duncan, Arbour-Nicitopoulos, et al., 2017; Faulkner et al., 2006), but only in regard to the moderate and vigorous PA scales of the IPAQ-SF. Additionally, Firth and colleagues (Firth et al., 2018) have expanded on this work by demonstrating that while PA scales of the IPAQ correlate with accelerometry, these scales are not suitable for comparing PA behaviours between individuals with schizophrenia to controls without schizophrenia. However, to our knowledge, and despite its apparent widespread use, similar assessments of the IPAQ sitting scale among individuals with schizophrenia has yet to be reported.

Contrary to Stubbs and colleagues' (2016) criticism of self-report questionnaires underestimating SB, existing agreement data in the general adult population between the IPAQ-SF sitting scale and accelerometer-defined SB (≤99 counts/min) indicate that the IPAQ-SF may overestimate the amount of sedentary time by +130 min/d with large 95% limits of agreement (LoA) -275 to 536 min/d (Hagstromer et al., 2010). Additionally, bias appeared to be proportional, with greater overestimation associated with greater reported sitting time. A systematic review of validation and reliability studies with the IPAQ-SF found Spearman's correlations (ρ) of 0.07 to 0.61 with accelerometer-defined SB and test-retest reliability ranging from Spearman's ρ of 0.18 to 0.95, and intraclass correlation coefficients (ICC) ranging from 0.80 to 0.97 (Healy et al., 2011).

However, cognitive deficits among people with schizophrenia are common. These include memory impairment (Keefe & Fenton, 2007; Reichenberg et al., 2006) which may impact participant's ability to accurately recall activities over an extended period, and deficits in attention and executive function (Keefe & Fenton, 2007; Reichenberg et

al., 2006) which may impact the reliability and validity of any self-report questionnaire. Given these concerns, performing additional assessments of the IPAQ-SF sitting scale is warranted, and will shed light on the validity of studies that seek to examine relationships between SB and other outcomes of interest among samples of individuals with schizophrenia. Furthermore, evaluating the sitting scale of the IPAQ-SF will help inform ongoing efforts to develop new self-report measures for use among individuals with serious mental illness such as those described by Rosenbaum and Ward (2016).

The IPAQ-SF is designed for population level assessment (Bauman, Ainsworth, et al., 2009; Bauman, Bull, et al., 2009; Craig et al., 2003) as "small sample studies may be under-powered to detect between-group differences" (Bauman, Ainsworth, et al., 2009, p. S6) due to the high variance in scores obtained. However, small samples comparing the IPAQ to accelerometry will offer some insight on how an instrument may perform at a population level. Furthermore, "establishing the reliability and validity of extant measures" as suggested by phase 2 of Sallis and colleague's behavioural epidemiological framework (Sallis et al., 2000, p. 295) may demonstrate new utility for existing tools even if not their initial intention – but this must be evaluated.

As such, this study is a follow-up to the study characterizing accuracy of the PA scales of the IPAQ-SF presented by Duncan and colleagues (Duncan, Arbour-Nicitopoulos, et al., 2017). In the current study, data from a pre-existing prospective study (Arbour-Nicitopoulos et al., 2017) were reanalyzed with the purpose of assessing: 1) the agreement between accelerometry and the IPAQ-SF sitting scale, 2) test-retest reliability of the IPAQ-SF sitting scale over a 4-week time period. Furthermore, with respect to the first objective, the IPAQ-SF sitting scale may be better at assessing

prolonged periods of sitting rather than shorter incidental bouts that participants may not remember. Thus, accelerometry derived SB is evaluated using bout (distinct periods of uninterrupted sitting) lengths of 1) ≥1-min, i.e. any 60 second epoch classified SB, regardless of total bout length; 2) ≥10-min, which correspond to common guidelines for what qualifies as a bout of physical activity (e.g. Canadian Society of Exercise Physiology, 2012); and 3) ≥30-min, which represent longer periods of sitting that may be easier to recall than incidental amounts of sitting accrued in smaller bouts, and correspond to the length of a typical TV-show, an archetypal form of SB, with each of these accelerometer-derived bout lengths compared to the IPAQ-SF.

Based on our previous work with PA accuracy in this population (Duncan, Arbour-Nicitopoulos, et al., 2017) and the criticism by Stubbs et al (2016), it was hypothesized that the IPAQ-SF sitting scale would underestimate accelerometer derived SB on average, but as per Hagstromer et al. (2010), the bias would be proportional to the amount of SB engaged in by participants. Further to this point, agreement would be improved when accelerometry derived SB is defined by longer bouts over shorter bouts. Given the longer reliability period of the study, test-retest reliability would be similar to what was previously reported for MVPA in the same sample (i.e., Spearman's ρ of 0.47) (Duncan, Arbour-Nicitopoulos, et al., 2017).

3.2 Methods

3.2.1 Participants

Research ethics boards at the Centre for Addiction and Mental Health in Toronto and the University of Toronto approved the larger 4-week prospective study (Arbour-

Nicitopoulos et al., 2017), the purpose of which was to evaluate psychological determinants of physical activity participation in the sample – no effort was made to change behaviour. Participants were required to: 1) be age 18-64 years (in line with the Canadian Physical Activity Guidelines recommendations for adults (Canadian Society of Exercise Physiology, 2012), and 2) have a diagnosis of schizophrenia or schizoaffective disorder. Diagnosis was confirmed with the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998). Participants were excluded for: 1) being hospitalized over the past 12 months for angina pectoris, myocardial infarction, or cardiac surgery of any kind; and/or 2) uncontrolled hypertension (defined as blood pressure > 140/90). If eligible, participants provided written consent prior to commencing the study and capacity to consent was assessed with the MacArthur Competence Assessment Tool for Clinical Research (MacCAT-CR; Appelbaum and Grisso, 2001). Participants were recruited by referrals from nurses, psychiatrists, and other studies involving persons with schizophrenia at the Centre for Addiction and Mental Health. Given that this is a secondary data analysis, achieved power (1-β) was calculated post-hoc using G*Power 3 (Heinrich Heine Universität Düsseldorf, Düsseldorf, Germany) with α set at 0.05.

3.2.2 Procedures

SB levels were assessed at baseline with the IPAQ-SF and again four weeks later. Participants were instructed to wear an ActiGraph (Pensacola, USA) wGT3X+ accelerometer for 7 days over their right hip three weeks after baseline assessment of the primary study. Proper wear was demonstrated to the participant when receiving the accelerometer. Participants were also instructed to keep a log tracking their accelerometer

wear time. Participants returned the accelerometer when completing the final IPAQ-SF. Thus, accelerometry data coincided with the period of time assessed by the Week 4 IPAQ-SF, allowing for direct comparison of the measurements of SB obtained by each instrument.

3.2.3 Analysis

Analysis methods are similar to those described previously for evaluating the PA in this sample (Duncan, Arbour-Nicitopoulos, et al., 2017): Accelerometry data were analyzed using ActiGraph's (Pensacola, USA) *Actilife* software (v6.12). Actilife's sleep analysis suite uses algorithms designed for wrist worn accelerometers (Cole et al., 1992; Sadeh et al., 1994) and attempts to validate sleep time measurement algorithms using waist worn actigraphy in adults has been limited and demonstrated poor agreement with gold standard measurement (polysomnography) (Slater et al., 2015). Therefore, to compensate for participants who did not remove the accelerometer before sleeping nor kept a record of their sleep and wake times, data from 0h00 to 5h59 from each day were not analyzed. Wear time was calculated using Choi and colleagues' (2011) algorithm. The accelerometer needed to register 600 minutes of wear time for a day to be considered a valid (Troiano et al., 2008). As the IPAQ-SF asks participants for time spent "sitting on a weekday," accelerometer data were averaged across all available valid weekdays. Three valid weekdays out of the 7-day period were required for participants' accelerometry data to be included in the analysis (Troiano et al., 2008; Trost et al., 2005).

Troiano and colleagues' (2008) adult physical activity cut-off points were used to determine time spent engaging in SB (≤99 counts/min) using the vector magnitude of all

three axes. While the wGT3X+ does have a built in inclinometer, it has not been shown to be better at detecting SB time than using count per minute thresholds in post-secondary students (Peterson et al., 2015). SB bouts were defined in 3 ways: ≥1 minute, ≥10 minutes, and ≥30 minutes. For all definitions, a tolerance of 2 minutes spent above the maximum counts per minute was permitted before a SB bout ended.

Agreement between measurements was assessed primarily through Bland-Altman mean-difference plots (Bland & Altman, 1986, 2003) with 95% Limits of Agreement (LoA). Positive differences represented overestimation by the Week 4 IPAQ-SF assessment versus the Week 1 IPAQ-SF or accelerometry. Regression analysis was performed on the BA plots to determine if any bias between measurements may be fixed or proportional. Paired t-tests were used to determine if the mean bias between measurements represented a statistically significant difference. Based on ICC guidelines presented by Koo and Li (2016), retest reliability coefficients were calculated using an absolute agreement single measure, two-way mixed-effect model ICC ($ICC_{A,1}$). $ICC_{A,1}$ is reported for agreement between IPAQ and accelerometry as well. Furthermore, as Spearman's rank correlation coefficients is a commonly reported indicator of validity and reliability in the literature, it is also reported here for comparability to other studies (missing values excluded pairwise and confidence intervals calculated by applying Fisher's z-transformation) but is not a primary outcome of this study.

All statistical analyses were performed in RStudio (v1.1.453; R v3.4.4) with package 'psych' (v.1.8.4) installed (Revelle, 2017). Bonferroni corrections ($\alpha = 0.0167$) were applied to account for the three definitions used to assess SB with the accelerometer.

3.3 Results

3.3.1 Demographics

Of 130 participants who enrolled in the parent study, 113 completed the study, and 103 participants had sufficient accelerometry data. At Week 1, 105 of 130 participants had valid IPAQ-SF sitting responses (did not select "Don't know/Not sure"). At Week 4, 84 of 113 participants had valid IPAQ-SF sitting responses, resulting in 74 participants with valid responses at both week 1 and 4. Meanwhile, 74 participants had both valid accelerometry data and IPAQ responses at week 4. Table 3.1 describes characteristics of the participant who had valid data, separated by whether data were available for the validity and reliability assessments as well as replicating overall sample data presented by Duncan and colleagues (Duncan, Arbour-Nicitopoulos, et al., 2017). Figure 3.1 illustrates the interrelationship of valid data collected.

3.3.2 IPAQ Validity

Table 3.2 summarizes both the IPAQ-SF sitting scale (at Week 1 and 4) and average daily SB as measured by accelerometry for all valid responses, as well as Spearman's correlation between methods. $ICC_{A,1}$ between methods was lowest and non-significant for ≥ 30-min bouts: .09 (95% CI: -.05 to .24), $p = .04$. $ICC_{A,1}$ between methods for ≥1-min and ≥ 10-min bouts were comparable at .21 (95% CI: .04 to .36), $p = .004$ and .23 (95% CI: .06 to .39), $p = .004$ respectively; and were statistically significant at the Bonferroni corrected alpha.

Figure 3.2a-c includes BA plots of the difference between each accelerometry SB definition and Week 4 IPAQ-SF against the mean of both methods. Regression lines are

displayed in Figure 3.2a-c as alternating short and long dashed lines. Linear regression was significant and explained 32% of the variance when accelerometry SB was measured in ≥1-min bouts, 29% for ≥10-min bouts, and 45% for ≥30-min bouts. Mean difference between IPAQ-SF and accelerometry ≥1-min SB bouts was -86.1-min, 95% LoA: -569.9-min to 397.6-min, Cohen's $d_z = 0.35$, $1\text{-}\beta = 0.91$; and for ≥ 30-min SB bouts was 220.7 minutes, LoA: -266.6 to 708.0, Cohen's $d_z = 0.89$, $1\text{-}\beta = 1.0$. T-tests revealed significant differences for both comparisons. Mean difference for ≥10-min bouts was 26.8 minutes, LoA: -458.7 to 512.3, Cohen's $d_z = 0.11$, $1\text{-}\beta = 0.23$, but t-test results were not statistically significant. Results from statistical analyses for regressions and paired t-tests are summarized in Table 3.3.

3.3.3 IPAQ Reliability

Figure 3.2d includes the BA plot of the difference between Week 1 and Week 4 administrations of the IPAQ-SF against the mean of both time points. Week 1 and Week 4 administrations differed by an average of -26.6 minutes, 95% LoA: -564.2 to 510.9, Cohen's $d_z = 0.10$, $1\text{-}\beta = 0.13$. but this was not statistically significant. Linear regression analysis of the plot was also not significant (results of statistical analyses also reported in Table 3.3), indicating that bias between measures is fixed. $ICC_{A,1}$ between time-points was 0.41 (95% CI: 0.21 to 0.59, $p<0.001$), representing a 'fair' association between measurement times (Cicchetti, 1994).

3.4 Discussion

Previous evidence of validity suggests that the IPAQ-SF overestimates SB by 130 minutes per day with LoA of ±6.75h per day when using an accelerometry definition of 100 counts/min and no minimum bout length (Hagstromer et al., 2010). The comparable analysis performed in this study (bouts ≥1-min) suggests that the IPAQ-SF may underestimate by 86 minutes compared to accelerometry defined SB, but with even larger LoA (±8h). This is in line with Stubbs and colleague's (2016) hypothesis that self-report tends to underestimate SB among people with severe mental illness, despite the opposite being evident in the general population. One plausible explanation may be related to symptoms such as apathy and amotivation (e.g. less effortful recall; activities of daily life perceived as more effortful and thus not SB). However, more exploration as to why this is the case is warranted in order identify accurate self-report methods for collecting information on SB. However, using a minimum 10-min bout length seems to improve the mean bias so that it is no longer significantly different, but LoA remain large. As well, much like the data presented by Hagstromer and colleagues (2010) from the general adult population, all BA analyses suggested proportional bias occurs when using the IPAQ-SF.

With regards to correlational evidence, our results suggest that relationships between accelerometry and the IPAQ-SF in individuals with schizophrenia is on the low end of what is reported in the literature from the general adult population. ICCs between the accelerometer and the IPAQ-SF ranged from 0.09 to 0.23, and Spearman correlations ranged from 0.25 to 0.30, compared to Spearman correlations of 0.07 to 0.61 reported by Healy et al. (2011). The ICC as a measure of absolute agreement provides a better indication of whether the minute values between the two methods align, which given the

agreement data presented in BA plots, is unsurprisingly low. Even Spearman's correlation – which is a better indicator of the general monotonic relationship (i.e. higher sitting scores on the scale are associated with higher levels of measured SB) – was low.

Reliability indices were also quite low compared to the general adult population. The ICC reported here of 0.41 is well outside the range of 0.80 to 0.97 reported in the literature for the general adult population (Healy et al., 2011). Spearman's correlation did however fall within the range reported by Healy and colleagues (2011) ($\rho = 0.41$ vs $\rho =$ 0.18 to 0.95) despite representing a 4-week span, though again, the ICC used here is a more appropriate indicator of whether the same values are reported at both time points.

Contrary to the hypothesis, the IPAQ-SF sitting scale best represented accelerometer derived SB when minimum bout lengths were set to ≥10-min, followed by ≥1-min, and finally ≥30-min bouts. While all three comparisons demonstrated statistically significant proportional bias such that higher amounts of SB resulted in more overestimation by the IPAQ-SF, this bias was most pronounced in the ≥30-min bout comparison where the slope of the regression line was 1.14 and explained 45% of the variance, compared to slopes of 0.85 and 0.86, and 29% to 32% of the variance explained when using the smaller minimum bout lengths.

With regards to mean bias, IPAQ-SF scores overestimated by 28.7 minutes per day but did not differ significantly from SB bouts of ≥10-min, whereas the IPAQ-SF scores did significantly underestimate compared to SB bouts of ≥1-min by 84.0 minutes on average, and scores significantly overestimated when compared to SB bouts of ≥ 30-min by 224.7 minutes on average. Based on these findings, when using the IPAQ-SF sitting scale as an indicator of SB among individuals with schizophrenia, researchers

should interpret this as an indicator of the amount of SB engaged in bouts of at least 10 minutes – much like how the IPAQ-SF asks for bouts of PA of at least 10 minutes. Interpreting IPAQ-SF sitting data as an indicator of all SB time (i.e. ≥1-min bouts) may also be acceptable, but less accurate. However, the sitting scale should not be interpreted as an indicator of just prolonged sitting time (i.e. ≥30-min bouts).

It is unclear why the IPAQ performs best as an estimate of 10-minute bouts of sitting time, though several explanations are plausible. Foremost, the IPAQ PA questions ask for minimum 10-minute bouts. While it does not do this for the sitting item, participants may be cued to assess their day for sitting in a similar fashion as when trying to recall PA. Additionally, it is possible that participants are able to factor in even short periods of SB to their estimate/recall such as a bus ride or waiting for an appointment. Research into the response processes used by participants to estimate PA and SB may help understand how activity profiling questions are responded to.

However, while the ≥10-min bout of SB interpretation does appear to be the best way to interpret IPAQ-SF sitting scale score, there are still substantial issues with the use of this self-report questionnaire as an indicator of SB. Firstly, the 95% LoA are quite large at approximately ±8h/day among all three bout definitions. Thus, while the bias did not differ significantly in the ≥10-min bout analysis on average, individual scores vary widely from accelerometry estimates. Secondly, the presence of proportional bias is a concern for validity as well, with overestimation tending to occur with higher levels of SB while underestimation occurs with lower levels. Finally, from a theoretical standpoint, while the prelude to the sitting scale includes the instructions: “This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch

television," the question itself only asks about sitting. As such, SB in a lying or reclining position may not be considered by participants when responding to this question.

Repeated administration of the IPAQ-SF resulted in no significant mean bias between time-points, nor was there significant proportional bias. However, LoA were approximately ±9h/day, and the ICC was much lower than those reported in (Healy et al., 2011). Thus, while there appears to be no bias between the two measurement periods, there was substantial inconsistency. However, this inconsistency may represent actual differences in individual sitting behaviours between the two measurement periods. Future studies aiming to verify the reliability of physical activity and SB measurement tools should consider methods to address this possibility, such as using objective measures to assess the behaviours over both periods to determine if the behaviour was actually similar (Duncan, Arbour-Nicitopoulos, et al., 2017), or by shortening the retest duration so that both administration cover similar time periods (e.g. two administrations on the same day, or one day apart so that the majority of days in the 7-day measurement period are covered).

The results presented here suggest that the IPAQ-SF sitting scale may be an unsuitable measure of SB for many scenarios when working with individuals with schizophrenia including at a population level or with smaller samples. While mean values do compare to objective accelerometry, other psychometric properties are lacking. LoA are large and bias is proportional, instilling a lack of confidence in whether the actual "minutes" of activity reported by the IPAQ-SF reflect the minutes of SB engaged in by any one *individual*. As a result, any correlational data based on the IPAQ-SF or cut points based on the number of minutes reported should be interpreted with caution. Given low

Spearman correlations, even interpretations about who has high vs low rates of SB within a sample cannot be made with confidence. At best, the IPAQ-SF can be used to indicate the mean SB of a *sample.* As Firth and colleagues (Firth et al., 2018) suggest regarding the PA scales, similar caution is likely required if comparing sitting time measured by the IPAQ-SF among individuals with schizophrenia and control. Given that BA analysis suggests similar characteristics (i.e., mean and proportional biases) when used in the general adult population (Hagstromer et al., 2010) the IPAQ-SF may be adequate at best when comparing to the general adult population. However, any further data generated should be considered preliminary or exploratory and would require replication with more objective methods in order to confirm whether any relationship exists between SB and a proposed outcome or predictor.

To our knowledge this is the first study to generate evidence of validity when using the IPAQ-SF sitting scale in a sample of individuals with schizophrenia. However, as a secondary analysis, some limitations were inherent. The use of accelerometers without the additional use of a purpose-built inclinometer only allows the use of counts per minute to classify SB but does not take into account posture. Thus, behaviours such as standing with low activity counts per minute may be improperly classified as SB.

Unfortunately, much data were unusable in this sample. While accelerometer adherence was high (only 8.9% of participants had insufficient wear time), IPAQ responses that could be compared against the accelerometry were minimal. At Week 4, 25.7% of the completing sample (and 29.5% of those with valid accelerometry) selected "Don't know/Not sure" option on the sitting item, which could not be compared to accelerometry as a result. Thus, the usable sample was reduced to 74. Given the large

effect size observed for SB bouts ≥1 and ≥30 minutes, these analyses were reasonably powered ($1\text{-}\beta \geq 0.91$ when $\alpha = 0.05$) to avoid both Type I and Type II errors. However, with the small effect observed for SB bouts ≥10-min a false negative may have resulted (i.e. no statistical difference between measures concluded despite a difference existing). That being said, whether the IPAQ differs *statistically* from accelerometry, may be less relevant than simply characterizing the relationship between the resultant measurements to indicate the implications of using the IPAQ as an assessment of SB. Furthermore, while this is a small sample compared to the initial intention of the IPAQ (Bauman, Ainsworth, et al., 2009; Bauman, Bull, et al., 2009; Craig et al., 2003), results are of a similar magnitude as reported by Hagstromer and colleague's (2010) from larger sample of the general population.

Additionally, while this study builds on previous efforts to develop validity evidence of the IPAQ-SF when used with samples of individuals with schizophrenia (Duncan, Arbour-Nicitopoulos, et al., 2017; Faulkner et al., 2006), the focus of these efforts has been predominantly on the evidence of convergence (American Educational Research Association et al., 2014) with other measurement methods, that is the accuracy of the tool relative to other methods. However, there are several other sources of evidence for validity that can also be considered, namely evidence based on: test content, response processes, internal structure, and the consequences of testing (AERA, APA, NCME, & JCSEPT, 2014). Having analyzed the accelerometry data in varying minimum bout lengths provides some insight on response processes (i.e., how participants approach responding to items (Padilla and Benítez, 2014; AERA, APA, NCME, & JCSEPT, 2014)). Even short bouts of sitting appear to be considered when individuals with

schizophrenia respond to the IPAQ-SF. However, other methods of assessing response processes such as eye-tracking (to identify what participants are attending to when presented with a questionnaire) and cognitive interviews (where participants are asked to describe aloud how they are interpreting the items, and their process behind generating a response) (Padilla & Benítez, 2014) may help inform the interpretation of IPAQ-SF data and help researchers design more accurate self-report questionnaires for intervention and surveillance work. In particular, identifying why participants opt to select "Don't know/Not sure" may help reduce lost data.

3.5 Tables

Table 3.1 *Summary of Participant Demographics Separated by Assessment*

Demographic	Enrolled (*n* = 130)	Completed (*n* = 113)	Validity Analysis (*n* = 74)	Reliability Analysis (*n* = 74)
Male:Female	80:50	68:45	47:27	48:26
Current smokers	63	55	33	33
Outpatient:Inpatient	124:6	108:5	70:4	70:4
Mean (*sd*) Age	40.1 (11.6) yrs	41.0 (11.7) yrs	42.3 (11.9) yrs	41.9 (11.8) yrs
Ethnicity				
African Descent	21	19	12	12
Asian/South Asian	20	16	11	11
Caucasian & Hispanic	74	68	44	43
Multi-ethnic	5	5	3	3
Other	10	5	4	4
Symptom Severity				
BPRS-A mean score (*sd*)	34.3 (7.6)	34.2 (8.5)	33.6 (7.1)	34.1 (9.0)
CGI-S mean score (*sd*)	3.5 (1.1)	3.5 (1.1)	3.4 (1.0)	3.4 (1.1)
AES mean score (*sd*)	32.0 (7.9)	31.7 (7.9)	30.6 (7.4)	31.8 (7.8)
CPZ equivalents mean (*sd*)	788.5 (1237.3) mg	748.1 (1216.6) mg	637.7 (819.9) mg	603.1 (820.4) mg
BMI*				
Mean (*sd*)	31.7 (8.1) kg/m^2	31.5 (8.4) kg/m^2	30.6 (7.2) kg/m^2	30.0 (6.1) kg/m^2
Underweight (BMI<18.5)	1	1	1	1
Normal Weight (18.5<BMI<25)	21	19	13	13
Overweight (25<BMI<30)	35	32	24	24
Obese (BMI>30)	72	60	35	35
Education				

Some High School (no diploma)	28	22	13	15
High School Diploma	36	31	14	16
At least some Postsecondary	62	59	45	41
Other (e.g. apprenticeship)	4	2	2	2
Employment				
Full-Time	3	2	2	2
Part-Time	36	33	23	24
Student	7	6	3	4
Unemployed	77	65	38	35
Other (e.g. retired, volunteer)	7	7	8	8

Note: Values are frequency unless otherwise specified. Higher score represent more severe symptoms. *One participant opted out of being weighed. BPRS-A = Brief Psychiatric Rating Scale 18-item Anchored version (Woerner et al., 1988), CGI-S = Clinical Global Impression Severity Scale (Guy, 1976b), AES = Apathy Evaluation Scale (Marin et al., 1991), CPZ = Chlorpromazine Equivalents (Gardner et al., 2010), BMI = Body Mass Index.

Table 3.2 *Sedentary Behaviour over Two 1-Week Periods*

Assessment	Descriptive Statistics						Correlation with IPAQ Week 4	
	Mean	Median	*sd*	*n*	Min	Max	$ICC_{A,1}$	Spearman's rho
Validity								
IPAQ Week 4	424.6	390.0	247.0	74	0	960.0	-	-
Accelerometer:								
≥1-min Bouts	515.7	524.8	135.8	74	142.8	827.2	0.21 (CI: 0.04 to 0.36), p = 0.004[B]	ρ = 0.25, p = 0.029
≥10-min Bouts	403.1	408.6	138.1	74	99.0	748.8	0.23 (CI: 0.06 to 0.39), p = 0.004[B]	ρ = 0.30, p = 0.009[B]
≥30-min Bouts	213.6	190.4	111.9	74	25.0	490.8	0.09 (CI: -0.05 to 0.24), p = 0.04	ρ = 0.26, p = 0.023
Reliability								
IPAQ Week 1	447.4	360.0	260.1	74	60.0	1210.2	0.41 (CI: 0.21 to 0.59), p < 0.001*	ρ = 0.49, p < 0.001*
IPAQ Week 4	420.8	390.0	245.9	74	0.0	960.0	-	-

Note: IPAQ Week 4 descriptive statistics are reported separately for the individuals with valid accelerometry and with valid week 1 IPAQ responses. For correlations * = statistically significant at α = 0.05 where appropriate; [B] = statistically significant at Bonferroni corrected α = 0.0167 where appropriate

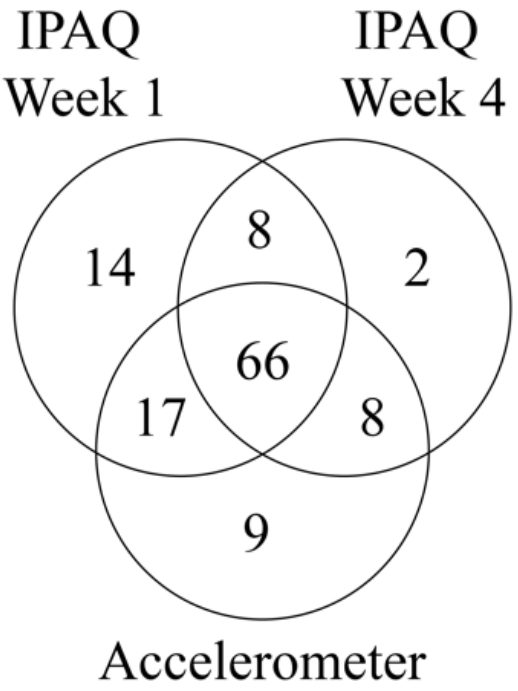
Table 3.3 *Bland-Altman Plot Statistical Analyses*

	Regression Analysis						Paired T-test			
Assessment	R^2	*df*	*F*	*p*	$\beta 1$	$\beta 0$	Δ *(sd)*	*t*	*df*	*p*
Validity										
≥1-min Bouts	0.32	1, 72	31.9	<.001[B]	0.86	-505.5	-86.1 (246.8)	3.00	73	0.004[B]
≥10-min Bouts	0.29	1, 72	29.8	<.001[B]	0.85	-334.2	26.8 (247.7)	.93	73	0.36
≥30-min Bouts	0.45	1, 72	59.9	<.001[B]	1.14	-155.2	220.7 (248.6)	7.63	73	<0.001[B]
Reliability										
IPAQ Week 1	0.01	1, 72	0.28	0.6	-0.08	7.9	-26.6 (274.3)	0.84	73	0.41

[B] = Significant with Bonferroni correction

3.6 Figures

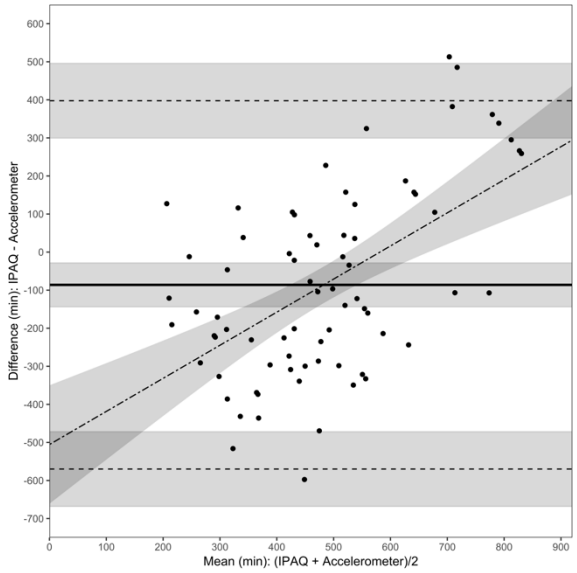
Figure 3.1 *Venn Diagram Representing Overlap of Sedentary Behaviour Data Sources Across Participants*



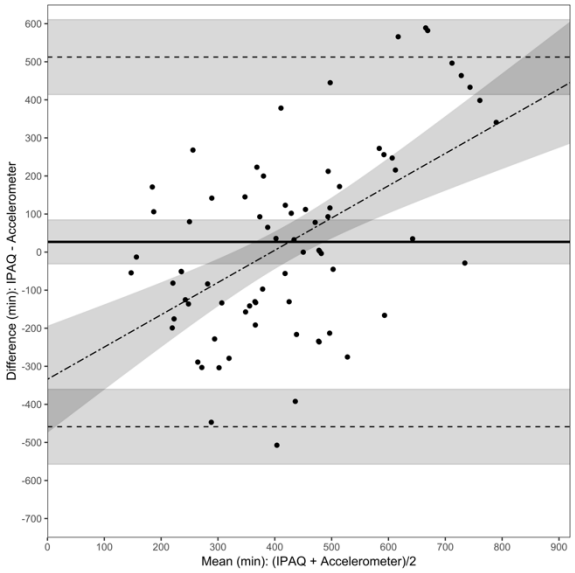
Note: Participants with no physical activity data = 6.

Figure 3.2a–d *Bland-Altman Plots of SB Measured by IPAQ-SF and Accelerometry*

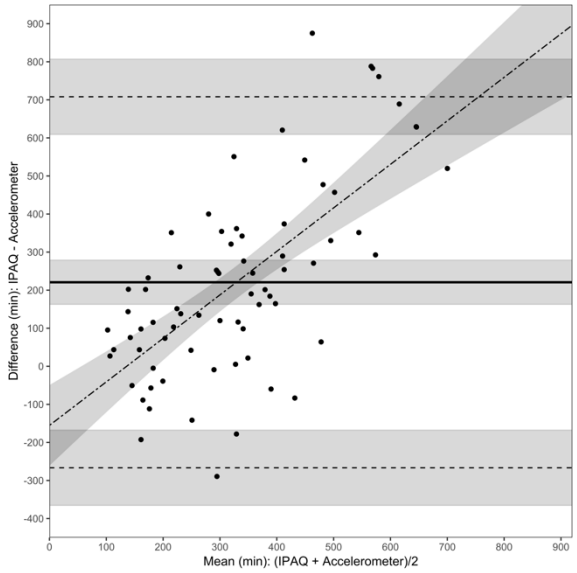
a) *IPAQ Week 4 vs ≥1-min bout Accelerometry*



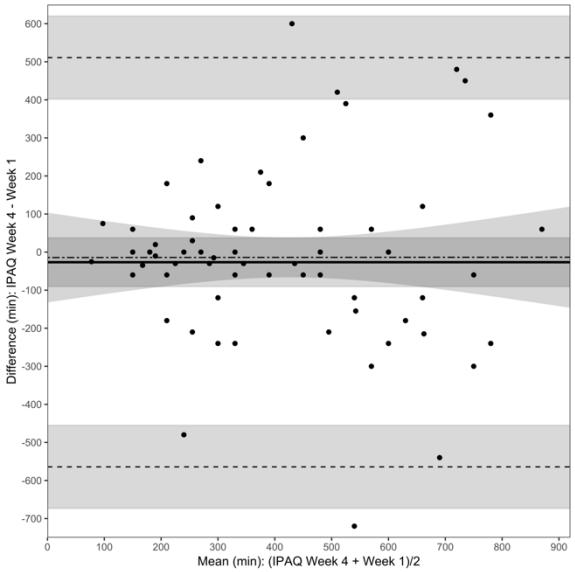
b) *IPAQ Week 4 vs ≥10-min bout Accelerometry*



c) *IPAQ Week 4 vs ≥30-min bout Accelerometry*



d) *IPAQ Week 4 vs Week 1*



Note: Solid lines indicate mean difference (bias), long dashed lines indicate 95% agreement limits, and alternating short and long dashes indicate the linear regression line. Grey areas represent 95% confidence intervals around respective lines.

Chapter 4: Correlates of Discrepancy Between Measurement Tools

4.1 Introduction

Increasing physical activity (PA) and reducing sedentary behaviour (SB; any waking activity characterized by an energy expenditure less than 1.5 metabolic equivalents in a sitting, reclining or lying posture (Tremblay, Aubert, et al., 2017)) are critical for reducing the high rates of physical comorbidity (such as obesity, diabetes, metabolic syndrome, and cardiovascular disease) in people with schizophrenia (Bueno-Antequera et al., 2017; McNamee et al., 2013; Vancampfort, De Hert, et al., 2014; Vancampfort, Probst, Knapen, et al., 2012; Vancampfort, Rosenbaum, et al., 2016). In turn, redressing these comorbidities may reduce the 15-25 year shortened lifespan among people with schizophrenia (Laursen et al., 2012, 2014), increase quality of life (Faulkner, Cohn, Remington, et al., 2007; Foldemo et al., 2014; Gorczynski & Faulkner, 2010; Guo et al., 2013; Sugawara et al., 2013), and improve antipsychotic medication adherence by reducing the impact of antipsychotic-associated weight gain (Faulkner, Cohn, Remington, et al., 2007).

In order to properly monitor and detect change in PA and SB, tools that accurately measure these behaviours are essential (Sallis et al., 2000) – especially tools that can be broadly administered at a population level or quickly administered in a health care setting. Sallis and colleague's (2000) behavioural epidemiology framework suggest that using existing behavioural measurement tools for novel purposes such as assessing interventions or identifying behavioural determinants may be appropriate but must be assessed for these purposes. The International Physical Activity Questionnaire (IPAQ) is among the most commonly used self-report questionnaires for assessing moderate-to-vigorous PA (MVPA) (Stubbs, Firth, et al., 2016) and SB (Stubbs, Williams, et al., 2016) in schizophrenia. While the IPAQ was originally designed for

population-level assessment (Bauman, Ainsworth, et al., 2009; Bauman, Bull, et al., 2009; Craig et al., 2003) to avoid being under-powered for between-groups analyses (Bauman, Ainsworth, et al., 2009), it is frequently used in for other purposes such as assessing PA as a secondary outcome or covariate in interventions (e.g. Aschbrenner et al., 2018; McIver, O'Halloran, & McGartland, 2009).

Recent work has called into question the validity of using the IPAQ as a tool to measure minutes of PA among individuals with serious mental illness such as schizophrenia. The initial paper to provide validity evidence for using the IPAQ in individuals with schizophrenia demonstrated a Spearman correlation of 0.37 between self-reported minutes of MVPA and accelerometer-derived metabolic equivalents/week; this correlation is similar to what is observed in the general population (Faulkner et al., 2006). Expanding on this, recent research revisiting the IPAQ has demonstrated that agreement between self-reported minutes and accelerometry-derived minutes of MVPA and SB show minimal differences between measures when averaged across the sample ($n \geq 74$), but large limits of agreement suggest that on an intra-individual basis, IPAQ reported minutes have concerning patterns of agreement with accelerometry (Duncan et al., 2019; Duncan, Arbour-Nicitopoulos, et al., 2017). Individual responses to the IPAQ vary widely in the degree of error from accelerometry-derived values, however, when averaged the self-reported values approach an accurate representation of mean objectively measured PA for the sample.

This issue limits the utility of the IPAQ beyond an estimate of a sample's mean MVPA and SB (Duncan et al., 2019), which can be useful for demographic sample descriptors or cross-sectional monitoring on a population level. However, due to the possible measurement variability of any one individual's response on the IPAQ, associations between the IPAQ and health

variables may not accurately represent the relationship. Furthermore, Firth and colleagues (Firth et al., 2018) have shown that when IPAQ derived PA data were compared between individuals with schizophrenia and healthy controls, the IPAQ may diminish differences that are detected when using accelerometry data, suggesting that the IPAQ may also not be appropriate for comparisons of mean values between populations.

Collectively, this evidence paints a gloomy picture for data accrued from one of the most popular and widespread tools for measuring PA and SB in schizophrenia, especially with respect to accuracy of individual measurements. That being said, self-report tools remain a necessity in PA and SB epidemiology because they are low-cost, easy to disseminate, and a minimal burden for participants and clinicians when compared to accelerometry. Currently, efforts are underway to develop and test a PA questionnaire specifically for measuring PA among clinical mental health populations (Rosenbaum & Ward, 2016) which may help improve accuracy by taking the unique needs of people with severe mental illness into account. However, for any such undertakings to develop better tools, it may be useful to understand why PA is being mis-measured in the first place by existing tools.

One issue to consider is whether individual differences contribute to measurement error. In populations without severe mental illness, age, sex, and BMI status have been used to account for differences between self-report and device-assessed measures of PA and SB (Saint-Maurice et al., 2014; Welk et al., 2017). Welk and colleagues (2017) found that higher BMI was a predictor of underestimating sitting time and overestimating MVPA, and hypothesized that for MVPA higher BMI status may contribute to increased perceived exertion of activity, and thus lower intensity PA may be inaccurately classified as MVPA when self-reported. Similarly, Welk and colleagues (2017) found that older age was associated with underestimating SB and

overestimating MVPA in adults, and Saint-Maurice et al. (2014) found that among children, younger age predicted overestimating MVPA in children. It is plausible that age may also influence perceived exertion, especially in older adults. Finally, female adults tended to underestimate both SB and MVPA (Welk et al., 2017), whereas female children tended to overestimate MVPA (Saint-Maurice et al., 2014). It is unclear what might be influencing these sex differences in reporting tendencies.

Individuals with schizophrenia do tend to have higher BMI status (Bueno-Antequera et al., 2017; McNamee et al., 2013; Vancampfort, De Hert, et al., 2014; Vancampfort, Probst, Knapen, et al., 2012; Vancampfort, Rosenbaum, et al., 2016), which may partially explain discrepancy between measurements in this population. Additionally, while incidence of schizophrenia is similar between males and females, the disease seems to affect males differently from females. Males tend to have earlier onset, more negative symptoms, fewer affective symptoms, and poorer social functioning compared to females (Li et al., 2016) which may compound sex differences in reporting compared to the general population.

In addition to possible variables observed in the general population, individuals with schizophrenia may have additional symptomatic factors that may contribute to error when self-reporting PA and SB. Cognitive deficits are a hallmark of individuals with schizophrenia (Reichenberg et al., 2006) (Harvey et al., 2006), including memory and executive function impairment (Keefe and Fenton, 2007; Reichenberg et al., 2006), which may influence recall or understanding of self-report questionnaires. Inherently, psychotic features are indicative of poor reality testing (Bentall et al., 1991), and may also interfere with accurate recall. Additionally, negative symptoms – especially apathy or amotivation which are strongly related to functioning in individuals with schizophrenia (Foussias & Remington, 2010; Kiang et al., 2003) – may

reduce the effort participants are willing to expend in order to complete the questionnaire accurately or could lead to greater perceived effort when engaging in PA, similar to individuals with high BMI. However, to our knowledge no studies have attempted to explore symptomatic correlates of error in PA and SB self-reporting.

Efforts to create new measurement tools may be able to use this information to create items that address these limitations. Additionally, identifying these measurable differences that contribute to error creates the possibility of performing regression calibration. Regression calibration is a method commonly used in nutrition (Agogo et al., 2014; Bennett et al., 2017) and occasionally PA (Metcalf et al., 2018; Saint-Maurice et al., 2014; Welk et al., 2017) measurement research where regression models are used to adjust the values obtained from an error prone measurement method (*Xw*) (e.g. self-report) to better align with another measurement (*Xz*) (e.g., a gold standard method or known value) (Guolo, 2008). *Xw* can be used as the sole predictor of *Xz*, however other variables can be added to the model as well to further explain the error between measurement methods. The resultant equation from the regression model can then be used to estimate *Xz* values in other samples where *Xw* was measured.

Thus, the purpose of this paper is to build on the validity evidence of how the IPAQ is likely to perform in samples with schizophrenia previously presented by Duncan and colleagues (2019; 2017) by identifying whether individual factors are related to the discrepancy between the IPAQ and accelerometry; specifically, age, gender, BMI, education, various measures of mental health status such as negative symptoms, and cognitive impairment were examined. Subsequently, an attempt to use these factors to correct for disagreement was performed. This consisted of three analyses: (1) First, an exploratory correlation approach was conducted to identify possible links between individual factors and MVPA and SB measurement differences.

(2) Subsequently, results of the initial analysis were used to inform a stepwise regression approach to generate parsimonious models to calibrate IPAQ self-reported data to accelerometry derived SB and MVPA. (3) Finally, data were resampled using a 5-fold 50 repeat cross-validation approach to estimate how selected regression calibration from the second analysis may generalize (Hastie et al., 2009). It was expected that the model predicted values of SB and MVPA generated by the regression calibration approach will reduce the range of limits of agreement with accelerometry compared to uncalibrated IPAQ values. Identifying the correlates of discrepancy will help researchers identify possible sources of error when using self-report PA tools and inform tool development. Developing a regression calibration equation may allow for calibrating data collected by the IPAQ to reduce measurement error, given that the IPAQ does not require an interviewer to administer, and can thus be easily distributed.

4.2 General Methods

4.2.1 Participants

The present study is a secondary data analysis of a study assessing the psychological determinants of objectively measured MVPA in individuals with schizophrenia and schizoaffective disorder (Arbour-Nicitopoulos et al., 2017). Research ethics boards at the Centre for Addiction and Mental Health in Toronto and the University of Toronto approved the original study. To be included participants had to: (1) be between age 18–64 years and (2) have a diagnosis of schizophrenia or schizoaffective disorder, which was confirmed with the Mini-International Neuropsychiatric Interview (MINI) (Sheehan et al., 1998). Exclusion criteria consisted of: (1) being hospitalized over the past 12 months for angina pectoris, myocardial infarction, or cardiac surgery of any kind; and/or (2) uncontrolled hypertension (defined as blood

pressure>140/90 mmHg). Participants were recruited by referrals from nurses, psychiatrists, and other studies involving persons with schizophrenia at the Centre for Addiction and Mental Health. Eligible participants provided written consent prior to commencing the study and the MacArthur Competence Assessment Tool for Clinical Research (MacCAT-CR; Appelbaum and Grisso, 2001) assessed capacity to consent.

Available sample descriptors included: age; sex; BMI; International Diabetes Federation Waist Circumference classification (IDF-WC; above or below sex and ethnicity cut points for metabolic syndrome criteria (IDF, 2006)); Chlorpromazine equivalents (CPZ) based on self-reported prescription (Gardner et al., 2010), and the 12-item Short Form Health Survey version 2 (SF12) Physical Composite Score (PCS) and Mental Composite Score (MCS) as indicators of quality of life (Ware, 2002). Symptom-specific factors included measures derived from: the Brief Psychiatric Research Scale 18-item Anchored version (BPRS) (Woerner et al., 1988) and subscales based on the schizophrenia appropriate 5-factor model identified by Shafer (2005); BPRS-Activation, BPRS-Affective, BPRS-Positive, BPRS-Negative, BPRS-Resistance; Clinical Global Impression Severity Scale (CGI-S) (Guy, 1976a) Apathy Evaluation Scale (AES) (Marin et al., 1991). Finally a subsample (n = 94) of the study had two cognitive measures available which were added in an amendment after the study had commenced: Symbol Coding (SC)– measuring speed of processing and executive function – and Letter Number Span (LNS– measuring working memory – with higher scores in each representing better functioning (Fervaha et al., 2014).

4.2.2 Data Analysis

Physical activity and SB data were processed using the methods described by Duncan and colleagues (2019; 2017). Specifically, accelerometry data were analyzed using Actigraph's Actilife software (v6.12). Wear time was calculated using Choi and colleagues' (2011a) algorithm. To be considered a valid day, the accelerometer needed to register 600 min of wear time (Troiano et al., 2008). A valid week was at least four valid days of the 7-day wear period (Troiano et al., 2008; Trost et al., 2005).

As participants did not keep a record of their accelerometer wear time and had the option to not remove their accelerometer before sleeping, data from 0h00 to 5h59 from each day were not analyzed. Troiano et al.'s (2008) adult PA cut-off points were used to determine time spent in MVPA (≥2020 counts per minute) and SB (≤99 counts per minute). IPAQ and SB bouts were required to be at least 10 minutes in length with a tolerance of 2 minutes spent above the maximum counts per minute permitted before a bout ended. Bouts of at least 10 minutes are in line with IPAQ operationalization for MVPA, and IPAQ sitting to be a closer representation of SB bouts of at least 10 minutes (Duncan et al., 2019).To compensate for varying valid wear days among participants, average daily MVPA and SB were calculated for both IPAQ and accelerometry data. Differences between IPAQ and accelerometry measures were calculated such that positive values represent overestimation by the IPAQ compared to accelerometry. All statistical analyses were performed in RStudio (v1.1.453; R v3.4.4).

4.3 Analysis 1: Exploratory Correlational Analyses

4.3.1 Methods

Exploratory Pearson correlations were performed for the differences (indicating magnitude and direction of differences between measurements) and absolute differences (magnitude of differences regardless of directionality) between the IPAQ and accelerometry measured MVPA and SB with each of the participant descriptor variables. Visual inspection of the bivariate relationship between variables was also performed to identify possible non-linear relationships between variables.

4.3.2 Results

Table 4.1a summarizes the correlations between the MVPA differences in measurements and the participant individual differences. Table 4.1b summarizes the correlations between the SB differences in measurements and the participant individual differences. Descriptive statistics presented in Table 4.1a-b represent the means, standard deviations and *n* among participants who had valid data for the IPAQ, accelerometer and the participant descriptors. The significant correlates of the difference between IPAQ and accelerometry measured MVPA were CPZ ($r = -0.34$, $p < 0.001$), IDF-WC ($r = -0.34$, $p < 0.001$), SF12-MCS ($r = -0.25$, $p < 0.05$), and BMI ($r = 0.22$, $p < 0.05$). CPZ ($r = 0.34$, $p < 0.001$), IDF-WC ($r = -0.35$, $p < 0.001$), and SF12-MCS ($r = 0.24$, $p < 0.05$) also significantly correlated to absolute MVPA differences. The significant correlates of SB differences were BPRS Affective Subscale ($r = 0.25$, $p < 0.05$) and AES Total Score ($r = 0.24$, $p < 0.05$). Only age ($r = 0.23$, $p < 0.05$) was correlated with absolute difference in SB. No apparent non-linear relationships emerged from visual inspection.

4.4 Analysis 2: Regression Modelling

4.4.1 Methods

A 3-step approach was chosen to minimize the number of predictor variables in the regression calibration. Minimizing the number of predictors prevents overfitting, reduces the number of variables that must be included when performing the calibration on another sample, and reduces the number of error terms when error prone predictors are used.

For both MVPA and SB variables, regression Stage 1 consists of a single model using IPAQ data as the sole predictor of accelerometry data. Stage 2 consists of multiple models which add one of the participant descriptors to the Stage 1 model. Based on the exploratory correlation data, all significant correlates were selected to be tested in a hierarchical regression approach to account for variance between the IPAQ and accelerometry. Additionally, the AES, CGI, BPRS total score, BPRS-Resistance, BPRS-Negative, CPZ, SC, and LNS were included as predictors regardless of significance during the exploratory phase, as they capture symptoms and factors particularly relevant to the schizophrenia population. For predicting MVPA the IPAQ sitting scale was included as a potential predictor and vice-versa.

Stage 3 uses the most predictive model from Stage 2 and adds a second variable which improved explained variance in Stage 2. Figures 4.1 and 4.2 summarize the model relations and notation methods for MVPA and SB respectively. For each model a listwise deletion approach was used for missing data, thus model improvement scores reported reflect nested model comparisons with the same number of complete cases where necessary.

4.4.2 Results

Table 4.2a summarize the results of the regression modelling process for MVPA. For MVPA, Model 1 explained 8% of the variance and was statistically significant. In Stage 2 adding CPZ, IDF-WC, SF12-MCS, and BMI improved significantly on Model 1. While adding IDF-WC explained slightly more variance than other variables added in Stage 2, the corresponding complete case of Model 1 only explained 2% of the variance and was not significant. MVPA Model 2a was chosen to improve upon in MVPA Model 3 because it was a significant improvement on MVPA Model 1 even though Model 1a was significant (as opposed to IDF-WC and Model 1b) and omitted little data. Models 3ab, 3ac, and 3ad then added IDF-WC, SF12-MCS and BMI respectively. Models 3ab, 3ac and 3ad all improved significantly on corresponding complete case models in Stage 2. Models 1, 2a, 3ab, 3ac and 3ad were selected for evaluation in Study 3, as these represented models that were significant improvements over the previous model at each level of model complexity.

Table 4.2b summarizes the results of regression modelling for SB. Similar to MVPA, Model 1 explained 8% of the variance. SB Models 2b and 2e improved upon Model 1b and 2e by adding AES and BPRS-Resistance scores respectively. Model 3eb was not a statistical improvement over 2e (but was an improvement over 2e $\Delta R^2 = 0.053$, $p = 0.04$). Models 1, 2b, and 2e were selected for evaluation in Study 3. Model 3eb was also selected for further evaluation, as it explained the maximum variance, even though it was not a statistically significant improvement over 2e.

4.5 Analysis 3: Predicting Activity Data

4.5.1 Methods

Model 1, as well as the most predictive version of Models 2 and 3, were used to predict MVPA and SB data. Additionally, a variant of Model 1 was calculated with an intercept fixed at 0, in order to ensure that 0 minutes was a possible predicted value. Due to the small sample size, as opposed to a typical single training/testing split for regression calibration (e.g. Metcalf et al., 2018; Saint-Maurice et al., 2014; Welk et al., 2017), a 5-fold resampling with 50 repeats was performed using a cross validation approach to estimate model performance at a population level. Bland-Altman analyses were performed comparing accelerometry data to model predicted activity levels. The 'caret' package for R (Kuhn, 2008) was used to perform repeated 5-fold cross validation.

4.5.2 Results

Table 4.3a-b summarizes the results of the 5-fold repeated resampling Bland-Altman comparison of the predicted MVPA/SB data to observed accelerometry data and regression cross-validation. Due to the nature of calculating regressions, the mean difference tended to approach 0, except when the intercept was set to 0 *a priori*. For MVPA a 0 intercept (y = 0 + 0.6*IPAQ-MVPA) increased the degree of underestimation to 20 minutes/day. While for SB (y = 0 + 0.75*IPAQ-SB, the IPAQ now underestimated by 79 minutes/day, instead of overestimating by nearly 30 minutes.

In almost all calibration equations the LoA range was reduced by more than 10%, with SB benefitting the most. However, the variant of MVPA Model 1 with a 0 intercept only

improved LoA range by 3.4%, and given the 95% CI for the LoA, it is possible that this change is lower.

Mean absolute error (MAE) for MVPA model cross-validation was approximately 30 minutes/day for each model, indicating that on average model derived predictions were within 30 minutes/day of observed values. Standard deviations of MAE ranged from 1.0-1.9minutes. Mean R^2 indicators of model fit for MVPA was similar to what was observed in Analysis 2, but slightly lower for the models with only the IPAQ as a predictor.

Similar patterns were observed in the SB cross-validation. MAE was approximately 104 minutes/day with standard deviations ranging from 15.4 to 16.7 minutes/day. The exception was with a 0 intercept, in which case MAE increased to 175.8 minutes/day with a larger standard deviation of 27.2. Overall, mean R^2 were also similar to the initial regression testing, but standard deviations were nearly equal to the mean, suggesting a large degree of variance in model fit.

4.6 Discussion

While previous studies have addressed the accuracy of the IPAQ in reference to accelerometry (Duncan et al., 2019; Duncan, Arbour-Nicitopoulos, et al., 2017), no studies have attempted to identify what factors may be associated with measurement error. This paper contributes to the validity evidence for using the IPAQ in samples with schizophrenia by first attempting to identify such individual factors and second by attempting to compensate for those errors with regression calibration.

With respect to identifying possible factors for MVPA miscalibration, the difference between the IPAQ and accelerometry was positively correlated with BMI and IDF-WC status,

such that overestimation was correlated with higher BMI and exceeding IDF-WC status guidelines, whereas underestimation was associated with lower BMI and being below IDF-WC guidelines, similar to what Welk and colleagues (2017) had identified in an otherwise healthy adult population. Ekkekakis and Lind (2006) have previously shown that overweight women tend to perceive moderate intensity physical activity more affectively negative than non-overweight peers – a response that is associated with more vigorous intensities (Ekkekakis et al., 2011). Thus, the explanation for this relationship suggested by Welk and colleagues (2017) that overweight individuals may experience even light physical activity as taxing, and thus remember and report it as MVPA when using the IPAQ, appears plausible. IDF-WC status also correlated negatively with absolute MVPA difference, suggesting that individuals with higher weight status may be prone to error in estimating MVPA in general.

Higher anti-psychotic medication doses (CPZ) were also associated with greater underestimation, and lower doses associated with greater overestimation. Additionally, worse quality of life mental composite scores (lower SF12-MCS) were related to overestimation and higher SF12-MCS scores (better mental health) related to underestimation. CPZ and SF12-MCS also correlated positively with absolute difference. Taken together, these results suggest that individuals with higher CPZ dosage and better SF12-MCS scores tend to make errors of greater magnitude when using the IPAQ to estimate MVPA measured by accelerometer; however, the direction of the error is also related to these variables, such that higher CPZ dosage and SF12-MCS scores are associated with underestimation. Additionally, the regression analyses demonstrate that adding IDF-WC, SF12-MCS, or BMI significantly improves a model that already accounts for CPZ, suggesting that these variables are tapping into different sources or correlates of error between measures. These results suggest that general mental health related

descriptors (antipsychotic dose and general mental well-being) are related to IPAQ MVPA reporting accuracy.

Despite a plausible link between error and either symptomology (negative symptoms may reduce the effort participants are willing to expend in order to complete the questionnaire accurately) or cognitive impairment (due to poor recall or understanding of self-report questionnaires), measures of mental health status and symptom severity such as the BPRS, CGI-S, and AES did not correlate with the difference between measures, nor did any measure of cognitive ability. In light of the association observed with CPZ, this suggests that it may be the medication dose that is interfering with accurate recall such as associated drowsiness or sedation. However, it is worth noting that some atypical anti-psychotic medications are more sedating, even at equal CPZ doses (Miller, 2004), and this was not taken into account; doing so may clarify these exploratory results. Furthermore, the general mental well-being of the SF12 measures a component of mental health that is not specific to illness symptoms. Specifically, the SF12 includes items asking about whether emotional issues have interfered with accomplishing daily activities, energy, and general mood. Thus, individuals with a more positive view of their mental well-being may tend to underestimate the amount of MVPA they accomplish although reasons why this may be can only be speculated.

For SB, only measures derived from the BPRS Affective subscale and the AES were correlated with the difference between the IPAQ and accelerometry data such that higher scores were related to over reporting. The affective subscale includes items for anxiety, guilt, depression, and somatic concern for health (Shafer, 2005), suggesting that greater levels of emotional distress may lead to an overestimation of sitting time. Additionally, higher levels of apathy may reduce the amount of effort participants are willing to expend on attempting to recall

and as a result tend to overestimate. Age was also marginally correlated with absolute differences between measures, such that older participants tended to make greater errors in SB estimation, but in both directions. Interestingly, during the regression analyses scores derived from the BPRS Affective scale were no longer significant predictors once the IPAQ sitting scores were controlled within the regression model, perhaps because the relationship between the IPAQ and accelerometry does not appear to be one of equivalence (i.e. 1min IPAQ sitting ≠ 1min accelerometer SB) as illustrated by SB-Model 1.

In contrast, adding BPRS Resistance measures improved the regression model, despite not being statistically significant in the exploratory correlation process. Adding the AES measures to SB-Model 1 improved the explained variance to 14%, while the BPRS Resistance measures improved the explained variance to 16% with higher scores predicting underestimation. However, including both in the model simultaneously did not provide a significant improvement over either scale alone, suggesting a degree of overlap in explained variance (despite a low correlation between the scales: $r = 0.20$, $p = 0.07$). The BPRS Resistance subscale includes the Uncooperativeness, Hostility, and Suspiciousness items of the scale, while the AES represents levels of amotivation and apathy. It appears, as though willingness to expend effort or cooperate is related to accuracy of the IPAQ sitting scale. However, the question of directionality (over versus underestimation) is unclear.

It is unclear why variables related to variance between SB measures were different than MVPA. Due to the volume of SB that individuals tend to engage in, it may require more effort to provide an accurate recall, thus the AES and BPRS Resistance measures were related to explained variance between the accelerometer and IPAQ Sitting measures but not MVPA which

is a lower volume and potentially easier to recall. Contrarily, while we posit that weight status may impact perceptions of PA intensity, it is unlikely to affect perceptions of SB.

Despite the regression modelling approach suggesting that weight status, CPZ dosage, and general well-being may be associated with self-reporting error in MVPA and measures of apathy and resistance with error in SB, the present data provides limited utility for calibrating the IPAQ scores to better reflect accelerometry derived data. The intercept for MVPA-Models 1 and 2a are well above 0, such that even individuals who claim no MVPA will be achieving 32 or 22 minutes of MVPA per day respectively. Multiplied across the week, this method would make it appear as though every participant is meeting MVPA guidelines of ≥150 minutes per week. Using an approach that forces a 0 intercept and rescaling the IPAQ MVPA minutes (1min IPAQ = 0.75 min accelerometry) is one approach to circumvent this issue, however, when tested through 5-fold cross validation to anticipate performance in the broader population, this approach demonstrated even greater MVPA underestimation and little-to-no improvement on LoA range. The other option that circumvents the intercept issue is Model 3ac(y = -36.35 + 0.31*(IPAQ) + 0.01*(CPZ) + 1.17*(SF12-MCS)) due to the negative intercept. This model performed relatively well when assessed with repeated 5-fold resampling. Mean bias was low, LoA were reduced by 14.6% and the model explained 22% of the variance. However, the potential downside of this calibration approach is that the SF12-MCS as a self-report tool is, from a measurement error theory perspective, prone to error (Guolo, 2008), and CPZ equivalents were also self-reported, causing inherent error as well. If CPZ equivalents can be acquired through chart review, this issue can be reduced.

Similarly, regression equations in the SB Models all included a positive intercept of >300 minutes, thus any adjustment applied to the sitting scale of the IPAQ would not be able to

capture very low levels of SB if they exist within the sample. This may be less of a concern given the increasingly sedentary lifestyle of most populations. However, this issue could limit the implications of the data (e.g. attempting to develop guidelines), and the 0-intercept reiteration of SB Model 1 performed poorly relative to other options. Adding the BPRS-Resistance or AES measures improved the explained variance (17% and 16% respectively) when undergoing the 5-fold repeated resampling assessment but provided little added benefit over the IPAQ alone in terms of reducing mean bias or LoA.

Ultimately, only small improvements can be made on the IPAQ accuracy through calibration. In particular a calibration equation including CPZ and SF12-MCS scores may be useful for MVPA, while recalibration of the IPAQ sitting scale alone appears to be the option of least compromise, so long as detecting low levels of SB (≤5h) are not a priority. Regardless, as the cross-validation process only simulates a population level approach, data should be tested in additional larger samples.

Overall, despite the small sample size of the study, this is, to our knowledge, the first attempt to understand the correlates of error when measuring MVPA and SB in people with schizophrenia. Identifying such correlates can help design better self-report measurement tools and compensate for error in existing tools. In particular, intensity guidelines and cut points do not appear to be well suited towards communicating the intended constructs to overweight and unfit individuals (Ekkekakis et al., 2011), thus contributing to the error when self-reporting MVPA. Interview-based protocols (e.g., think-aloud procedures (Beauchamp & McEwan, 2017)) and other qualitative techniques may also help identify sources of error to improve measurement tools or delivery methods. Among individuals taking antipsychotics, assessing MVPA by self-report remains problematic. While the calibration approach described here provides limited

utility, statistically controlling for CPZ dosage appears to be necessary before attempting to relate IPAQ MVPA data to other health correlates. Ultimately, where feasible, device-assessed measures of PA and SB are still warranted when detail beyond a sample's mean activity level is of interest.

4.7 Tables

Table 4.1a *Complete Case Data for Sample Descriptors and Moderate-to-Vigorous Physical Activity (MVPA) Measurement Differences*

Descriptor	Mean	sd	n	MVPA Δ		MVPA |Δ|	
				r	*p*	*r*	*p*
Age	41.1	11.5	97	0.107	0.298	-0.144	0.160
Sex	58 male	39 female		0.149	0.144	-0.072	0.481
BMI	31.1	7.7	96	0.219	0.032*	-0.138	0.181
IDF-WC	21 below	72 above		0.339	0.001*	-0.347	0.001*
CPZ	781.3	1302.5	93	-0.340	0.001*	0.342	0.001*
SF12-PCS	30.1	4.8	92	0.157	0.136	-0.196	0.061
SF12-MCS	51.7	11.0	92	-0.252	0.015*	0.238	0.022*
Symptoms							
BPRS	33.7	7.4	97	-0.030	0.771	-0.063	0.540
BPRS-Activation	4.4	1.5	97	-0.115	0.263	0.038	0.711
BPRS-Affective	9.0	3.3	97	-0.025	0.806	-0.118	0.250
BPRS-Positive	7.7	3.3	97	-0.049	0.635	0.022	0.831
BPRS-Negative	7.2	3.1	97	0.039	0.707	-0.049	0.637
BPRS-Resistance	5.4	2.0	97	0.040	0.699	-0.030	0.768
CGI-S	3.4	1.1	97	-0.043	0.676	0.024	0.813
AES	31.3	7.9	97	0.055	0.592	-0.078	0.446
Cognitive Functioning							
Symbol Coding	45.1	13.2	80	-0.012	0.915	-0.004	0.974
Letter Number Span	13.0	3.6	80	0.111	0.326	-0.027	0.813

Δ = difference between measures (IPAQ-Accelerometer), |Δ| = absolute difference between measures |IPAQ-Accelerometer|, * = p<.05

Table 4.1b *Complete Case Data for Sample Descriptors and Sedentary Behaviour (SB) Measurement Differences*

Descriptor	Mean	sd	n	SB Δ	SB |Δ|

				r	p	r	p
Age	42.1	12.0	75	-0.012	0.918	0.230	0.047*
Sex	48 male	27 female	75	-0.044	0.711	0.121	0.302
BMI	30.5	7.2	74	-0.011	0.925	0.171	0.146
IDF-WC	16 below	57 above	73	-0.128	0.280	-0.153	0.197
CPZ	631.3	815.9	71	-0.076	0.530	-0.040	0.741
SF12-PCS	30.5	4.8	71	0.004	0.972	-0.131	0.276
SF12-MCS	51.3	11.4	71	-0.126	0.294	0.047	0.699
Mental Health							
BPRS	33.8	7.2	75	0.219	0.060	0.040	0.736
BPRS-Activation	4.5	1.5	75	-0.094	0.420	-0.068	0.560
BPRS-Affective	9.0	3.3	75	0.248	0.032*	0.070	0.550
BPRS-Positive	7.5	3.5	75	0.076	0.519	0.055	0.638
BPRS-Negative	7.4	3.3	75	0.062	0.600	0.039	0.737
BPRS-Resistance	5.5	2.1	75	0.202	0.083	-0.085	0.470
CGI-S	3.4	1.0	75	0.202	0.082	0.168	0.150
AES	30.8	7.5	75	0.239	0.039*	-0.106	0.364
Cognitive Functioning							
Symbol Coding	44.4	12.5	63	-0.006	0.964	-0.080	0.530
Letter Number Span	13.1	3.6	63	0.087	0.500	0.159	0.213

Δ = difference between measures (IPAQ-Accelerometer), |Δ| = absolute difference between measures |IPAQ-Accelerometer|, * = p<.05

Table 4.2a *Moderate-to-Vigorous Physical Activity Regression Model Results*

Model	R^2	*F*	df	*p*	R^2 improvement	R^2 improvement *p*	Equation
1	0.08	7.73	1,95	0.007*	-	-	32.57 +0.26 (IPAQ)
2a	0.19	10.75	2,90	<0.001*	0.11	<0.001*	22.10 +0.30(IPAQ) +0.01 (CPZ)
2b	0.22	12.54	2,90	<0.001*	0.20	<0.001*	70.73 +0.13 (IPAQ) -44.19 (IDF-WC)
2c	0.19	10.33	2,89	<.001*	0.11	0.001*	-33.06 +0.28 (IPAQ) +1.29 (SF12-MCS)
2d	0.14	7.82	2,93	<.001*	0.06	0.007*	78.76 +0.26(IPAQ) -1.49 (BMI)
2e	0.10	5.25	2,94	0.007*	.02	0.11	60.53 +0.24(IPAQ) -0.88 (AES)
2f	0.08	3.86	2,94	.025*	0	0.80	29.13 +0.26 (IPAQ) +0.99 (CGI)
2g	0.08	3.86	2,94	0.024	0	0.78	38.08 +0.25 (IPAQ) -0.16 (BPRS)
2h	0.08	3.94	2,94	0.023*	0	0.64	37.83 +0.26(IPAQ) -0.97 (BPRS-Resistance)
2i	0.08	4.12	2,94	0.019*	0	0.46	40.04 +0.26 (IPAQ) +1.03 (BPRS-Negative)
3ab	0.31	12.57	3,85	<.001*	0.17	<0.001*	59.61 +0.18 (IPAQ) +0.01 (CPZ) -41.38 (IDF-WC)
3ac	0.29	11.4	3,84	<.001*	0.08	0.003*	-36.35 + 0.31 (IPAQ) + 0.01 (CPZ) + 1.17 (SF12-MCS)
3ad	0.25	9.65	3,88	<.001*	0.06	0.013*	65.24 +0.30 (IPAQ) +0.01 (CPZ) -1.36 (BMI)

* = $p<.05$

Table 4.2b *Sedentary Behaviour Regression Model Results*

Model	R^2	*F*	df	*p*	R^2 improvement	R^2 improvement *p*	Equation
1	0.082	6.48	(1,73)	0.013*	-	-	342.40 +0.16(IPAQ)
2a	0.085	3.36	(2,72)	0.04*	0.003	0.59	362.27 0.17 (IPAQ) -2.67 (BPRS-Affective)
2b	0.144	6.10	(2,72)	0.004*	0.062	0.02*	479.00 +0.18(IPAQ) -4.68 (AES)
2c	0.091	3.60	(2,72)	0.03*	0.009	0.39	383.48 +0.17 (IPAQ) -13.24 (CGI)
2d	0.114	4.65	(2,72)	0.01*	0.032	0.11	453.99 +0.17 (IPAQ) -3.49 (BPRS)
2e	0.157	6.68	(2,72)	0.002*	0.075	0.01*	434.93 +0.17(IPAQ) -17.68 (BPRS-Resistance)
2f	0.090	3.55	(2,72)	0.03*	0.008	0.43	370.06 + 0.16 (IPAQ) -3.77 (BPRS-Negative)
2g	0.086	3.39	(2,72)	0.04*	0.004	0.55	310.54 + 0.16 (IPAQ) + 0.77 (Age)
2h	0.057	2.07	(2,68)	0.13	0.002	0.69	371.83 +0.12 (IPAQ) -0.01 (CPZ)
2i	0.115	3.98	(2,60)	0.025*	0.000	0.94	316.18 + 0.20 (IPAQ) + 0.10 (BACS)
2j	0.128	4.40	(2,60)	0.016*	0.012	0.35	265.82 + 0.18 (IPAQ) + 4.59 (LNS)
3eb	0.197	5.79	(3,71)	0.001*	0.040	0.06	531.87 +0.18(IPAQ) -14.99 (BPRS-Resistance) -3.80 (AES)

* = $p<.05$

Table 4.3a *Average Daily MVPA 5-fold Repeated Resampling Bland Altman Results & Cross Validation*

Adjustment	Bias			Low LoA			High LoA			LoA	RMSE	R^2	MAE
	Mean	Lower CI	Upper CI		Lower CI	Upper CI		Lower CI	Upper CI	Range Change	(sd)	(sd)	(sd)
IPAQ-Accel	-9.4	-20.1	1.4	-113.7	-132.2	-95.3	95.0	76.5	113.9	-	-	-	-
Model 1	0.9	-1.0	2.8	-92.4	-95.7	-89.2	94.2	90.9	97.4	-10.6%	47.6 (1.4)	0.03 (0.03)	29.9 (1.0)
Model 1 w/0 intercept	-19.6	-21.6	-17.5	-120.4	-123.9	-116.9	81.2	77.7	84.7	-3.4%	55.0 (1.7)	0.03 (0.03)	31.7 (1.3)
Model 2a (CPZ)	6.4	4.5	8.4	-85.5	-88.8	-82.1	98.3	94.9	101.6	-12.0%	47.2 (2.1)	0.14 (0.02)	32.2 (1.3)
Model 2b (IDF)	-6.5	-8.3	-4.7	-91.1	-94.1	-88.0	78.0	74.9	81.1	-19.0%	43.6 (1.7)	0.24 (0.04)	26.9 (1.2)
Model 3ab (CPZ + IDF)	-2.0	-3.9	-0.2	-85.3	-88.5	-82.2	81.3	78.2	84.5	-20.1%	42.5 (1.8)	0.29 (0.05)	28.8 (1.4)
Model 3ac (CPZ + SF12_MCS)	3.9	1.9	5.9	-85.3	-88.7	-81.8	93.0	89.6	96.5	-14.6%	45.5 (3.0)	0.22 (0.05)	30.0 (1.9)
Model 3ad (CPZ+BMI)	3.2	1.3	5.2	-86.3	-89.6	-83.0	92.8	89.5	96.1	-14.2%	45.8 (1.9)	0.19 (0.02)	30.8 (1.3)

Table 4.3b *Average Daily SB 5-fold Repeated Resampling Bland Altman Results & Cross Validation*

Model		Bias			Lower LoA			Upper LoA		LoA	RMSE	R^2	MAE
	Mean	Lower CI	Upper CI		Lower CI	Upper CI		Lower CI	Upper CI	Range Change	(sd)	(sd)	(sd)
IPAQ-Accel	28.7	-27.96	85.3	-453.8	-551.0	-356.6	511.1	413.9	608.3	-	-	-	-
Model 1	0.0	-4.3	4.2	-261.9	-269.2	-254.5	261.8	254.5	269.1	45.7%	132.4 (17.5)	0.13 (0.12)	104.3 (15.4)
Model 1 w/0 intercept	-79.1	-85.5	-72.7	-471.2	-482.1	-460.2	313.1	302.1	324.0	18.7%	213.2 (29.3)	0.13 (0.12)	175.8 (27.2)
Model 2b (AES)	-0.8	-5.0	3.4	-258.7	-265.9	-251.5	257.1	249.9	264.3	46.5%	130.4 (16.6)	0.16 (0.13)	103.8 (15.4)
Model 2e (BPRS Resistance)	-0.4	-4.6	3.7	-255.3	-262.4	-248.2	254.4	247.3	261.5	47.2%	128.6 (18.8)	0.17 (0.12)	103.3 (16.5)
Model 3eb (BPRS Resistance +AES)	-1.4	-5.6	2.8	-256.1	-263.2	-249.0	253.3	246.2	260.4	47.2%	128.6 (17.6)	0.17 (0.11)	103.8 (16.7)

4.8 Figures

Figure 4.1 *MVPA Model Notations, Predictors, and Progression*

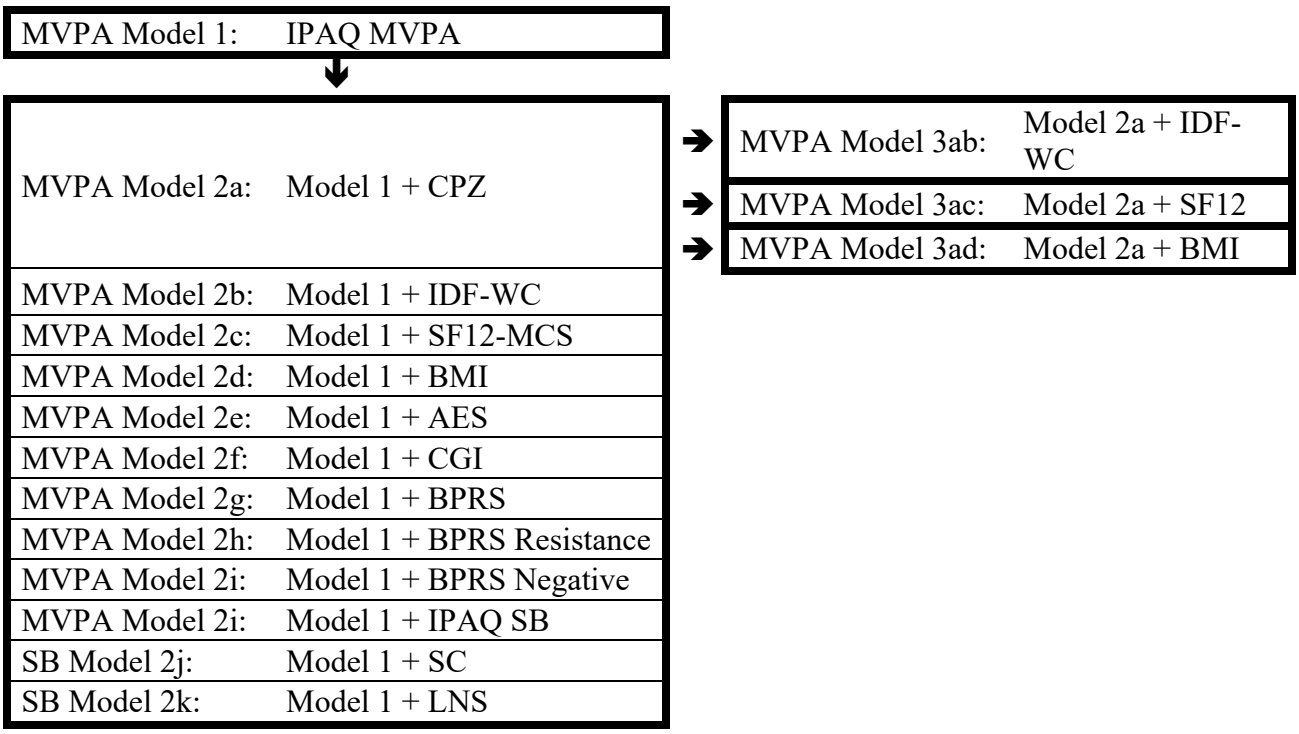


Figure 4.2 SB Model Notations, Predictors, and Progression

Chapter 5: Accelerometry Wear and Movement Behaviour Analysis

5.1 Introduction

Behavioural approaches are an essential component of addressing the physical health needs of individuals with schizophrenia. The high rates of obesity, diabetes, metabolic syndrome, and cardiovascular disease in people with schizophrenia (Bueno-Antequera et al., 2017; McNamee et al., 2013; Vancampfort, De Hert, et al., 2014; Vancampfort, Probst, Knapen, et al., 2012; Vancampfort, Rosenbaum, et al., 2016) contribute to a 15-25 year shortened lifespan among people with schizophrenia (Laursen et al., 2012, 2014). Movement behaviours (MBs), such as increasing physical activity (PA), are one such behavioural target used to redress these chronic physical health conditions in the general population.

MBs of the waking day consist of the amount of time spent along a continuum of energy expenditure ranging from low/basal energy expenditure to high expenditure and is commonly measured by metabolic equivalents (METs, the amount of energy expended relative to a resting state). MBs of the waking day consist of sedentary behaviour (SB) – any waking activity characterized by an energy expenditure less than 1.5 METs in a sitting, reclining or lying posture (Tremblay, Aubert, et al., 2017) – and PA (any bodily movement produced by skeletal muscles that requires energy expenditure) (World Health Organization, 2014). PA is further categorized by the intensity of energy exertion ranging from light intensity PA (LPA; 1.5-to <3 METs) such as standing, casual walking, and other incidental activities of daily living, moderate intensity (MPA; 3-6 METs) such as yoga, brisk walking, or hiking, and vigorous intensity PA (VPA; >6 METs) such as jogging or swimming laps (American College of Sports Medicine, 2014).

The cardiovascular and physical health benefits of engaging in moderate-to-vigorous PA (MVPA) are well understood in the general population (Warburton et al., 2010) and MVPA appears to also benefit the physical health of individuals with schizophrenia and severe mental

illness (Caemmerer et al., 2012; Faulkner, Cohn, Remington, et al., 2007). Currently, the World Health Organization suggests adults accrue ≥150min of MPA per week or ≥75 min of VPA for optimal health benefits (World Health Organization, 2014).

The impact of SB on physical health appears to be equivocal and is an emerging area of research. Early evidence suggests that SB seems to be related to weight, cardiovascular health, and all-cause mortality in adults, after adjusting for MVPA (Biswas et al., 2015; Warren et al., 2010). However, a common criticism of this research is that when assessing the relationship of SB with health outcomes *only* MVPA is adjusted for and not LPA (Pedišić, 2014; van der Ploeg & Hillsdon, 2017). Additionally, observation periods have a limited time budget available and therefore increasing any one MB inherently requires that another be decreased (Pedišić, 2014). As van der Ploeg and Hillsdon (2017) state: "If no adjustment is made for light physical activity then it is difficult to determine whether the association with poor health is due to more time sedentary or less time in light activity." Thus, the benefits of decreasing SB must be considered with regards to which PA behaviour replaces SB. Similarly, the benefits of increasing the amount of time spent in MPA or VPA may be different depending on the behaviour it replaces; that is MPA may have a greater benefit when replacing SB than when it replaces LPA. Assessing all MBs of daily life simultaneously also presents a statistical problem, as this codependence between behaviours results in the issue of multicollinearity when including MB variables in statistical models.

In response to these issues, isotemporal substitution analysis approaches – that is, approaches which examine the effects of swapping time spent in one MB for another – have become an established method of assessing the relationship between MBs and health and well-being. The seminal paper by Mekary and colleagues (2009) outlined the application of

isotemporal substitution using absolute values (i.e. minutes of behaviour). Absolute isotemporal substitution modelling estimates the effect of substituting time spent in one MB for another by creating multiple models for the same outcome, each model excludes a different MB in exchange for the total observation time (the sum of all observed MBs). The effect of replacing a unit of time in the omitted behaviour corresponds with the model's effect estimates (β) for each of the remaining MBs included in the model. Repeated modelling with each MB excluded allows for the estimation of every possible time substitution combination (i.e. if four MBs are included, 4 separate models for each dependent variable are required). While absolute modelling estimates the effect of substituting one behaviour for another it has been criticized for not adequately accounting for the constrained nature of MB observation periods (Dumuid et al., 2019; Pedišić et al., 2017). As an alternative, compositional data isotemporal substitution analysis accounts for the proportion of time spent in each MB within a simplex of limited time by addressing issues of collinearity and quantifying the effect of behaviour substitution, between all behaviours.

Both absolute and compositional isotemporal substitution approaches have grown in popularity for assessing the epidemiology of time use and MB (e.g. Biddle et al., 2018; Gupta et al., 2018). When both absolute and compositional data isotemporal substitution approaches are compared in the same data set, they reach similar conclusions about whether MBs are associated with dependent variables (e.g. Biddle et al., 2018; Chastin et al., 2015; Gupta et al., 2018) and the resultant models have been shown to be mathematically similar (Mekary & Ding, 2019). Absolute isotemporal substitution has the advantage of describing behaviour in units of time which are useful for interpreting results (Mekary & Ding, 2019), however the proportional results of compositional data can also be readily converted to a 24h day (Biddle et al., 2019; Chastin et al., 2015). Several jurisdictions, such as Canada and Australia (Australian

Government & Department of Health, 2019; Tremblay et al., 2016; Tremblay, Chaput, et al., 2017), have already adopted 24-hour guidelines for time use allocation in children and youth recommending how much of the day should be spent in sleep and various MBs during the waking day and 24-hour guidelines for adults will be released in 2020 (Ross et al., In Review). In addition to statistically accounting for the constrained nature of time use and guidelines adopting proportional allocation of MBs within a 24h period, compositional data approaches are also able to identify the co-dependent relationship between MBs, that is how much time in one MB influences time in another. Standard correlations may not adequately capture these relationships due to the enclosure of MBs within the limited time simplex (Chastin et al., 2015).

However, when examining the effect of time reallocation using recommendations for compositional analysis, the results are asymmetrical (e.g. Biddle et al., 2018; Chastin et al., 2015; Gupta et al., 2018). That is, for example, substituting 10 minutes of SB with MPA, is not equal to the reciprocal substituting of 10 minutes of MPA with SB. Nor are the effects linear, exchanging 30 minutes of MPA does not have 3 times the effect of exchanging 10 minutes of MPA. This is not the case in absolute isotemporal substitution modelling. This observed asymmetry has been suggested to be the result of log-ratio data transformation required to conduct compositional data (Mekary & Ding, 2019) as the reallocation of time from a MB in large quantity to a small quantity has a more drastic relative change than a reallocation in the other direction (Chastin et al., 2015). For example, in a 16h observation period of waking day MBs, if SB constitutes 50% of the period and MPA constitutes 3%, exchanging 10 minutes of SB (~1% of the day) for MPA (while keeping all other behaviours constant) results in MPA having a 33% increase relative to the original proportion, whereas the proportion of SB only decreases by 2% relative to the initial proportion. Intuitively, from an energy expenditure

perspective, replacing one behaviour for another should have equal and inverse effects. As a result of these differences, there is little consensus on which approach to isotemporal substitution is most appropriate (Biddle et al., 2019; Mekary & Ding, 2019). Given that energy expenditure is at the core of MB classifications, substitution effects should be coherent with this underlying system of classification. It may therefore be prudent to use both absolute and compositional isotemporal approaches in tandem, as compositional analyses may be best able to detect if, overall, MBs are associated with a variable or differ between groups with fewer violations of statistical assumptions, while absolute isotemporal substitution approaches provide untransformed estimates of exchanging one MB with another, and thus are easier to interpret and align with the presumption that increasing or decreasing energy expenditure in equal amounts should have equal but opposite effects on health outcomes.

While isotemporal substitution analysis approaches have been used to simultaneously control for SB, LPA and MVPA in a variety of populations, it has yet to be applied among individuals with schizophrenia. LPA may be a particularly beneficial behavioural target among individuals with schizophrenia if it is indeed associated with better health and well-being, as more effortful PA intensities may be more difficult to maintain among individuals with high levels of amotivation – a characteristic symptom of schizophrenia. Additionally, for individuals with high body mass index or low aerobic capacity, the affective experience of lower intensity PA tends to be more pleasant than higher intensities (Ekkekakis et al., 2011), and thus LPA may be a more amenable target for such individuals to commence and sustain than ≥MPA. Even if LPA only has modest effects on health and well-being, it may also be easier to replace LPA as an intermittent stage than moving directly from SB to MPA.

Accelerometry, when feasible in terms of expense and delivery, provides an optimal solution for assessing and profiling daily MBs when trying to quantify time and intensity among individuals with schizophrenia compared to self-report. Accelerometry allows measurement of timestamped movement frequency and does not rely on the capacity of an individual to recall how much activity they did and at what intensity. The analysis of accelerometry data does not require subjective judgement on the categorical intensity of PA by the participant as these variables are assessed post-hoc by adjustable algorithms; thus, accelerometry also requires fewer *a priori* judgements by the researcher of what types of PA are relevant to assess. Unlike self-report methods, the only recall required is to wear the accelerometer device and return it for analysis.

Furthermore, the International Physical Activity Questionnaire (IPAQ), which is currently the most frequently used self-report tool used to assess PA in samples of people with schizophrenia (Stubbs, Firth, et al., 2016; Stubbs, Williams, et al., 2016), only explicitly asks for walking time as a form of LPA (Craig et al., 2003), and thus may not capture tasks like light chores or standing. In our recent evaluations of IPAQ derived estimates of ≥MPA and SB in individuals with schizophrenia compared to accelerometry (Duncan et al., 2019; Duncan, Arbour-Nicitopoulos, et al., 2017) the overall sample means reported on the IPAQ tend to be quite close to accelerometer derived values, which is in line with the IPAQ's original intention as a population level assessment of MBs (Bauman, Ainsworth, et al., 2009; Bauman, Bull, et al., 2009; Craig et al., 2003). However, when individual level estimates are evaluated, there are wide limits of agreement (Bland & Altman, 1986, 2003) between measurements from accelerometry and the IPAQ (Duncan et al., 2019; Duncan, Arbour-Nicitopoulos, et al., 2017), suggesting that any two individuals who report the same amount of PA on the IPAQ may have very different

measurements of PA when accelerometry is used instead. This is particularly problematic when attempting to evaluate associations between MBs and sample characteristics such as health status.

The use of accelerometry to assess MBs in individuals with schizophrenia has numerous benefits over self-report approaches. However, while it is easy to assume that device-based measurement is more accurate than a self-report questionnaire, this assumes that the devices are being worn appropriately. Sufficient device wear time is essential to accurately estimate daily MBs; a consistently adopted guideline has been ≥600 minutes of wear per day and ≥3 weekdays and ≥1 weekend day to be considered representative of a week (Troiano et al., 2008; Trost et al., 2005) (although higher amounts of daily wear have been suggested to improve MB estimates (Herrmann et al., 2013). If individuals are systematically not achieving these wear time protocols due to individual or external factors, estimates of MBs may be biased or generalize poorly to other populations. Given the nature of the population, a focused analysis of wear time is needed as it might provide insight for other researchers planning to use accelerometry.

The overall purpose of this study was to assess accelerometer derived movement data in a sample of individuals with schizophrenia, using a combination of absolute and compositional isotemporal substitution approaches. First, accelerometer protocol adherence was assessed to identify potential sources of bias in the data consisting of: a) whether observation characteristics (e.g. season, day of the week) affected the likelihood of each day having sufficient wear among participants receiving an accelerometer; b) among days with sufficient wear identify whether observation characteristics influenced total daily accelerometer wear time; and c) determine if observation and participant factors were associated with meeting the minimum number of days with sufficient wear. Second, for the sample that met the minimum protocol adherence,

participant MB during the observation period was assessed to describe how the composition of MBs were distributed across the week. Third, the relationship between MBs and participant factors were assessed using a compositional data approach to identify whether time use differed between participants in the sample. Finally, the relationship between MB composition and health factors in the sample were examined, and an absolute isotemporal substitution approach was used to describe the effects of MB substitution.

5.2 Methods

5.2.1 Participants

The present study is a secondary data analysis of a study assessing the psychological determinants of achieving Canadian PA guidelines for adults (≥150min/week of ≥MPA) in individuals with schizophrenia and schizoaffective disorder (Arbour-Nicitopoulos et al., 2017) and has been previously used to compare estimates of ≥MPA and SB time between accelerometry and self-report (Duncan et al., 2019; Duncan, Arbour-Nicitopoulos, et al., 2017). Research ethics boards at the Centre for Addiction and Mental Health in Toronto and the University of Toronto approved the original study. To be included participants had to: (1) be between age 18–64 years and (2) have a diagnosis of schizophrenia or schizoaffective disorder, which was confirmed with the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998). Exclusion criteria consisted of: (1) being hospitalized over the past 12 months for angina pectoris, myocardial infarction, or cardiac surgery of any kind; and/or (2) uncontrolled hypertension (defined as blood pressure>140/90 mmHg). Participants were recruited by referrals from nurses, psychiatrists, and other studies involving persons with schizophrenia at the Centre for Addiction and Mental Health. Eligible participants provided written consent prior to

commencing the study and the MacArthur Competence Assessment Tool for Clinical Research (MacCAT-CR; Appelbaum and Grisso, 2001) assessed capacity to consent.

Available sample descriptors included: age; sex; smoking status, diagnosis, ethnicity; self-reported medication dose measured by chlorpromazine equivalents (CPZ) (Gardner et al., 2010), employment status, education, marital status and living arrangements. Available health variables were: body mass index (BMI); International Diabetes Federation Waist Circumference classification (IDF-WC) dichotomized as high or low relative to sex and ethnicity cut points for metabolic syndrome criteria (IDF, 2006); self-reported health related quality of life assessed by the 12-item Short Form Health Survey version 2 (SF-12) Physical Composite Score (PCS) and Mental Composite Score (MCS) (Ware, 2002); symptom-specific variables including: the Brief Psychiatric Research Scale 18-item Anchored version (BPRS) (Woerner et al., 1988), Clinical Global Impression Severity Scale (CGI-S) (Guy, 1976a), Apathy Evaluation Scale (AES) (Marin et al., 1991). Subscales for the BPRS were also derived based on the schizophrenia appropriate 5-factor model identified by Shafer (2005): BPRS-Activation, BPRS-Affective, BPRS-Positive, BPRS-Negative, and BPRS-Resistance. Additionally, a subsample (n = 93) of the study had two cognitive measures available which were added in an amendment after the study had commenced: a symbol coding task – measuring speed of processing and executive function – and letter number span task – measuring working memory – with higher scores in each representing better functioning, a composite cognitive score was calculated by summing the respective z-scores for each test (Fervaha et al., 2014).

5.2.2 Accelerometer Data Processing

Participants were given an accelerometer to wear on their right hip for seven full days (i.e. provided an accelerometer on a Monday and instructed to return on a following Tuesday). If a participant was late to return the device, additional recorded days beyond the intended return period were excluded from the current analyses.[2] Accelerometers sampled movement at 30Hz and data were analyzed in 60s epochs, using Actigraph's ActiLife software (v6.12). Wear time was calculated using Choi and colleagues' (2011a) algorithm (≥90 minutes of consecutive of zero counts, with tolerance for ≤2 minutes of nonzero counts outside the first or last 30 minutes of the period). To be considered a day with sufficient wear, the accelerometer needed to register 600 minutes of wear time (Troiano et al., 2008).

Often, ≥3 weekdays and ≥1 weekend day with sufficient wear is required for accelerometry data to be considered representative of a week (Troiano et al., 2008; Trost et al., 2005); in the previous studies from this project ≥4 days regardless of status (weekday vs weekend) were considered representative under the assumption that weekends were unlikely to differ substantially from weekdays in a population with low levels of employment (Duncan et al., 2019; Duncan, Arbour-Nicitopoulos, et al., 2017). Again, the more liberal approach of ≥4 days with sufficient wear was taken, in part to allow this assumption to be examined.

Troiano et al.'s (2008) adult PA cut-off points were used to evaluate intensity of movement within each epoch of wear: LPA (100-2019 counts per minute; cpm), MPA (2020-5998 cpm), VPA (≥5999 cpm). Anything below ≤99 counts per minute was classified as possible

[2] This may result in slight variations in PA compared to those reported in (2019; 2017), as days analyzed were those that aligned with the 7 days which the IPAQ sought recall for.

SB; while this operationalization does not account for body position (sitting or lying) it has been shown to agree well with thigh worn accelerometry estimates of SB that do. As participants did not keep a record of their accelerometer wear time and had the option to wear their accelerometer while sleeping if it helped them remember to wear it during the waking day, wear strategies inherently varied among participants, capturing sleep in some but not others. Furthermore, at the time of analysis, ActiLife did not offer sleep detection for waist worn accelerometers that had been assessed for evidence of validity in an adult sample. To try and minimize the amount of sleep misclassified as SB, a filter was applied to omit possible SB between the hours of 0h00 and 5h59 as likely sleep, as was previously done when assessing the accuracy of the IPAQ-SF sitting scale in this sample (Duncan et al., 2019). Overall, valid wear time between 6am-12pm was classified into SB, LPA, MPA, and VPA categories.

5.2.3 Statistical Analysis

Statistical analyses were performed in R v3.6.1 (R Core Team, 2019). To test whether distribution of dichotomous response variables was equal between nominal independent variable categories Chi-square goodness of fit tests were conducted when response variables were not repeated measures. The "*rcompanion*" package (Mangiafico, 2020) was used to perform pairwise post-hoc Chi-square tests with Bonferroni adjustment to p-values when an omnibus Chi-square test detected significant difference between multiple categories.

For repeated measures response variables (i.e. multiple days of observation per participant), mixed model regression approaches were used. Mixed model regression models were assessed using the "*lme4*" package (Bates et al., 2019) including subjects as a random intercept factor, with restricted maximum likelihood used to optimize standard linear mixed

models. To assess differences between independent variable categories within mixed model regressions, the "*car*" package (Fox et al., 2019) was used to apply analysis of variance (ANOVA) tests (for standard linear mixed models) and Wald chi-square tests (for logistic mixed models) with type II Sum of squares to allow for unbalanced categories and Kenward-Rogers *F*-test with Satterthwaite adjustment for residual degrees of freedom. The "*emmeans*" package (Lenth, 2019) was used to assess pairwise differences in model estimated marginal means (least square means/predicted marginal means) between categories in statistically significant ANOVA and Wald tests using Tukey's adjustment to *p*-values for multiple comparisons. All *p*-values for post-hoc tests are reported with the adjustment applied such that $\alpha = 0.05$ remains the threshold for being statistically significant.

5.2.3.1 Compositional Analysis of MBs

Current guides for performing and analyzing MB using compositional analysis were followed to conduct analyses (Chastin et al., 2015; Comas & Thio-Henestrosa, 2011; Dumuid et al., 2018, 2019; McGregor et al., 2017); while compositional analysis of MBs often uses a 24h observation period, any simplex where available time is finite, such as the waking day, can be analyzed (Chastin et al., 2015; McGregor et al., 2019; Štefelová et al., 2018). First, the composition of MBs were summarized by transforming the ratio of time spent in each activity to a set of pivot coordinates using isometric log ratio (ILR) transformations. The "*robComposition*" package (Templ et al., 2011) was used to produce ILR pivot coordinates. With four behaviours, a set consists of three ILR pivot coordinates; Figure 5.1a-b illustrates the process for two possible pivot coordinate sets. The first pivot coordinate describes the natural logarithm of the ratio of the proportion of time spent in one behaviour (e.g. SB) to the geometric mean of all three remaining

behaviours. Each subsequent pivot omits the previous numerator, and compares the remaining behaviours in a similar manner, until the final pivot, which is the ratio between just two behaviours. As ILR transformations are not conducive to 0s, the "*zCompositions*" package (Palarea-Albaladejo & Martin-Fernandez, 2015) was used to impute totals less than the 1-minute detection limit of the accelerometer for any movement behaviours with cumulative 0 minutes; VPA was the only MB variable affected.

Due to the nature of compositional analysis, any permutation of ILR pivot coordinates between the same variables will produce statistical models with identical overall fit (explaining equal variance (R^2), p-value, y-intercept), so long as all coordinates within the set are included in the model. However, statistical testing with multiple permutation sets is commonly used to examine whether the first pivot coordinate in each model is statistically associated with the outcome (Chastin et al., 2015; Dumuid et al., 2018, 2019). Thus, four sets of pivot coordinate permutations were produced in order to isolate each of the four MBs as the numerator in the ratio of the first ILR pivot coordinate (specifically $ILR_{SB:LPA\cdot MPA\cdot VPA}$, $ILR_{VPA:MPA\cdot LPA\cdot SB}$, $ILR_{MPA:VPA\cdot LPA\cdot SB}$, $ILR_{LPA:VPA\cdot MPA\cdot SB}$). Multivariate ANOVA (MANOVA) tests were used to assess whether the MB composition differed between categorical variables while controlling for relevant covariates, with follow-up ANOVA tests for the first ILR pivot variable. For continuous variables, regression models consisting of the ILR coordinates representing the MB composition predictors with covariates were compared to a null model with any covariates as predictors using ANOVA comparison and whether the Akaike Information Criterion (AIC) decreased despite adding additional terms.

The relationships between MB composition and observation factors (wear time, number of valid days, season, and month) were assessed first. Potential confounding observation factors

were used as covariates in subsequent models of participant descriptors (sex, age, ethnicity, diagnosis, marital status, education, employment, antipsychotic CPZ dose, current smoking status). Observation and participant factors were used as covariates in models of health and well-being factors (BMI, waist circumference, symptom scores, cognitive health scores, and general physical and mental health status).

Absolute isometric substitution modelling approaches were subsequently adopted for health metrics to identify untransformed effects of MB substitution. Again, daily average time spent in each MB was used. Four models for each variable were used, each dropping a different MB for total time with classified MBs.

5.3 Results

5.3.1 Participants

Table 5.1 describes the sample that received an accelerometer compared to those that met the protocol adherence requirement of ≥4 days with sufficient wear. For subsequent analyses, categorical variables with low group membership were collapsed into larger categories. Employment status was dichotomized based on whether time may be occupied by working (part or full time) or studying ($n_{all} = 41$, $n_{adherence} = 36$) compared to not working or studying (unemployed, retired, other; $n_{all} = 72$, $n_{adherence} = 65$). The one underweight individual was grouped with other individuals below the overweight threshold (BMI<25) ($n_{all} = 20$, $n_{adherence} = 19$). As a high school diploma is not required for trade schools and can be earned concurrently, respondents were grouped as having a high school/trade diploma. Marital status was dichotomized as currently married ($n_{adherence} = 7$) or unmarried ($n_{adherence} = 94$).

5.3.2 Wear Time Assessment

Table 5.2 summarizes the frequency of days with sufficient wear time separated by weekdays and weekends. Of 113 participants to be given an accelerometer, 101 met the inclusion criteria of at least 4 days with ≥600 minutes of wear; all but 3 of these participants would not have met traditional 2-criteria inclusion requirement of 3 weekdays and one weekend day with ≥600 minutes of wear, in all cases it was due to insufficient number of weekdays with sufficient wear (2 weekdays and 2 weekend days). Table 5.2 is shaded to illustrate participants who met inclusion criteria for wear.

A Wald test found that there was no significant difference in the likelihood of having sufficient wear time between days of the week ($\chi^2(6) = 10.36$, $p = 0.11$). However, mixed model logistic regression accounting for random participant effects found a small but significant increased likelihood of having sufficient wear on a weekend day (OR: 1.71, 95% CI: 1.01 – 2.87, $\chi^2(1) = 4.03$, $p = 0.045$) compared to weekdays. A Wald test also indicated that observation day order affected likelihood of meeting sufficient wear ($\chi^2(6) = 39.06$, $p \leq 0.001$); post-hoc tests indicated observation day 1 vs 2 (OR = 0.08, 95% CI: 0.02 – 0.37, z= -4.91, $p_{adjusted} < .001$), 1 vs 3 (OR = 0.09, CI: 0.02 – 0.37, z = -4.91, $p_{adjusted} < .001$), 1 vs 4 (OR = 0.19, CI: 0.05 – 0.67, z = -3.88, $p_{adjusted} = 0.002$), and 1 vs 6 (OR = 0.26, 95% CI: 0.08 – 0.86, z = -3.33, $p_{adjusted} = 0.015$) differed significantly, Day 1 vs 5 (OR = 0.34, 95% CI: 0.11 – 1.08, z = -2.76, $p_{adjusted} = 0.084$), 2 vs 5 (OR = 4.01, 95% CI: 0.89 – 18.1, z = 2.71 $p_{adjusted} = 0.095$), 2 vs 7 (OR = 4.36, 95% CI: 0.97 – 19.6, z = 2.89, $p_{adjusted} = 0.059$), and 3 vs 7 (OR = 4.36, 95% CI: 0.97 – 19.6, z = 2.89, $p_{adjusted}$ = 0.060) approached statistical significance. Wald tests found no significant difference in the number of wear days per season ($\chi^2(3) = 4.07$, $p = 0.254$) or month ($\chi^2(11) = 10.29$, $p = 0.504$).

Among days with sufficient wear, mean wear time was 1058.1 (309.5) per day. Repeated measures mixed model type II ANOVA found significantly more mean wear time on weekends than weekdays (Δ = 39.5, 95% CI: 5.0-74.1, $F(1, 553.1) = 5.0$, $p = 0.025$). No significant difference on wear time by specific day of the week was detected, $F(6, 549.6) = 1.6$, $p = 0.16$. However, observation day order significantly affected wear time among days with sufficient wear time, $F(6, 548.05) = 28.17$, $p < 0.001$. Post-hoc pairwise comparison of estimated marginal means with Tukey's adjustment for multiple comparisons indicated that the first day of observation was significantly lower ($p < 0.001$) compared to all other days (ranging from $\Delta_{\text{Day 1-7}}$ = - 220.8, 95% CI: -305.1 to -136.4, t(550) = -7.7, $p_{\text{adjusted}} < .001$ to $\Delta_{\text{Day 1-2}}$ = -302.9, -384.6 to -221.2, t(549) = -11.0, $p_{\text{adjusted}} < .001$); pairwise comparisons between day 2 and 7 also achieved statistical significance ($\Delta_{\text{Day 2-7}}$ = 82.1, -4.7 to 159.5, t(547) = 3.1, $p_{\text{adjusted}} = 0.029$), the difference between day 5 and 7 approached significance as well ($\Delta_{\text{Day 5-7}}$ = 76.2, -3.5 to 155.9, t(547) = 2.8, $p_{\text{adjusted}} = 0.07$). Neither season ($F(3, 200.3) = 0.57$, $p = 0.634$) nor month of observation period ($F(11, 257.78) = 0.72$, $p = 0.719$) significantly affected daily wear time among days with sufficient wear.

Chi-square tests found that individuals were no more likely to achieve ≥4 days of adequate wear regardless of season ($\chi^2(3) = 2.26$, $p = 0.52$) or month (χ^2 (11) = 8.58, $p = 0.661$) of observation period. Meeting the ≥4 day criteria also did not differ by sex ($\chi^2(1) < 0.01$, $p = 1$), IDF-WC status ($\chi^2(1) = 0.62$, $p = 0.43$), living arrangements ($\chi^2(3) = 3.44$, $p = 0.33$), employment ($\chi^2(1) = 0.10$, $p = 0.76$), marital status (χ^2 (1) = 0.60, $p = 0.81$), diagnosis ($\chi^2(1) <$ 0.01, $p = 0.98$), smoking status ($\chi^2(1) < 0.01$, $p = 1$), ethnicity ($\chi^2(3) = 4.48$, $p = 0.21$). An overall difference between education statuses $\chi^2(2) = 9.30$, df = 2, $p = 0.01$) was significant. Post-hoc pairwise Chi-square tests indicated that individuals with at least some post-secondary

education were more likely to have met the ≥4 day requirement than individuals with a high-school/trade diploma (OR = 13.4, 95% CI: 1.5 to 116.9, $\chi^2(1) = 6.27$, df = 1, $p_{adjusted} = 0.037$) and approached significance compared to those without a high-school diploma (OR = 13.6, 95% CI: 1.43 to 130.39, χ^2 (1) = 5.27, $p_{adjusted} = 0.065$). No difference existed between those with a high-school/trade diploma compared to those without (OR = 1.02, 95% CI: 0.25 to 4.16 ,χ^2 (1)< 0.01, $p_{adjusted} = 1$).

Logistic regression found no significant effect for age, BMI, CPZ, or AES total on likelihood of having ≥4 sufficient days. In the subset of participants (n = 93), symbol coding task, letter number span and combined z-scores on both tests also showed no association with having ≥4 days with sufficient wear. Higher CGI (OR = 0.45, 95% CI: 0.24 to 0.84, $p = 0.012$) and BPRS total score (OR = 0.93, 95% CI: 0.87 to 0.99, $p = 0.028$) both predicted lower likelihood of having ≥4 days of wear. Among BPRS subscales, BPRS Positive Symptoms was the only subscale predictive of lower likelihood of having at least 4 days with sufficient wear (OR = 0.85, 95% CI: 0.73 to 0.99, $p = 0.033$). Higher reported physical health on the SF12 Physical health composite score was predictive of greater likelihood of having ≥4 days of sufficient wear (OR =1.17, 95% CI: 1.03 to 1.32, $p = 0.019$) but the Mental health composite score was not (OR = 0.99, 95% CI: 0.94 to 1.06, $p = 0.86$).

5.3.3 Daily MB Assessment

Table 5.3 summarizes the geometric mean of movement behaviours as a proportion of the observation period excluding days with insufficient wear, as well as arithmetic daily means of the sample. The category “Non-wear time and sleep filter” represents all time not captured by wear or SB excluded from 0h00 to 5h59. Minute equivalents represent the geometric mean

applied to a 24h day. Table 5.4 quantifies the codependence between movement behaviours. Values are the natural logarithm of the ratio between time spent in each behaviour; values below the diagonal are means, and above is the variance for each comparison. Mean values closer to 0 are indicative of a stronger relationship between behaviours. For example, the proportion of time in SB is inherently lower when the proportion of LPA, MPA, and VPA was higher, but the relationship with LPA was strongest suggesting that the proportions of SB and LPA remain similar relative to each other even when a greater proportion of MPA or VPA is observed. VPA was most independent of the behaviours, varying most closely with MPA, however MPA was more closely (and about equally) related to LPA or SB than it was to VPA.

Regressing accelerometer wear time across all valid days on an ILR pivot coordinate set with random participant intercept was a significant improvement over a null model with random participant intercept alone, $\chi^2(3) = 10.82$, $p = 0.013$. Among first pivot coordinates, only $ILR_{SB:LPA\cdot MPA\cdot VPA}$ was statistically significantly related to wear time ($\beta = 61.54$, 95% CI: 14.25 – 108.83, $p = 0.011$), such that on each day, wear time was associated with a greater proportion of SB relative to other behaviours.

Multilevel models cannot currently be conducted with multivariate dependent variables in "*lme4*", and standard MANOVA excludes missing values listwise, therefore an omnibus MANOVA test was omitted for the effect of observation factors on MB compositions.[3] Instead multilevel model linear models were produced with each of the first four ILR pivot coordinates ($ILR_{SB:LPA\cdot MPA\cdot VPA}$, $ILR_{VPA:MPA\cdot LPA\cdot SB}$, $ILR_{MPA:VPA\cdot LPA\cdot SB}$, $ILR_{LPA:VPA\cdot MPA\cdot SB}$) as dependent

[3] Multivariate general linear mixed models are however possible with Markov chain Monte Carlo generalized linear modelling (e.g. using the MCMCglmm package for R (Hadfield, 2010)) which use a Bayesian statistical approach. Given that all other statistical analyses used probabilistic statistical paradigms, the omnibus test was omitted in favour of testing the relevant ILR coordinates.

variable in a model with day of the week, observation day, day type (weekend or weekday), season and month as independent predictors after controlling for accelerometer wear time – as would have been done if a difference in the overall composition had been detected. ANOVA results are summarized in Table 5.5.

Statistically significant results were observed between weekends and weekdays, with weekdays having slightly smaller proportion of classified time spent in LPA relative to other MBs, $F(1, 536.4) = 3.96$, $p = 0.047$, $ILR_{LPA:VPA \cdot MPA \cdot SB}\Delta = -0.16$, 95% CI: -0.32 to -0.01. Figure 5.2 illustrates the daily composition of MBs on all days of the week, as well as a summary for all weekday and weekend days, relative to the geometric composition of all days after data centering to the grand geometric mean (which causes the large visual difference as VPA usually comprises a very small component of the composition). A trend was also observed for the proportion of LPA relative to other behaviour by observation day, $F(6, 538.5) = 1.95$, $p = 0.071$. Exploratory post-hoc estimated marginal means comparisons suggested that the first day of observation may report slightly lower proportions of LPA relative to other MBs compared to subsequent days ($ILR_{LPA:VPA \cdot MPA \cdot SB}\,\Delta_{Day1\text{-}2} = -0.43$, 95% CI: -0.88 to 0.03, $t(570) = -2.78$, $p_{adjusted} = 0.082$; $ILR_{LPA:VPA \cdot MPA \cdot SB}\,\Delta_{Day1\text{-}3} = -0.40$, 95% CI: -0.85 to 0.05, $t(567) = -2.61$, $p_{adjusted} = 0.126$; $ILR_{LPA:VPA \cdot MPA \cdot SB}\,\Delta_{Day1\text{-}4} = -0.46$, 95% CI: -0.92 to -0.01, $t(564) = -3.05$, $p_{adjusted} = 0.039$).

5.3.4 Mean Daily Participant MB Assessment

As random intercept mixed models accounting for multiple days of observation per participant failed to converge in subsequent models, MBs were averaged within participants per valid day. Table 5.6 repeats the compositional and arithmetic summary of Table 5.3, after averaging time in MBs per day for each participant. Figure 5.3a-b are ternary plots of the

composition of measured MBs, with (Figure 5.3a) and without (Figure 5.3b) the sleep filter applied, for each participant; percentages represent the amount of time spent in each behaviour as a proportion of total classified time. Scales range from 0% at edges to 100% at each vertex for SB, LPA and ≥MPA; VPA was illustrated using a color scale as a percentage of >MPA as an additional dimension. Thus, plot points closer to the right vertex indicate individuals where time was predominantly comprised of inactivity (estimated SB or Non-PA), any remaining time would be comprised of a combination of LPA and ≥MPA, with higher plot points indicating more ≥MPA. Comparing the plots with and without the sleep filter slightly reduced the estimate of inactivity, but participants remained in similar rank order. Overall, most individuals spent 40-80% of classified time in inactivity with few individuals having more than 20% of their time spent in ≥MPA while nearly all individuals spent at least 20% of classified time in LPA.

Regressing mean daily accelerometer wear time on the ILR coordinates representing the MB composition was not a significant improvement over the null model, $F(3,97) = 2.34$, $p = 0.08$; however, within the model an increased ratio of sedentary behaviour to all other behaviours ($ILR_{SB:LPA\cdot MPA\cdot VPA}$, $\beta = 137.0$ minutes/day, $p = 0.025$) was related to mean daily wear, such that the proportion of time spent in SB was larger in individuals with more total wear time, potentially confounding associations with other variables. Omnibus MANOVA tests found no overall difference in the composition of movement behaviours based on the number of days with sufficient wear with (Pillai's Trace = 0.02, $F(9, 288) = 0.23$, $p = 0.99$) or without (Pillai's Trace = 0.08, $F(9, 291) = 0.84$, $p = 0.58$) controlling for mean wear time. Thus, the total weekly composition of movement appears to be unaffected by additional observation days, but longer periods of wear per day tends to capture greater amounts of SB relative to other behaviours. As well, MANOVA test controlling for wear time found no significant difference in the overall

composition of movement behaviours by season (Pillai's Trace = 0.09, $F(9, 288) = 0.94$, $p = 0.49$) or month (Pillai's Trace = 0.36, $F(33, 264) = 1.11$, $p = 0.32$).

5.3.4.1 Relationships between MBs and Participant Descriptors

Regressing age on the PA composition while controlling for mean daily wear time was a significant improvement over mean daily wear time alone ($\Delta R^2 = 0.18$, $F(3, 96) = 7.35$, $p < 0.001$). Specifically, age was inversely related to the proportion of VPA relative to all other behaviours ($ILR_{VPA:MPA.LPA.SB}$), that is the proportion of observation time spent in VPA tended to be replaced with other MBs as age increased. Adding the PA composition as a predictor to a model of CPZ equivalents regressed on mean daily wear did not statistically improve over the null model ($\Delta R^2 = 0.05$, $F(3, 92) = 1.47$, $p = 0.23$).

With respect to categorical participant descriptor variables, MANOVA after controlling for daily wear found no significant influence of sex on the PA composition (Pillai's Trace = 0.05, $F(3,96) = 1.57$, $p = 0.20$). MANOVA controlling for mean wear time found no significant effects on the PA composition for employment (Pillai's Trace = 0.02, $F(3, 96) = 0.52$, $p = 0.67$), living arrangements (Pillai's Trace = 0.09, $F(9, 288) = 1.04$, $p = 0.41$), education (Pillai's Trace = 0.10, $F(6, 192) = 1.64$, $p = 0.14$), diagnosis (Pillai's Trace = 0.04, $F(3, 96) = 1.18$, $p = 0.32$), or smoking status (Pillai's Trace = 0.01, $F(3, 96) = 0.41$, $p = 0.74$) as well. MANOVA controlling for mean wear time found a significant difference among ethnic groups (Pillai's Trace = 0.18, $F(9, 288) = 2.05$, $p = 0.03$). Figure 5.4 displays relative composition of MB by ethnicity after data centering. ANOVA testing of ILR-pivot permutations only found a significant difference between ethnic groups in the ILR of SB to other behaviours ($ILR_{SB:LPA \cdot MPA \cdot VPA}$, $F(3, 96) = 3.29$, $p = 0.024$). Post-hoc estimated marginal means comparison found a significant lower proportion

of SB among individuals identifying as African Origin/Black compared to individuals identifying as White/Caucasian (Δ = -0.65, 95% CI: -1.25 to -0.04, t(96) = -2.80, $p_{adjusted}$ = 0.031). MANOVA controlling for mean wear time found a difference between marital status group (Pillai's Trace = 0.09, $F(3, 96) = 3.19$, $p = 0.03$) however, it seemed likely that marital status was at least partially a function of age, adding an additional control for age resulted in the effect no longer being significant (Pillai's Trace = 0.06, $F(3, 95) = 2.19$, $p = 0.09$).

5.3.4.2 Relationships between MBs and Health Well-being

Ethnicity, sex and age were added to mean daily wear as common covariates for health and well-being variables, even though the difference between sex groups was non-significant. There was a significant difference in the overall MB composition between those who exceeded IDF-WC thresholds and those who did not after controlling for covariates, Pillai's Trace = 0.14, $F(3, 87) = 4.59$, $p = 0.005$. ANOVA testing of ILR coordinate permutations found only the $ILR_{SB:LPA \cdot MPA \cdot VPA}$ was significant, $F(1,89) = 10.64$, $p = 0.002$. Figure 5.5 displays relative composition of MB by IDF-WC category after centering the composition. BMI was tested categorically as well as continuously. No difference was found between BMI categories, and adding the MB composition to a model predicting BMI continuously was not a significant improvement over the null model. Statistical results for compositional models of continuous health variables, as well as the effects identified for the first pivot coordinate of each permutation are described in Table 5.7. In fact, only the model assessing the symbol coding task was statistically significant compared to a model without predictors, however this final was not an improvement over the covariate only null model, $F(3,72) = 0.67$, $p = 0.58$, $\Delta R^2 = 0.019$, $\Delta AIC = 3.76$. However, the only model that improved on the null model by adding the MB composition

had the SF12-MCS as the dependent variable, $F(3, 86) = 3.14$, $p = 0.029$, $\Delta R^2 = 0.09$, $\Delta AIC = -3.98$; however the overall model fit was non-significant, $F(9,86) = 1.63$, $p = 0.1185$, $R^2 = 0.15$, AIC = 741.1. Among the primary ILR pivots only the ratio of MPA:VPA·LPA·SB was significantly associated with SF12-MCS ($\beta = 7.65$, $p = 0.004$). While absolute isotemporal substitution model is used to identify the linear effects of changing between MB categories based on energy expenditure, the estimated effects based on compositional data are available in Appendix B.

For absolute isotemporal substitution, the modeling approach inherently includes a term for observation period which stands in for the dropped MB, thus the null models only accounted for age, gender, and ethnicity of the participant. Model statistics are summarized in Table 5.8a, with the effect of specific temporal substitutions described in Table 5.8b. Results largely agreed with the compositional approach. Overall fit of models predicting BMI and symbol coding task score were significant, however neither were a statistically significant improvement of the null model, though adding MB times slightly reduced the AIC for the BMI model. Again, adding MBs to the SF12-MCS model was a statistical improvement over the null model, but AIC did not change.

A logistic model was used to estimate the effects of substituting MBs on the odds of exceeding IDF-WC recommendations, including MBs resulted in an overall good fit ($\chi^2(9,87) = 25.93$, $p = 0.002$, AIC = 102.7) , as well as a statistically significant improvement over the covariate null model ($\chi^2(4,87) = 16.73$ $p = 0.002$, $\Delta AIC = -8.7$). Models to estimate the effect of behaviour swapping suggest that there is decreased likelihood of exceeding IDF-WC guidelines when a minute of SB is replaced with MPA (OR = 0.98, 95% CI: 0.97 – 0.99, $p = 0.01$) or VPA (OR = 0.94, 95% CI: 0.89 – 1.00, $p = 0.04$). Cumulative link logistic models failed to converge

with all 4 BMI categories or by combining obese and severe obese individuals. Two dichotomizations for BMI categories were tested using binary logistic regressions: ≥Overweight and ≥Obese. The overall fit for $BMI_{\geq Overweight}$ ($\chi^2(9, 90) = 24.96$ $p = 0.003$, AIC = 92.3) was better than the model for $BMI_{\geq Obese}$ ($\chi^2(9, 90) = 16.67$ $p = 0.054$, AIC = 141.8). However, neither model was an improvement over the null ($\chi^2_{BMI\geq Overweight}$ $(4, 90) = 7.49$ $p = 0.11$, AIC = 0.51; $\chi^2_{BMI\geq Obese}$ $(4, 90) = 7.53$, $p = 0.11$, AIC = 0.47). As the better of the two models, the effects of absolute isotemporal substitution for the $BMI_{\geq Overweight}$ are summarized in Table 5.8b, there was some indication that replacing SB with VPA (OR = 0.95, 95% CI: 0.90 to 1.00, $p = 0.042$) or LPA with VPA (OR = 0.95, 95% CI: 0.91 to 1.00, $p = 0.033$) reduces the likelihood of being in the ≥Overweight category.

Despite the poor fit observed in both compositional and absolute isotemporal models, Table 5.8b still suggests some possible associations between MB substitution and health variables. For example, replacing any MB with VPA appeared to be associated with lower levels of BMI, at approximately -0.16 BMI units/minute (95% CI -0.03 to -0.28, $p = 0.007$ to 0.008) regardless of what was being replaced. There were small benefits to SF12-MCS scores when MPA replaces SB (0.06 units/min, 95% CI: 0.01 – 0.12, $p = 0.019$) or LPA (0.07 units/min, 95% CI: 0.00 – 0.13, $p = 0.049$); this agrees with the compositional model which found the $ILR_{MPA:VPA\cdot LPA\cdot SB}$ coordinate to be significantly associated with SF12-MCS scores in that model.

5.4 Discussion

The purpose of the study was to assess accelerometry movement data of a sample of individuals with schizophrenia to identify potential sources of bias due to measurement protocol adherence, evaluate the composition of daily MBs, and explore if isotemporal substitution of

waking day MBs were related to differences in health status between participants. Accelerometer wear requirements were well adhered to with 89.4% of participants meeting sufficient wear criteria, indicating high acceptability and feasibility of using accelerometry and identified few sources of potential bias attributable to protocol adherence. The high rates of protocol adherence and daily wear time may have been influenced by incentivizing sufficient wear time ($20, assessed upon device return) and allowing participants to choose whether to wear the device in bed if preferred. Other than accelerometer wear time, MB composition was largely unaffected by observation factors such as season, although weekends tended to have a higher proportion of LPA than weekdays after controlling for wear time, which may skew estimates of average daily MBs among the few individuals (n = 3) who had less than 3 weekdays of data. Unsurprisingly, better adherence can be expected among individuals with more education, lower symptom severity and fewer physical issues impeding quality of life; additional supports may be necessary to use a similar protocol in more disabled samples. Overall, results indicated that accelerometers, when available, are well adhered to when used for MB surveillance in this population.

With regards to daily MBs observed in this sample, wear time was significantly associated with an increased proportion of SB relative to other behaviours, suggesting that days with longer accelerometer wear time are likely to report more total amounts of SB in the composition and that increasing wear time on any given day is likely to skew the composition towards being comprised of more SB. After statistically controlling for the influence of wear time on composition, weekends tended to have a greater proportion of LPA than weekdays. Thus, for individuals with schizophrenia, weekends may skew accelerometry data slightly toward a greater total proportion of LPA. The first day of observation also appeared to have slightly lower proportions of time spent in LPA relative to other MBs despite adjusting for wear

time, although the initial ANOVA test only began to approach statistical significance. Interestingly, despite the data being collected in Toronto, Canada, which has large seasonal climate differences, neither month nor season had a significant effect associated with MB composition. This may be indicative that much of the PA accrued by this sample including both LPA and MPA is due to the activities of daily life, such as using active transport to attend appointments or traverse the city due to financial barriers to using public transit.

Active transit may have also contributed to the relatively high levels of PA observed in this sample compared to previous studies of individuals with severe mental illness. A large meta-analysis found people with severe mental illness engaged in an average of 476.0 min/day (95% CI: 407.3-545.4) SB and 38.4 min/day (95% CI: 32.0-44.8) (Vancampfort, Firth, et al., 2017). Similar values have been observed in meta-analyses of studies comprising people with schizophrenia: 47.1 min moderate-vigorous PA (95% CI 31.5–62.8, n = 559) (Stubbs, Firth, et al., 2016), and psychosis: 660.8 min SB (95% CI 523.2-798.4) (Stubbs, Williams, et al., 2016). Time spent in SB were similar in this sample (arithmetic mean: 479.1 (137.2) min/day/person) however MPA (89.0 (51.0) min/day/person) and VPA (7.1 (14.8)) were much higher. The differences between this sample and meta-analyses of similar populations may also be attributable to methodological considerations. These meta-analyses identify that self-report methods tend to report lower levels of SB (Stubbs, Williams, et al., 2016) and VPA, while reporting higher levels of LPA and MPA (Stubbs, Firth, et al., 2016). Validation work conducted on IPAQ based estimates of time spent in MBs suggests that self-report tools underestimate MVPA (Duncan, Arbour-Nicitopoulos, et al., 2017) but overestimates SB compared to accelerometry (Duncan et al., 2019). Pooling self-report tools with accelerometry may therefore

heavily skew data, especially when the number of self-report data points in the pool outweighs the number of accelerometer based data.

Average daily MBs differed little between participant subgroups aside from VPA decreasing with age and White/Caucasian individuals exhibiting a higher proportion of the waking day spent in SB than African Origin/Black individuals. No other demographic factors were significantly associated with MBs, other than marital status, which was no longer significant after accounting for age. It is unsurprising that VPA decreases with age, as age related decreases in activity have been well documented in the general population (Ayabe et al., 2009; Speakman & Westerterp, 2010).

After controlling for covariates, absolute and compositional isotemporal substitution models showed few associations between MBs and various health measures available, however there was some evidence of an association between MBs and indices of body weight distribution. Using a compositional approach, IDF-WC categories differed significantly with follow-up testing indicating that the proportion of SB relative to other MBs differed between groups. Using an absolute isotemporal approach agreed with this to an extent, suggesting that those who spent less time in SB and more time in either MPA or VPA, in exchange, were less likely to exceed IDF-WC criteria for metabolic syndrome. LPA was not a beneficial replacement for SB in this regard, but neither was replacing LPA with more intense PA. Some evidence for the association between MBs and BMI in this sample exists as well; modelling for BMI as a continuous variable using isotemporal substitution resulted in a significantly good overall fit, even if not an improvement on a covariate null model, with substitution effects supporting increased VPA in exchange of SB, LPA, and even MPA. Moreover, after dichotomizing BMI categories to improve model fit, individuals were less likely to be overweight if they engaged in less SB or

LPA in exchange for VPA as well. Finally, while overall models indicated a poor fit for overall mental wellbeing as measured by the SF12-MCS, there was some evidence that MPA may be associated with greater wellbeing when displacing SB or LPA.

Abdominal obesity, as a component of metabolic syndrome and as indicated by waist circumference status, may be more strongly associated with early mortality and poor metabolic outcomes than overall BMI (Paley & Johnson, 2018). The results suggest that adopting more ≥MPA, so long as it replaces SB, may be sufficient to help reduce waist circumference. Targeting overall BMI, may prove more challenging, as the current results suggest that anything less than VPA may not be beneficial, at least statistically speaking. Unfortunately, in the current sample, results showed little support for the possible benefits of LPA replacing SB with the current health measures. Rather, LPA may serve as an intermediary between transitions from SB to more beneficial MPA and VPA, but the end goal is likely to remain to accrue at least MPA instead of SB, or preferably VPA to address issues of obesity. Of course, there are additional factors important to cardiometabolic health relevant to this population such as cardiorespiratory fitness, lipid profiles, hypertension and blood sugar levels which may be positively associated with exchanging SB for lower intensities of PA. Unfortunately, as a secondary data analysis these sample descriptors were not available.

From an energy expenditure perspective any additional energy expenditure should help contribute to lower body mass and reduced obesity risk and may carry other mental and physical health benefits, LPA showed minimal associations with better health outcomes when replacing SB, and MPA had few significant effects when replacing either SB or LPA. A larger sample may be adequately powered to detect smaller effects that might emerge when exchanging SB for LPA or MPA; for example G*Power 3.1 (Erdfelder et al., 1996) estimates a fixed model would

require n = 395 to detect a small effect (Cohen's $f^2 = 0.02$) at $\alpha = 0.05$ and $1-\beta = 0.80$. Alternatively, higher intensity VPA may actually be necessary to have an impact on targeting body composition, as studies have repeatedly shown significantly higher post-exercise metabolic activity after high-intensity exercise compared to rest or low intensity exercise (Børsheim & Bahr, 2003; Gore & Withers, 1990), which may contribute to the added benefit of VPA over other forms of PA observed for body composition variables. Regardless, this study indicates that greater efforts to promote VPA specifically may be necessary to address weight related health concerns over just ≥MPA and future research should continue to evaluate VPA separately from MPA where statistically feasible. This is not to say that there is not a benefit to simply moving more. For example, there was indication that MPA may be beneficial for overall mental wellbeing, and at least one study has suggested that overall step counts rather than time spent in LPA or MVPA was associated with better cardiorespiratory fitness in individuals with schizophrenia (Engh et al., 2019).

A limitation of the present analysis is that wear time varied across participants, which poses a significant challenge for using a compositional isotemporal approach. For example, increasing LPA by 5% for someone with 600 minutes of wear time is a 30 minute increase, while for someone with 1000 minutes of wear a 5% increase equates to a 50 minute increase; simply put a 5% displacement of SB for LPA may not represent similar metabolic expenditure among all included participants. The implicit assumption is that any unrecorded time of the individual with less wear time should be similar to what was recorded, though this study has demonstrated that this assumption is likely false, as individuals who had longer wear periods had a higher proportion of SB relative to other MBs. This is somewhat addressed by having a minimum wear requirement per day as well as statistically controlling for wear time in models. Thus, in addition

to aligning with energy expenditure, the absolute isotemporal approach may more appropriately judge the effects of behavioural substitution when total wear time differs among participants. This benefit may however be moot for future studies, as wrist worn accelerometry is becoming more common for capturing 24h/day wear due to the device being less obtrusive during sleep, and subsequent algorithms have been developed for classifying PA (Van Hees et al., 2015). On the note of 24h wear protocol, a further limitation of this study is that it does not control for sleep duration or quality as these data were not collected. Sleep has shown several benefits to health and wellbeing and takes up a substantial amount of the available 24h daily time use (Pedišić et al., 2017). Assessing for both sleep quantity and quality would be ideal for examining the associations between time use behaviours and health.

A final limitation is, due to the exploratory nature of these analyses, multiple tests were performed to evaluate the association between MBs and a variety of factors of interest in this population, leading to an increased risk of falsely rejecting null hypotheses. Type 1 errors were adjusted for during post-hoc testing between groups; however, in order to avoid missing associations worthy of further exploration there was no overall adjustment for the number of tests performed. Therefore, replication studies with explicit a priori hypotheses are necessary, and should include sleep and dietary data where appropriate to the outcome. However, this study indicates that effect sizes are likely to be small, and such replication studies should use this data as a basis for ensuring adequate power.

In conclusion, this study provides an in-depth exploratory analysis of the accelerometer wear time in a sample of individuals with schizophrenia and, to our knowledge, is the first to use an isotemporal approach to evaluate the effects of average daily MB exchange on health factors

in this population. Overall, adherence to the wear protocol was good with few participant characteristics associated with lower adherence. Daily MBs were similar thorough-out the observation period, with some reactivity in the first observation day leading to lower wear and less recorded LPA. Additionally, weekends tended to have less than an hour of additional wear but had a statistically significantly larger proportion of LPA detected, compared to weekdays. Accelerometry protocol adherence in individuals with schizophrenia appears to be feasible, with little prompting, and few participant or external factors influencing wear adherence. Participant age and ethnicity had significant effects on MB composition, and after controlling for these variables as well as sex, few associations with health emerged. However, replacing SB with MPA or VPA was associated with a lower risk of exceeding waist circumference guidelines, and replacing either SB or LPA with VPA was associated with a reduced risk of being overweight. Overall, the study suggests an emerging important role for VPA as a behavioural target for addressing obesity, and thus downstream deleterious health risks, in individuals with schizophrenia.

5.5 Tables

Table 5.1 *Summary of Participant Characteristics*

	Received Accelerometer (n = 113)	Accelerometer Protocol Adherence (n = 101)
Participant Descriptors		
Male : Female	68:45	60:41
Current smokers	55	49
Age [years]	41.0 (11.7)	41.5 (11.7)
Schizophrenia : Schizoaffective	76:37	68:33
CPZ equivalents [mg]	748.1 (1216.6)	748.1 (1216.6)
Ethnicity		
African origin/Black	19	19
Asian/South Asian	16	14
Caucasian/White	67	57
Other (Including Multiple Ethnicities)	11	11
Education		
Some High School (no diploma)	21	17
High School Diploma	31	24
At least some Postsecondary	59	58
Trade School	2	2
Employment		
Full-Time	2	2
Part-Time	35	31
Student	4	3
Unemployed	70	63
Other (e.g. retired, volunteer)	2	2
Marital Status		
Single	100	88
Married	7	7
Separated/Divorced	6	6
Living arrangements		
Independent	55	51
Family/Spouse	7	6
Group (meals provided)	14	11
Group (no meals provided)	37	33
Health & Well-Being Descriptors		
BMI (kg/m^2)*		
Mean	31.5 (8.4)	31.2 (7.8)
Underweight (BMI<18.5)	1	1
Healthy Weight (18.5≤BMI<25)	19	18
Overweight (25≤BMI<30)	32	29
Obese (30≤BMI<40)	49	42

Severe Obese (40≤BMI)	11	10
Waist Circumference Category**		
Low:High	25:84	24:73
SF-12 Health Survey		
Physical Composite Score	29.6 (5.1)	29.9 (4.9)
Mental Composite Score	52.0 (11.1)	51.9 (11.1)
Symptom Severity		
BPRS-A total	34.2 (8.5)	33.5 (7.3)
BPRS-A: Affective	9.1 (1.6)	9.0 (3.3)
BPRS-A: Activation	4.4 (3.4)	4.4 (1.5)
BPRS-A: Negative	7.2 (1.6)	7.1 (3.1)
BPRS-A: Positive	7.9 (3.8)	7.6 (3.5)
BPRS-A: Resistance	5.5 (2.2)	5.4 (2.0)
CGI-S	3.5 (1.1)	3.4 (1.1)
AES	31.7 (7.9)	31.2 (7.9)
Cognitive tests	93 (82.3%)	82 (81.2%)
Symbol Coding Task	44.5 (13.6)	44.9 (13.0)
Letter Number Span Task	12.9 (3.7)	13.0 (1.7)
Z-score total	0.0 (1.8)	0.1 (1.7)

Note: * One participant opted out of being weighed, **4 opted out of waist measurement, low vs high designation based on International Diabetes Foundation sex and ethnicity-based cut points. BPRS-A = Brief Psychiatric Rating Scale 18-item Anchored version (Woerner et al., 1988), CGI-S = Clinical Global Impression Severity Scale (Guy, 1976b), AES = Apathy Evaluation Scale (Marin et al., 1991), higher scores represent more severe symptoms for all scales. CPZ = Chlorpromazine Equivalents (Gardner et al., 2010), BMI = Body Mass Index. Higher cognitive scores represent better functioning. Higher cognitive scores and SF-12 scores represent better functioning and wellbeing respectively; SF-12 composite scores are norm based with 50 representing the mean in the general population.

Table 5.2 *Frequency of Days with Sufficient Accelerometer Wear Time*

Number of Valid Weekdays	Number of Valid Weekend Days			Weekday Totals:
	0	1	2	
0	1	0	0	1
1	1	1	0	2
2	1	3	3	7
3	4	4	8	16
4	0	5	25	30
5	0	0	56	56
Weekend Totals:	7	13	92	

Note: Light and dark grey shaded cells indicate participants who meet ≥4 days with sufficient wear. Light grey shaded cells indicate the subset that would meet ≥1 weekend and ≥3 weekdays with sufficient wear.

Table 5.3 *Summary of Daily Behaviour Composition Across All Days*

	Geometric Centre	Percentiles					24h Equivalent [minutes]	Arithmetic Mean (sd) [minutes/day]
		5	25	50	75	95		
Non-Wear Time & Sleep Filter Exclusion	36.8%	19.2%	23.8%	35.6%	48.8%	57.4%	529.8	523.4 (198.6)
Sedentary Behaviour	34.5%	14.4%	25.1%	33.5%	42.5%	53.8%	496.7	488.7 (173.4)
Light PA	23.7%	11.3%	17.4%	21.9%	27.9%	36.9%	341.3	331.2 (118.8)
Moderate PA	5.0%	1.0%	3.1%	5.2%	8.3%	14.9%	72.3	89.5 (63.7)
Vigorous PA*	0.001%	0.00%	0.00%	0.00%	0.4%	3.2%	0.1	7.3 (17.5)

Note: *Imputed values for totals below 1-min detection threshold. “24h Equivalent” multiplies geometric centre composition across a 24h day. PA = Physical Activity.

Table 5.4 *Variation Matrix of Classified Movement Behaviours*

	SB	LPA	MPA	VPA*
Sedentary Behaviour (SB)	0	0.38	1.30	17.02
Light PA (LPA)	-0.38	0	0.79	16.38
Moderate PA (MPA)	-1.93	-1.55	0	12.49
Vigorous PA* (VPA)	-8.04	-7.66	-6.11	0

Note: *Imputed values for totals below 1-min detection threshold.
Values are the natural log of the ratio column:row; under the diagonal 0s is the mean, above is the variance. 0s represent perfect codependence, with mean values closer to 0 indicative of a stronger co-dependence
PA = Physical Activity

Table 5.5 *Type II Analysis of Variance Results for Differences Between Observation Factors on Movement Behaviour Compositions*

ILR Coordinate:	SB: LPA·MPA·VPA	LPA: VPA·MPA·SB	MPA: VPA·LPA·SB	VPA: MPA·LPA·SB
<u>Independent Variable</u>				
Observation Day	$F(6, 539.3) = 1.29$, $p = 0.26$	$F(6, 538.5) = 1.95$, $p = 0.071$	$F(6, 540.0) = 0.82$, $p = 0.56$	$F(6, 539.5) = 1.60$, $p = 0.14$
Day of the Week	$F(6, 530.9) = 0.78$, $p = 0.59$	$F(6, 530.5) = 1.01$, $p = 0.42$	$F(6, 531.2) = 0.42$, $p = 0.86$	$F(6, 531.0) = 0.66$, $p = 0.68$
Weekend vs Weekday	$F(1, 536.8) = 3.27$, $p = 0.071$	$F(1, 536.4) = 3.96$, $p = 0.047$	$F(1, 537.1) = 0.45$, $p = 0.50$	$F(1, 536.9) = 3.06$, $p = 0.081$
Season	$F(3, 173.5) = 0.83$, $p = 0.48$	$F(3, 184.2) = 0.37$, $p = 0.78$	$F(3, 168.2) = 0.75$, $p = 0.52$	$F(3, 172.3) = 0.51$, $p = 0.68$
Month	$F(11, 218.8) = 1.02$, $p = 0.43$	$F(11, 230.4) = 0.98$, $p = 0.47$	$F(11, 211.2) = 1.03$, $p = 0.42$	$F(11, 216.0) = 0.96$, $p = 0.48$

Table 5.6 *Summary of Average Daily Movement Behaviour Composition per Participant*

	Geometric Centre	Percentiles 5	25	50	75	95	24h Equivalent [minutes]	Arithmetic Mean (sd)** [minutes/day]
Non-Wear Time & Sleep Filter Exclusion	37.8%	24.3%	28.2%	34.2%	47.0%	52.0%	544.5	536.4 (147.9)
Sedentary Behaviour	33.4%	16.7%	26.9%	33.7%	39.8%	48.4%	481.5	479.1 (137.2)
Light PA	23.2%	13.9%	18.0%	22.4%	26.3%	33.4%	333.8	328.4 (91.8)
Moderate PA	5.5%	1.6%	3.6%	5.5%	7.9%	13.0%	79.2	89.0 (51.0)
Vigorous PA*	0.07%	0.00%	0.01%	0.08%	0.49%	2.42%	1.0	7.1 (14.8)

Note: *Imputed values for totals below 1-min detection threshold. "24h Equivalent" multiplies geometric centre composition across a 24h day. PA = Physical Activity. **Values represents means and dispersion after daily movement behaviours were averaged across participants.

Table 5.7 *Compositional Analysis Model Summary*

Dependant Variable	Permutation 1st ILR-coordinate (*Estimate, p-value*) SB: LPA·MPA·VPA		LPA: VPA·MPA·SB		MPA: VPA·LPA·SB		VPA: MPA·LPA·SB		Overall Model Fit					Comparison to Null Model				
									df	*F*	*p*	R^2	AIC	*df*	*F*	*p*	ΔR^2	ΔAIC
Body Mass Index	1.19	0.54	-1.18	0.65	0.64	0.72	-0.65	0.15	9, 87	1.19	0.31	0.16	699.2	3, 90	0.77	0.52	0.02	3.5
BPRS - Total	-0.37	0.85	0.91	0.73	-1.01	0.57	0.47	0.28	9, 91	0.71	0.70	0.07	703.1	3, 91	0.39	0.76	0.01	4.7
CGI-S	0.31	0.26	-0.46	0.21	0.05	0.83	0.09	0.13	9, 91	1.25	0.27	0.11	306.1	3, 91	1.79	0.16	0.05	0.2
Apathy Evaluation Scale	-1.9	0.35	5.3	0.05	-3.36	0.08	-0.04	0.93	9, 91	1.01	0.44	0.09	714.4	3, 91	1.63	0.19	0.05	0.7
SF-12 Health Survey PCS	1.57	0.24	-0.15	0.93	-1.34	0.27	-0.07	0.80	9, 86	0.60	0.79	0.06	591.1	3, 86	1.05	0.37	0.03	2.5
SF-12 Health Survey MCS	-0.51	0.86	-6.93	0.07	7.65	**0.004**	-0.21	0.74	9, 86	1.63	0.12	0.15	741.1	3, 86	3.14	**0.03**	0.09	-4.0
BPRS-Activation	-0.49	0.22	0.06	0.91	0.36	0.33	0.07	0.43	9, 91	0.75	0.67	0.07	384.3	3, 91	1.28	0.29	0.04	1.8
BPRS-Affective	-0.01	0.99	-0.11	0.93	0.11	0.88	0.01	0.97	9, 91	1.22	0.29	0.11	536.0	3, 91	0.01	1.00	0.00	6.0
BPRS-Resistance	-0.87	0.09	1.25	0.07	-0.48	0.32	0.1	0.39	9, 91	1.34	0.23	0.12	435.2	3, 91	1.25	0.30	0.04	1.9
BPRS-Positive	0.03	0.98	0.06	0.96	-0.34	0.69	0.26	0.22	9, 91	0.75	0.66	0.07	553.7	3, 91	0.54	0.65	0.02	4.2
BPRS-Negative	0.97	0.22	-0.37	0.72	-0.62	0.39	0.02	0.89	9, 91	1.26	0.27	0.11	521.3	3, 91	0.83	0.49	0.02	3.3
Symbol Coding Task	3.09	0.34	-4.66	0.33	2.47	0.46	-0.89	0.25	9, 72	3.86	**<0.001**	0.32	650.6	3,72	0.67	0.58	0.02	3.8
Letter Number Span Task	0.17	0.87	0.24	0.88	-0.64	0.55	0.23	0.36	9, 72	0.43	0.916	0.05	468.8	3,72	0.36	0.78	0.01	4.8
Cognitive Total Z-Score	0.33	0.49	-0.32	0.65	0	1.00	-0.01	0.94	9,72	1.22	0.296	0.13	327.0	3,72	0.17	0.92	0.01	5.4

Note : BPRS = Brief Psychiatric Rating Scale, CGI-S = Clinical Global Impression Severity Scale, PCS = Physical Composite Score, MCS = Mental Composite Score; Confidence intervals omitted from effect estimates as they do not account for substituted behaviour.

Table 5.8a *Absolute Isotemporal Substitution Results: Model Summary*

Dependant Variable	Overall Model Fit					Comparison to Null Model				
	df	*F*	*p*	R^2	AIC	*df*	*F*	*p*	ΔR^2	ΔAIC
Body Mass Index	9, 90	2.61	**0.01**	0.21	693.6	3, 90	2.09	0.09	0.07	-0.9
SF-12 Health Survey: Physical Composite Score	9, 86	1.01	0.44	0.10	587.3	3, 86	1.87	0.12	0.08	-2.6
SF-12 Health Survey: Mental Composite Score	9, 86	1.54	0.15	0.14	741.9	3, 86	2.50	**0.048**	0.10	0.00
Clinical Global Impression - Severity Scale	9, 91	1.11	0.37	0.10	307.4	3, 91	1.55	0.19	0.06	1.3
Apathy Evaluation Scale	9, 91	1.02	0.43	0.09	714.3	3, 91	1.97	0.11	0.08	-0.4
Brief Psychiatric Rating Scale (BPRS) - Total	9, 91	0.74	0.67	0.07	702.8	3, 91	0.63	0.64	0.03	5.3
BPRS-Activation	9, 91	0.86	0.56	0.08	383.3	3, 91	1.36	0.25	0.06	2.1
BPRS-Affective	9, 91	1.28	0.26	0.11	535.6	3, 91	0.48	0.75	0.02	5.9
BPRS-Resistance	9, 91	1.22	0.29	0.11	436.2	3, 91	1.18	0.32	0.05	2.9
BPRS-Positive	9, 91	0.83	0.59	0.08	553.0	3, 91	0.59	0.67	0.02	5.4
BPRS-Negative	9, 91	1.28	0.26	0.11	521.1	3, 91	0.76	0.56	0.03	4.7
Symbol Coding Task	9, 72	3.72	0.00	0.32	651.6	3, 72	0.33	0.86	0.01	6.5
Letter Number Span Task	9, 72	0.31	0.97	0.04	470.0	3, 72	0.06	0.99	0.00	7.7
Cognitive Total Z-Score	9, 72	1.21	0.30	0.13	327.1	3, 72	0.19	0.95	0.01	7.2
	df	χ^2	*p*		AIC	*df*	χ^2	*p*		ΔAIC
Waist Circumference Category	9, 87	25.93	**0.002**		102.7	4, 87	16.73	**0.002**		**-8.7**
BMI Category (≥Overweight)	9, 90	24.96	**0.003**		92.3	4, 90	7.49	0.054		0.5
BMI Category (≥Obese)	9, 90	16.67	0.054		141.8	4, 90	7.53	0.11		0.5

Table 5.8b *Absolute Isotemporal Substitution Results: Effects Summary*

Displaced:	Sedentary Behaviour (SB)			Light Physical Activity (LPA)			Moderate Physical Activity (MPA)			Vigorous Physical Activity (VPA)		
	Estimates	*95% CI*	*p*	*Estimates*	*95% CI*	*p*	*Estimates*	*95% CI*	*p*	*Estimates*	*95% CI*	*p*
Replacement						**Body Mass Index**						
SB				0.00	-0.02 – 0.02	0.71	0.00	-0.03 – 0.03	0.99	0.16	0.04 – 0.27	**0.01**
LPA	0.00	-0.02 – 0.02	0.71				0.00	-0.05 – 0.04	0.86	0.15	0.04 – 0.26	**0.01**
MPA	0.00	-0.03 – 0.03	0.99	0.00	-0.04 – 0.05	0.86				0.16	0.03 – 0.28	**0.02**
VPA	-0.16	-0.27 – -0.04	**0.01**	-0.15	-0.26 – -0.04	**0.01**	-0.16	-0.28 – -0.03	**0.02**			
						SF-12 Health Survey - Physical Composite Score						
SB				0.01	-0.01 – 0.02	0.28	0.02	-0.00 – 0.04	0.12	0.05	-0.02 – 0.13	0.18
LPA	-0.01	-0.02 – 0.01	0.28				0.01	-0.02 – 0.04	0.44	0.04	-0.03 – 0.12	0.23
MPA	-0.02	-0.04 – 0.00	0.12	-0.01	-0.04 – 0.02	0.44				0.03	-0.05 – 0.12	0.44
VPA	-0.05	-0.13 – 0.02	0.18	-0.04	-0.12 – 0.03	0.23	-0.03	-0.12 – 0.05	0.44			
						SF-12 Health Survey - Mental Composite Score						
SB				0.00	-0.03 – 0.03	0.91	-0.06	-0.12 – -0.01	**0.02**	-0.09	-0.26 – 0.08	0.31
LPA	0.00	-0.03 – 0.03	0.91				-0.07	-0.13 – -0.00	**0.05**	-0.09	-0.25 – 0.07	0.28
MPA	0.06	0.01 – 0.12	**0.02**	0.07	0.00 – 0.13	**0.049**				-0.02	-0.21 – 0.16	0.80
VPA	0.09	-0.08 – 0.26	0.31	0.09	-0.07 – 0.25	0.28	0.02	-0.16 – 0.21	0.80			
						Brief Psychiatric Rating Scale (BPRS) - Total						
SB				0.00	-0.02 – 0.02	0.79	0.01	-0.02 – 0.05	0.43	-0.04	-0.15 – 0.08	0.53
LPA	0.00	-0.02 – 0.02	0.79				0.02	-0.03 – 0.06	0.45	-0.03	-0.15 – 0.08	0.55
MPA	-0.01	-0.05 – 0.02	0.43	-0.02	-0.06 – 0.03	0.45				-0.05	-0.18 – 0.08	0.44
VPA	0.04	-0.08 – 0.15	0.53	0.03	-0.08 – 0.15	0.55	0.05	-0.08 – 0.18	0.44			
						Clinical Global Impression - Severity Scale						
SB				0.00	-0.00 – 0.00	0.22	0.00	-0.00 – 0.01	0.79	-0.01	-0.02 – 0.01	0.42
LPA	0.00	-0.00 – 0.00	0.22				0.00	-0.01 – 0.01	0.73	-0.01	-0.02 – 0.01	0.29
MPA	0.00	-0.01 – 0.00	0.79	0.00	-0.01 – 0.01	0.73				-0.01	-0.03 – 0.01	0.43
VPA	0.01	-0.01 – 0.02	0.42	0.01	-0.01 – 0.02	0.29	0.01	-0.01 – 0.03	0.43			

						Apathy Evaluation Scale						
SB				-0.02	-0.04 – 0.00	0.08	0.02	-0.02 – 0.06	0.31	0.04	-0.09 – 0.16	0.56
LPA	0.02	-0.00 – 0.04	0.08				0.04	-0.01 – 0.08	0.11	0.05	-0.06 – 0.17	0.36
MPA	-0.02	-0.06 – 0.02	0.31	-0.04	-0.08 – 0.01	0.11				0.02	-0.12 – 0.15	0.81
VPA	-0.04	-0.16 – 0.09	0.56	-0.05	-0.17 – 0.06	0.36	-0.02	-0.15 – 0.12	0.81			
						BPRS-Activation						
SB				0.00	-0.01 – 0.00	0.52	0.00	-0.01 – 0.00	0.35	-0.02	-0.05 – 0.00	0.08
LPA	0.00	-0.00 – 0.01	0.52				0.00	-0.01 – 0.01	0.64	-0.02	-0.04 – 0.00	0.08
MPA	0.00	-0.00 – 0.01	0.35	0.00	-0.01 – 0.01	0.64				-0.02	-0.04 – 0.01	0.18
VPA	0.02	-0.00 – 0.05	0.08	0.02	-0.00 – 0.04	0.08	0.02	-0.01 – 0.04	0.18			
						BPRS-Affective						
SB				0.00	-0.01 – 0.01	0.88	0.00	-0.02 – 0.01	0.90	0.01	-0.04 – 0.06	0.61
LPA	0.00	-0.01 – 0.01	0.88				0.00	-0.02 – 0.02	0.87	0.01	-0.04 – 0.06	0.61
MPA	0.00	-0.01 – 0.02	0.90	0.00	-0.02 – 0.02	0.87				0.01	-0.04 – 0.07	0.62
VPA	-0.01	-0.06 – 0.04	0.61	-0.01	-0.06 – 0.04	0.61	-0.01	-0.07 – 0.04	0.62			
						BPRS-Positive						
SB				0.00	-0.01 – 0.01	0.83	0.01	-0.01 – 0.03	0.26	-0.03	-0.09 – 0.02	0.23
LPA	0.00	-0.01 – 0.01	0.83				0.01	-0.01 – 0.03	0.31	-0.03	-0.09 – 0.02	0.23
MPA	-0.01	-0.03 – 0.01	0.26	-0.01	-0.03 – 0.01	0.31				-0.04	-0.10 – 0.02	0.17
VPA	0.03	-0.02 – 0.09	0.23	0.03	-0.02 – 0.09	0.23	0.04	-0.02 – 0.10	0.17			
						BPRS-Negative						
SB				0.00	-0.00 – 0.01	0.37	0.01	-0.01 – 0.02	0.39	0.01	-0.04 – 0.06	0.67
LPA	0.00	-0.01 – 0.00	0.37				0.00	-0.01 – 0.02	0.77	0.01	-0.04 – 0.05	0.77
MPA	-0.01	-0.02 – 0.01	0.39	0.00	-0.02 – 0.01	0.77				0.00	-0.05 – 0.06	0.88
VPA	-0.01	-0.06 – 0.04	0.67	-0.01	-0.05 – 0.04	0.77	0.00	-0.06 – 0.05	0.88			
						BPRS-Resistance						
SB				0.00	-0.01 – 0.00	0.09	0.00	-0.01 – 0.01	0.60	0.00	-0.04 – 0.03	0.78
LPA	0.00	-0.00 – 0.01	0.09				0.01	-0.00 – 0.02	0.24	0.00	-0.03 – 0.03	1.00
MPA	0.00	-0.01 – 0.01	0.60	-0.01	-0.02 – 0.00	0.24				-0.01	-0.04 – 0.03	0.69

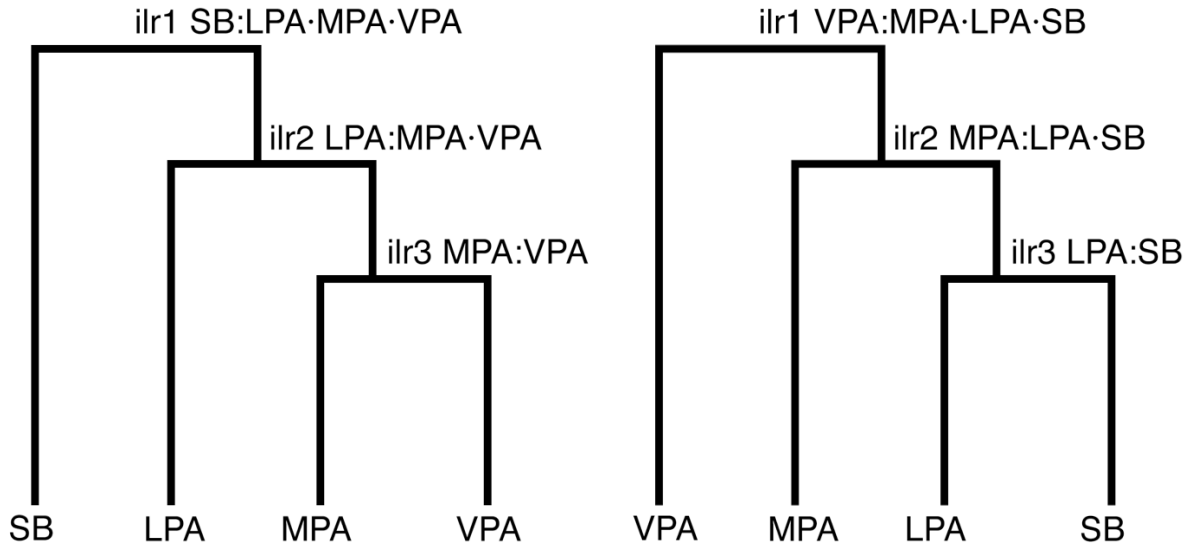
VPA	0.00	-0.03 – 0.04	0.78	0.00	-0.03 – 0.03	1.00	0.01	-0.03 – 0.04	0.69			
Symbol Coding Task												
SB				0.01	-0.02 – 0.04	0.56	0.01	-0.05 – 0.07	0.75	0.07	-0.12 – 0.26	0.47
LPA	-0.01	-0.04 – 0.02	0.56				0.00	-0.08 – 0.08	1.00	0.06	-0.12 – 0.24	0.52
MPA	-0.01	-0.07 – 0.05	0.75	0.00	-0.08 – 0.08	1.00				0.06	-0.15 – 0.27	0.58
VPA	-0.07	-0.26 – 0.12	0.47	-0.06	-0.24 – 0.12	0.52	-0.06	-0.27 – 0.15	0.58			
Letter Number Span Task												
SB				0.00	-0.01 – 0.01	0.95	0.00	-0.02 – 0.02	0.80	-0.01	-0.07 – 0.05	0.76
LPA	0.00	-0.01 – 0.01	0.95				0.00	-0.02 – 0.03	0.86	-0.01	-0.07 – 0.05	0.74
MPA	0.00	-0.02 – 0.02	0.80	0.00	-0.03 – 0.02	0.86				-0.01	-0.08 – 0.06	0.73
VPA	0.01	-0.05 – 0.07	0.76	0.01	-0.05 – 0.07	0.74	0.01	-0.06 – 0.08	0.73			
Cognitive Total Z-Score												
SB				0.00	-0.00 – 0.01	0.65	0.00	-0.01 – 0.01	0.74	0.00	-0.02 – 0.03	0.83
LPA	0.00	-0.01 – 0.00	0.65				0.00	-0.01 – 0.01	0.95	0.00	-0.02 – 0.03	0.89
MPA	0.00	-0.01 – 0.01	0.74	0.00	-0.01 – 0.01	0.95				0.00	-0.03 – 0.03	0.93
VPA	0.00	-0.03 – 0.02	0.83	0.00	-0.03 – 0.02	0.89	0.00	-0.03 – 0.03	0.93			
	Odds Ratio	*CI*	*p*	*Odds Ratio*	*CI*	*p*	*Odds Ratio*	*CI*	*p*	*Odds Ratio*	*CI*	*p*
Waist Circumference Category												
SB				1.00	1.00 – 1.01	0.26	1.02	1.01 – 1.03	**0.01**	1.06	1.00 – 1.12	**0.04**
LPA	1.00	0.99 – 1.00	0.26				1.01	1.00 – 1.03	0.07	1.06	1.00 – 1.12	0.06
MPA	0.98	0.97 – 0.99	**0.01**	0.99	0.97 – 1.00	0.07				1.04	0.98 – 1.10	0.17
VPA	0.94	0.89 – 1.00	**0.04**	0.95	0.90 – 1.00	0.06	0.96	0.91 – 1.02	0.17			
BMI Category (≥overweight)												
SB				1.00	0.99 – 1.01	0.88	1.00	0.98 – 1.01	0.72	1.05	1.00 – 1.11	**0.04**
LPA	1.00	0.99 – 1.01	0.88				1.00	0.98 – 1.02	0.73	1.05	1.00 – 1.10	**0.03**
MPA	1.00	0.99 – 1.02	0.72	1.00	0.98 – 1.02	0.73				1.06	1.00 – 1.12	0.06
VPA	0.95	0.90 – 1.00	**0.04**	0.95	0.91 – 1.00	**0.03**	0.95	0.90 – 1.00	0.06			

Note: Effect sizes are per minute of exchange. Displaced movement behaviours are listed across columns, replacement behaviours are listed by row for each dependant variable. Odds ratios listed for dichotomized variables. Odds ratios for waist circumference category above 1 indicate greater likelihood of being above International

Diabetes Federation Waist Circumference classification sex and ethnicity cut points for metabolic syndrome criteria; odds ratio for BMI category, above 1 indicates greater likelihood of having a BMI ≥ 25.

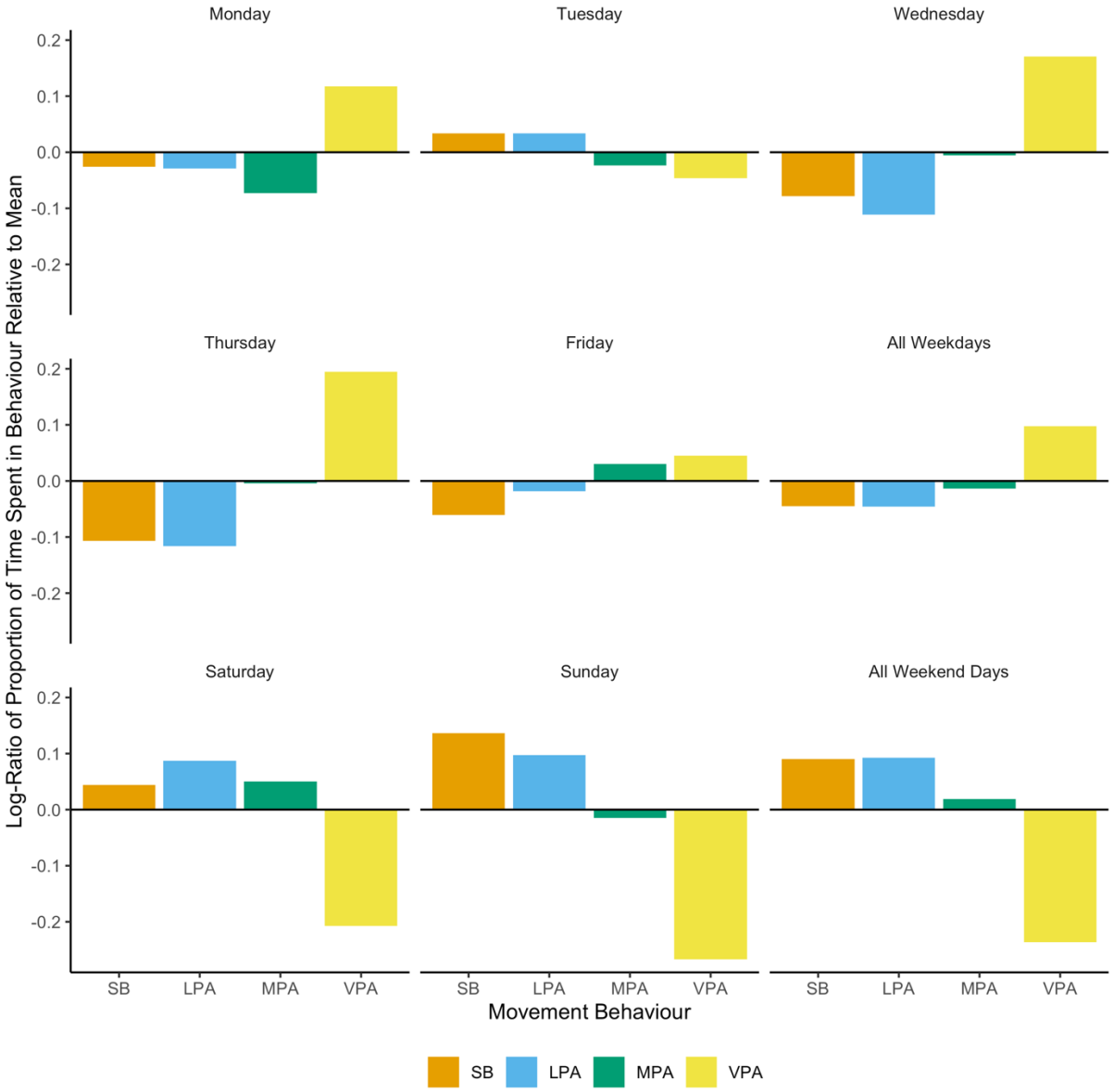
5.6 Figures

Figure 5.1a-b *Dendrogram Illustrations How Proportion of Time Spent in Movement Behaviours are Compared to Generate Two Equivalent Pivot Coordinate Sets*
a) *Set 1* **b)** *Set 2*



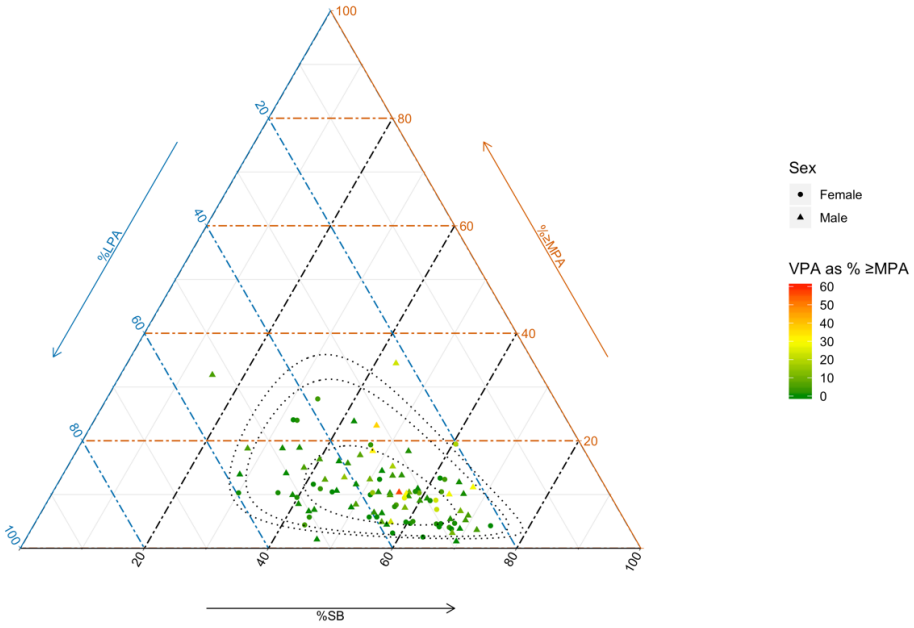
Note: The sets exemplified here are, perhaps, the least arbitrary approaches to pivot coordinate generation; Figure 5.1a compares movement behaviours (MBs) in increasing intensity while 5.1b compares MBs in decreasing intensity. SB = Sedentary Behaviour, LPA = Light Physical Activity (PA), MPA = Moderate PA, VPA = Vigorous PA, ILR = isometric log ratio pivot coordinate (with numbers indicating 1st through 3rd coordinate)

Figure 5.2 *Comparison of Centered Movement Behavior Composition by Day and Day Type*

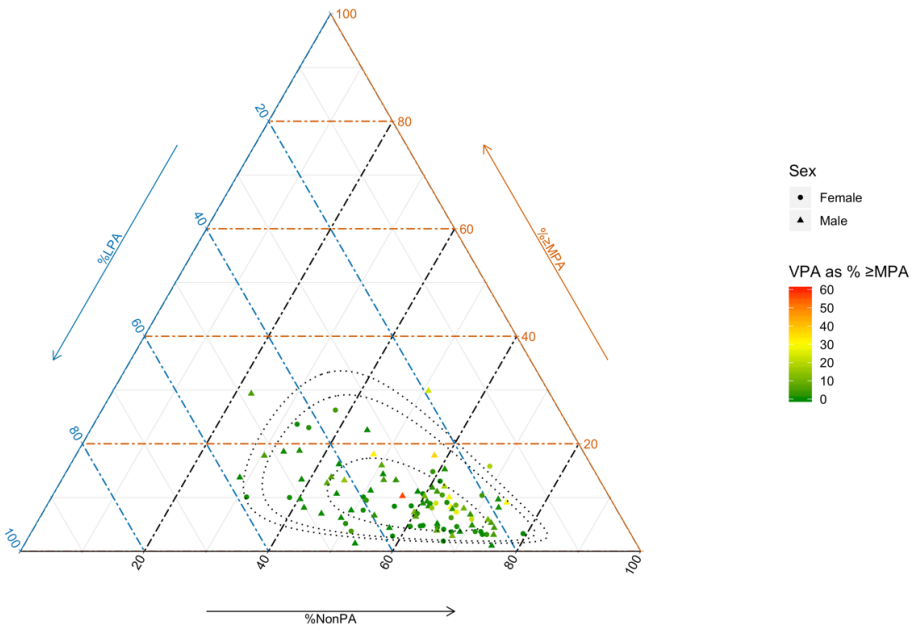


SB = Sedentary Behaviour (Orange), LPA = Light Physical Activity (PA)(Blue), MPA = Moderate PA (Green), VPA = Vigorous PA (Yellow). Data were centered such that the daily geometric mean was equivalent to equal parts for each movement behaviour (25% each).

Figure 5.3a-b *Ternary Plots of Movement Behaviour Composition With (A) and Without (B) Sleep Filter Applied to Accelerometry Data*
a) *Movement Behaviour Composition with Sleep Filter Applied*

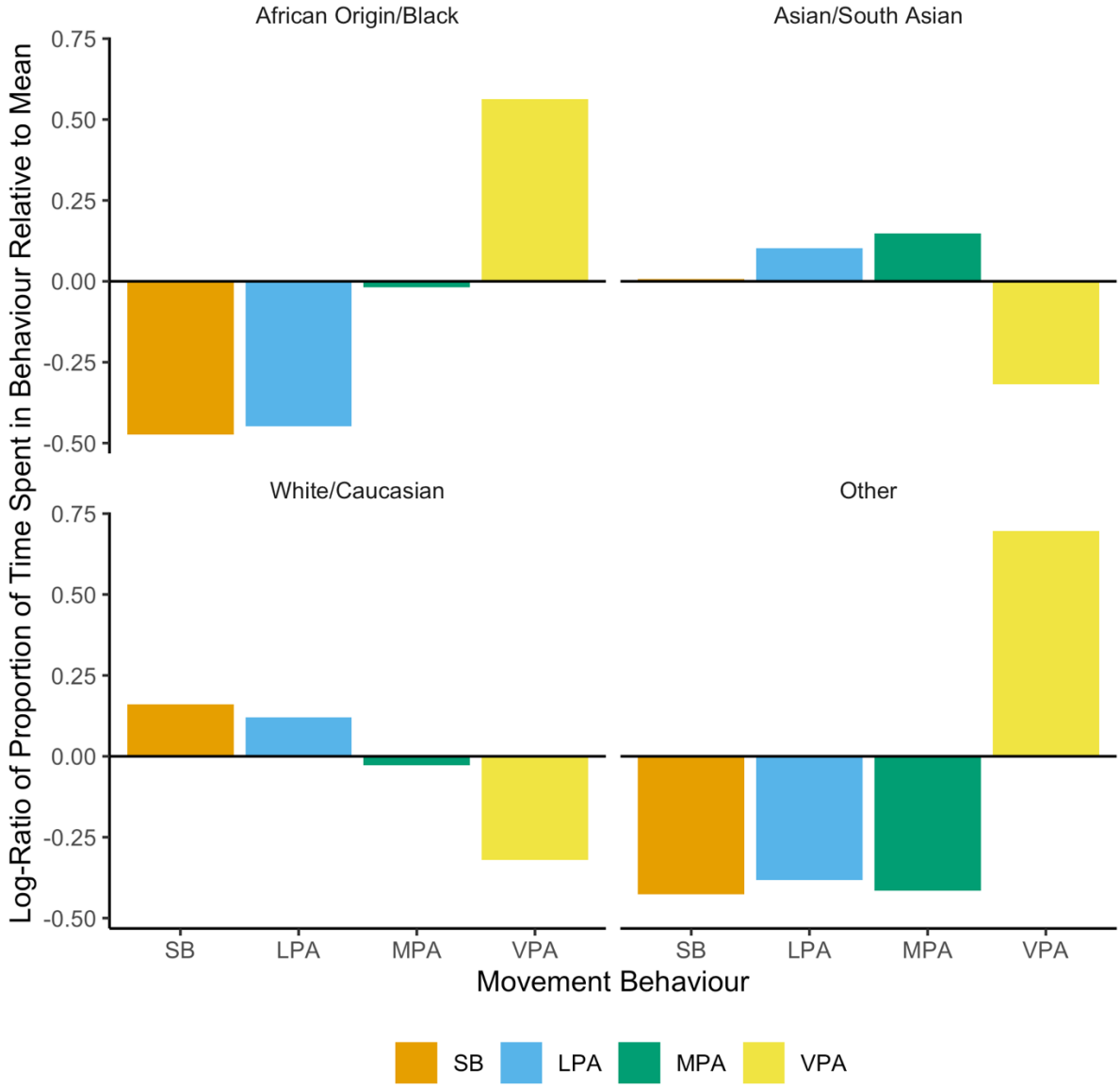


b) *Movement Behaviour Composition Without Sleep Filter Applied*



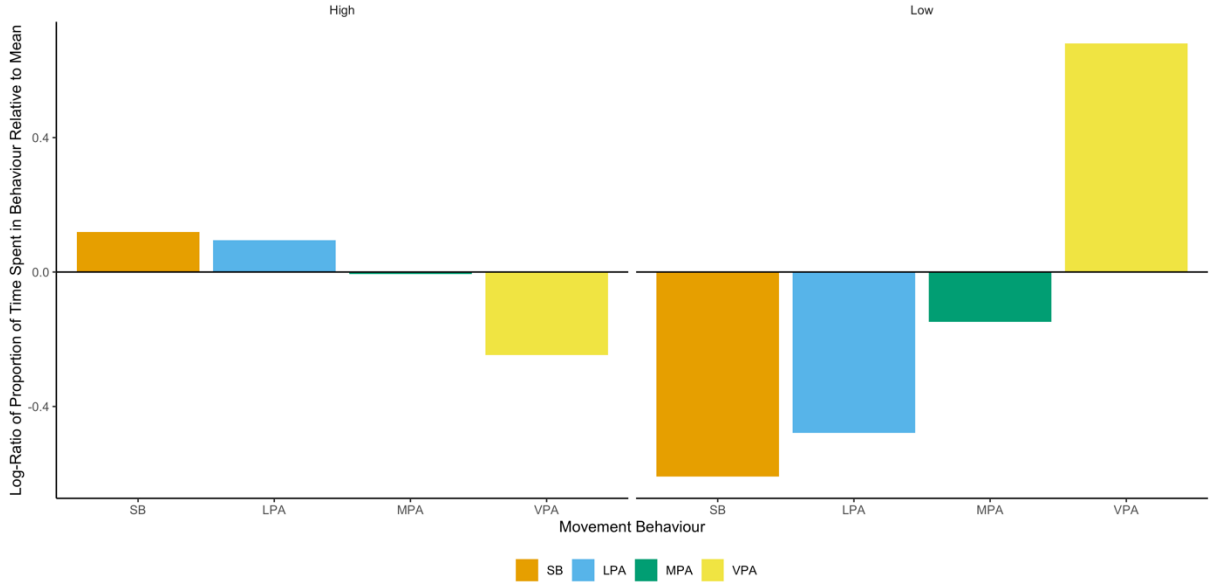
Note: LPA (Light Physical Activity [PA])(blue axes), ≥MPA (at least Moderate PA)(orange axes). The black axis on Figure 5.3a represents estimated SB (Sedentary Behaviour) after removing likely sleep from 0h00 to 5h59, and represents Non-PA without the sleep filter on Figure 5.3b (excluding sleep filter for figure 5.3a). Axis values are % of the composition of the behaviours included in the plot (i.e. LPA + ≥MPA + SB/NonPA = 100%). For 2-dimensional visualization Vigorous PA (VPA) is represented as a color scale applied to plot points as a percentage of ≥MPA; green representing lower proportions of VPA and red representing higher proportions. Dots indicate females, triangle points indicate males. Dotted lines indicate the 50, 90 and 95 confidence intervals (expanding outwards), calculated by Mahalnobis Distance and use of the Log-Ratio Transformation. Figures generated with ggtern and ggplot2 for R.

Figure 5.4 *Centered Average Daily Movement Behaviour Composition by Ethnic Group*



SB = Sedentary Behaviour (Orange), LPA = Light Physical Activity (PA)(Blue), MPA = Moderate PA (Green), VPA = Vigorous PA (Yellow). Data were centered such that the average daily geometric mean was equivalent to equal parts for each movement behaviour (25% each).

Figure 5.5 *Centered Average Daily Movement Behaviour Composition by Waist Circumference Category*



SB = Sedentary Behaviour (Orange), LPA = Light Physical Activity (PA)(Blue), MPA = Moderate PA (Green), VPA = Vigorous PA (Yellow). Data were centered such that the average daily geometric mean was equivalent to equal parts for each movement behaviour (25% each).

Chapter 6: Conclusion

Based on Sallis et al.'s (Sallis et al., 2000) behavioral epidemiological framework, the overall purpose of this dissertation therefore, is to work backwards along this framework starting by adding to evidence of validity for a common method of assessing PA in people with schizophrenia (Phase 2) and assessing whether a more holistic approach to assessing daily movement behaviours can improve the understanding of the relationships between PA and sedentary behaviour, and various health outcome (Phase 1). Given the already widespread use of the IPAQ to assess time spent in MBs in this population the first step was to build on existing evidence of validity for measures of time use derived from IPAQ-SF scores with the explicit purpose of evaluating the accuracy of minutes of MVPA (Chapter 2) and SB (Chapter 3) derived from IPAQ-SF scales relative to device-based measurements available for surveillance of daily life – namely accelerometry. After having quantified the discrepancy between measures derived from the IPAQ and accelerometry, the subsequent step sought to explore quantifiable sources of discrepancy, and whether certain participants were more prone to error than others, and using that knowledge to identify potential statistical models for correcting IPAQ-SF estimates to better align with accelerometry (Chapter 4). Finally, having demonstrated the limitations of the IPAQ for individual level surveillance, accelerometry data were analyzed to determine whether systematic bias in device wear might influence estimates of MBs, evaluate daily composition of MBs, and determine if average daily MBs were associated with demographic and health factors (Chapter 5). Table 6.1 updates the summary of studies presented in Chapter 1 (Table 1.2) with key findings, strengths and limitations for each study.

6.1 IPAQ Validation

The first two studies of this dissertation (Chapter 2 & 3) build on the seminal work by Faulkner and colleagues (2006), which assessed whether measures derived from the IPAQ-SF correlated with accelerometry derived estimates of energy expenditure in a sample of individuals with schizophrenia to evaluate its potential use for clinicians and researchers to quantify an individual's physical activity. The authors concluded that the IPAQ-SF had similar psychometric properties in that sample as when used in the general population. Since publication, the article has been cited over 150 times[4] and the IPAQ has become the most frequently used self-report tool in research for assessing time spent in MVPA and SB (Stubbs, Williams, et al., 2016) among samples of individuals with schizophrenia-like illnesses. However, while minutes of time spent in PA correlate with energy expenditure, heretofore, no studies had examined whether a minute of MVPA as scored by the IPAQ reflects a minute of MVPA in real life, despite being interpreted as such, nor has the sitting item on the IPAQ been evaluated as to how well it indicates SB.

This evaluation of whether IPAQ based estimates of time spent in MBs are valid is particularly timely. A paradigm shift is occurring in PA research. Researchers, clinicians, public health agencies and other stakeholders are interested not just in how much MVPA is accrued over a time period, but whether engaging in large amounts of uninterrupted SB can be detrimental to health even when meeting PA guidelines (Biswas et al., 2015; Warren et al., 2010). As such, analysis approaches which account for SB and varying PA intensities simultaneously have become necessary (Dumuid et al., 2019; Pedišić, 2014; Pedišić et al., 2017;

[4] As of July 9th, extracted from ScienceDirect webpage: https://www.sciencedirect.com/science/article/abs/pii/S0920996405004822

van der Ploeg & Hillsdon, 2017) and guidelines are shifting toward recommending how much of the waking day should be spent in various MBs (in addition to sleep recommendations to cover the full 24h period) (Australian Government & Department of Health, 2019; Tremblay et al., 2016; Tremblay, Chaput, et al., 2017, Ross et al., in press). In order to perform such analyses, methods to accurately measure time spent in MBs are necessary. Sallis and colleagues (2000) suggest that phases of their behavioural epidemiology framework be revisited based on feedback from later stages. It is therefore an opportune time to assess whether the IPAQ scores can be used to measure time spent in both MVPA and SB.

In addition to being timely, Messick's (American Educational Research Association et al., 2014; Messick, 1995) unified framework of construct validity would deem such additional validation work necessary. Under this framework, validation is an ongoing process drawing from multiple sources of evidence (American Educational Research Association et al., 2014; Messick, 1995). While the outcomes of the original Faulkner and colleagues (2006) study and the first two studies of this dissertation (Chapter 2 & 3) all contribute to convergent evidence of the external aspect of construct validity – that is whether the measures derived from the scale of interest is related to other theoretically related constructs (time in MVPA with METs in the former study) or alternative measurement techniques of the same construct (minutes of MVPA or SB in the latter studies) – the validation studies undertaken here align more closely with how the IPAQ-SF would be used in the context of time use analysis. As the IPAQ-SF was administered as a self-report questionnaire with only minor assistance as opposed to the interviewer-administered structure recall process used by Faulkner and colleagues (2006), these studies also provide some evidence of generalizability of IPAQ-SF derived scores when administered with limited contact in a study of individuals with schizophrenia, such as a mail-away approach to data collection

(although data completion rates would likely decrease without supervision). The fourth chapter of this dissertation also contributes to evidence of generalizability aspect of IPAQ-SF derived scores in this population – "the extent to which score properties and interpretations generalize to and across population groups" (Messick, 1995). While the original IPAQ validation studies (Craig et al., 2003) were conducted across multiple countries and in a variety of languages to assess generalizability across population groups, Chapter 4 examined why the IPAQ-SF may perform differently within the population, that is why two individuals may report different time spent in MBs on the IPAQ despite device-based measures indicating that time spent in MBs were similar. Put in the context of Messick's (1995) description of the generalizability aspect, this study examined the extent to which score properties and interpretations generalize to, and across *sub*population groups, as this may be indicative of how valid IPAQ-SF derived scores may be in other samples of individuals with schizophrenia with different demographic profiles. A review of validation studies published between 2002 and 2012 in the *Journal of Sport and Exercise Psychology* suggested that evidence of the generalizability aspect may be widely overlooked in the field of physical activity research (Chan et al., 2014). Therefore, Chapters 2-4 contribute needed evidence as to whether IPAQ-SF minute-based scores obtained from individuals with schizophrenia can be interpreted as measuring time spent in MVPA or SB, not only in terms of convergence with accelerometry data, but also in terms of the underdeveloped aspect of generalizability to administration methods and between subpopulations.

In addition to the shift towards MB measurement focusing on time use and the necessity of additional validation work that is in line with current frameworks, this update was well-timed in the context of mental health-oriented MB research. For the last five years, an effort to develop a novel MB self-report questionnaire with the needs of people with severe mental illness

specifically in mind, the Simple Physical Activity Questionnaire (SIMPAQ), has been underway (Rosenbaum & Ward, 2016), with results of a multi-country validation study recently published (Rosenbaum et al., 2020). The SIMPAQ uses an approach that attempts to account for all 24h of the day by measuring sleep, MVPA (as a combination of items measuring (1) walking and (2) sport/exercise), "incidental activity" (accrued through work or household chores), and SB. When assessed relative to an Actigraph GT3X+ results for the measures were similar to the results obtained in Chapters 2-3 for the IPAQ-SF. SIMPAQ measured MVPA correlated with accelerometry at $\rho_{spearman} = 0.25$, $p < 0.001$ in the full sample, similar to the $\rho_{spearman} = 0.30$, $p = 0.003$ observed between IPAQ-SF MVPA and accelerometry in Chapter 2. However, when the data from the subgroup of individuals with schizophrenia were analyzed, SIMPAQ MVPA Spearman correlation with accelerometry was 0.04, $p = 0.66$, n =130. SIMPAQ derived SB also correlated with accelerometry in the full sample $\rho_{spearman} = 0.19$, $p < 0.001$ and among individuals with schizophrenia $\rho_{spearman} = 0.26$, $p <0.01$, n =140. Again, this was similar to the results obtained in Chapter 3 where the IPAQ sitting item correlated with accelerometry derived $SB_{\geq 10 \text{ min bouts}}$, $\rho_{spearman} = 0.30$, $p = 0.009$. Additionally, much like my analyses of the IPAQ-SF, mean bias of Bland-Altman analyses showed small differences in mean bias between the SIMPAQ and accelerometry (MVPA = ~20min/week, SB = ~ -3h/week; estimated from figures); LoA were wide with MVPA ranging from approximately -2.5h to 3h/week and SB from -11h to 5h/week (estimated from figures), albeit smaller than the LoA observed in Chapters 2 & 3. However, the pattern of dispersion of Bland-Altman plots was similar with differences between measures increasing for MVPA in both directions as mean MVPA increased, and SB dispersion appearing somewhat homogeneous. Overall, the indices of accuracy and convergence with accelerometry for SIMPAQ derived MVPA and SB scores data were similar to those I obtained with the IPAQ-

SF, albeit with less problematically wide limits of agreement. However, the poor correlation between the SIMPAQ MVPA scores and accelerometry in the subgroup of individuals with schizophrenia is concerning.

6.2 Death to the IPAQ?

The studies described in Chapters 2 and 3 of this dissertation clearly indicate that while the IPAQ-SF MVPA and SB scores tend to correlate acceptably – or at least no worse than in the general population – with accelerometry data, the large LoA call in to question whether or not individual scores reflect genuine minutes of MVPA or SB, which is particularly important to quantify waking day MBs under the daily time use paradigm. Furthermore, regression calibration methods to correct for these wide LoAs in Chapter 4 showed small improvements, but with significant drawbacks for detecting low levels of MVPA and SB, while the novel SIMPAQ (at least across all individuals with mental illness) also appears to have less severe LoA when compared to accelerometry (Rosenbaum et al., 2020). Firth and colleagues (Firth et al., 2018) also demonstrated that IPAQ derived metabolic expenditure is not suitable for comparing PA behaviours between individuals with schizophrenia to controls without schizophrenia as IPAQ scores between groups were not significantly different despite accelerometry indicating otherwise. With mounting evidence of its flaws, is it time to stop using the IPAQ as a measure of activity in individuals with schizophrenia?

The answer depends on the research question at hand. As an indicator of mean MVPA or SB for a sample, the IPAQ seems to be adequate. It is also unclear how large a sample has to be in order for the sample mean estimate to be accurate. However, due to the wide LoA the scores from any one individual in the sample are unlikely to represent the actual minutes spent in a particular MB. This suggests that IPAQ-SF “minutes” do not represent minutes in real life, with

the exception of a sample mean – that is to say they are best thought of as a score with arbitrary units. This is further complicated by the findings of Chapter 4, which found regression-calibration approaches to IPAQ-scores to be of minimal use, and potentially obscure low levels of MVPA or SB. As a result, any index of variation, or analytic techniques that rely on spread, such as regression-based analysis (including ANOVA testing), should be interpreted cautiously. For example, if a cross-sectional regression finds a decrease in BMI of 0.1 for every 10 minutes of MVPA as scored by the IPAQ-SF, it will be difficult to determine if that observed association or effect size is indicative of any true association. Ultimately, the evidence suggests the IPAQ-SF should not be used as a measure of time spent in MVPA or SB by people with schizophrenia except to estimate group mean activity.

On a more positive note, rank order correlations between IPAQ-SF scores and accelerometry measured SB and MVPA were in the medium effect size range (both $\rho = 0.30$) indicating that IPAQ-SF scores may be adequate to indicate rank order within a sample (albeit with room for improvement), thus non-parametric approaches to analysis may be viable. However, Chapter 2 also indicates that when used to categorize individuals based on meeting Canadian Physical Activity Guidelines for MVPA, IPAQ scores were only slightly better than chance at 56% (Cohen's $\kappa = 0.12$, $p = .21$). While the ≥150 minutes per week guideline is being phased out for a 24h time use approach (Ross et al., in press) it does suggest that setting a specific minute-based cut-off to facilitate non-parametric analysis may be a flawed approach. As the research for this dissertation focuses on whether IPAQ-SF estimates of time use are accurate, alternative scoring approaches were not evaluated; however, the IPAQ does provide a method of categorical scoring into high, moderate or low activity levels based on MET expenditure and

activity frequency, which may be appropriate to broadly categorize activity levels, though validation work for this approach is necessary.

Finally, while these studies evaluated whether IPAQ derived minutes of MVPA and SB were accurate relative to accelerometry cross-sectionally it did not evaluate whether the IPAQ would be able to detect change accurately. Test-retest approaches to assess score reliability of questionnaires assumes that MBs remain consistent from one administration to the next. If this assumption was the case, there is some indication that IPAQ-SF scores remain relatively stable. Chapter 2 and 3 indicate that retest-reliability correlations for MVPA and SB scores met Cicchetti's (Cicchetti, 1994) threshold to be considered "fair" at 4-weeks apart, and this improves to "good" at 1 week apart for MVPA scores (Faulkner et al., 2006). Additionally, mean bias between administrations was low, albeit LoA were once again quite large. Chapter 4 also indicated that error in IPAQ-SF scores were associated with some participant factors, suggesting that individuals may be relatively consistent in how they respond to the IPAQ-SF. Future research to test this possibility could use accelerometry surveillance for two consecutive periods that cover both the initial IPAQ administration and the follow-up retest (as opposed to once); if the change detected by the IPAQ between test and retest is similar to the change detected by accelerometry, the IPAQ scores may be a suitable measurement tool for repeated observations to detect change over time in individuals with schizophrenia.

In summary, MVPA and SB scores from the IPAQ should not be used as a measure of MB time use for research in individuals with schizophrenia. Additional validation work would be necessary to determine whether scoring methodology are indicative of overall activity level or individual level changes. However, as physical activity guidelines communicate evidence-based

MB prescriptions in terms of time use (be it weekly or daily), these alternatives approaches have limited utility for researchers.

6.3 Assessing Movement Behaviours of the Waking Day for Individuals with Schizophrenia

In light of the conclusion that IPAQ-SF scores should not be used as a measure of time spent in MVPA and SB (Chapter 2-3), and the limited success in calibrations to correct for inaccuracy (Chapter 4), alternative methods for assessing time spent in waking day MBs in individuals with schizophrenia is necessary. Other report based measures appear to have similar levels of correlation with accelerometry data (Duncan, Arbour-Nicitopoulos, et al., 2017; Lindamer et al., 2008; Rosenbaum et al., 2020; Soundy et al., 2007) with few studies reporting accuracy compared to a more objective measurement (Duncan, Arbour-Nicitopoulos, et al., 2017; Lindamer et al., 2008; Rosenbaum et al., 2020). As a result, time use analyses such as isotemporal substitution approaches should rely on more objective approaches to movement data collection. Options are however limited. Direct calorimetry cannot be used to observe behaviour in the individual's natural environment, and indirect calorimetry units are not practical for continuous everyday surveillance as they measure ventilatory gas exchange and thus need to be worn on the face in order to estimate energy expenditure. Additionally, calorimetry and doubly labelled water methods only assess energy expenditure rather than time use. Total energy expenditure is an important component of body mass and composition related health, as unused calories absorbed by the digestive system are converted into long-term storage in the forms of glycogen and adipose tissue. However, total energy expenditure may not capture whether prolonged bouts of low energy expenditure (i.e. periods of SB) are deleterious even when total

caloric expenditure remains constant. Nutritional science has been wrestling with the question of whether every calorie ingested is the same (Mozaffarian, 2017). Similarly, time use analysis allows for MB researchers to ask whether every calorie expended is the same.

Doubly labelled water methods could be modified to sample body water more frequently to create more epochs (e.g. hourly) which would allow epochs to be classified as low or high active based on energy expenditure, however this would add considerable burden to participants and analysis expense on the researcher. As a result, direct observation or device-based measurement of research subjects are likely the most feasible methods of assessing daily time use without relying on report-based approaches. Direct observation is most frequently used in school aged children as the observer is able to monitor many several children simultaneously in classroom and recess settings, and such an approach may also be possible in inpatient settings. However, for community living adults, direct observation is invasive of privacy and person-power intensive to capture full-day activity of a sufficient sample. Device based measurements, such as accelerometry, thus appear to be the optimal solution from a practical standpoint for measuring time spent in MBs for free-living populations of individuals with schizophrenia.

While accelerometry as a tool for MB time measurement appears to balance the need for objective measurement and feasibility, device-based approaches can only represent daily behaviour if they are worn sufficiently. The study protocol for this dissertation required participants to wear their devices for 600 minutes per day on at least 4 of 7 days. Longer daily wear requirements have been suggested to better capture MBs (Herrmann et al., 2013), however at the time of data collection, 600-min per day for ≥4 days was a common requirement based on analyses of the National Health and Nutrition Examination Survey accelerometry data (Troiano et al., 2008; Trost et al., 2005). Chapter 5 assessed this fundamental component of device-based

measurement and found 89.4% of participants met this *a priori* criteria, indicating high acceptability and feasibility of using accelerometry and identified few sources of potential bias attributable to protocol adherence. Unsurprisingly, better adherence was observed among individuals with more education, lower symptom severity and fewer physical issues impeding quality of life. Participants were incentivized ($20) to meet these criteria and provided with a daily log to help remind them to wear the device while awake, but no other supports to enhance wear time were provided. Overall, results indicated that accelerometers are well suited for MB surveillance in this population with minimal intervention to enhance adherence, though more support may be necessary for those with greater needs (e.g., reminder calls to wear the device). This is an important contribution of this work as such detailed examination of accelerometer compliance has not been reported in this clinical population.

Furthermore, Chapter 5 found that differences in accelerometer wear patterns between participants had only minor impact on MB composition. Specifically, increased daily wear time was significantly associated with an increased proportion of SB relative to other behaviours and weekends tended to have a greater proportion of LPA than weekdays. As accelerometer wear protocols have shifted towards a 24h surveillance approach, the issue of differing wear time between participants will likely be less of an issue although researchers should remain aware that waking day non-wear is, generally speaking, more likely to be comprised of SB.

Having returned to Phase 2 of Sallis and colleague's (2000) behavioural epidemiological framework and identified that the IPAQ-SF is ill suited to measure time spent in MVPA and SB, but accelerometry provides a feasible and acceptable method of measuring waking day MB, the analyses of Chapter 5 returned to Phase 1 of the framework to explore the links between health and behaviour using an isotemporal approach to evaluating associations between MBs and

health. While novel in this population and exploratory in nature, the isotemporal analyses found several interesting results warranting future consideration. Foremost, results found that VPA may be necessary to reduce BMI, as exchanging time spent in any other MB with VPA was associated with lower BMI, whereas exchanging SB time with MPA or LPA was not associated with significant improvement. However, replacing SB with either MPA or VPA was associated with being below the IDF-WC cut point, suggesting that while MPA may not influence overall body weight it may help prevent adipose tissue from accumulating in locations associated with greater health risk. Additionally, there was some indication that replacing SB or LPA with MPA but not VPA was associated with improved mental wellbeing. Collectively these results show that MPA and VPA may have different health outcomes. MPA and VPA are commonly combined in MB research and PA guidelines as MVPA, but these results suggest researchers may want to analyze MPA and VPA as separate variables to determine the differing effects on health outcomes of interest. In the context of Sallis and colleague's (2000) epidemiological framework, this could inform later research phases to focus on evaluating the determinants of VPA (which may be different from MPA) and promoting MBs most relevant to key health outcomes. It may be the case that long term supports through a deliberate exercise program is necessary to have individuals with schizophrenia accrue sufficient amounts of VPA in order to manage weight, whereas encouraging regular autonomous MPA as a replacement for SB or LPA has benefits for other health domains.

For health practitioners working with individuals with schizophrenia, the increased prevalence of diabetes and CVD is an important concern. Reducing body weight reduces the risk of both, as well as the severity of insulin tolerance (Abushamat et al., 2019; Colberg et al., 2010). However, reduced risk of diabetes and improved glycemic profiles can be accrued through PA

even without weight loss (Abushamat et al., 2019; Colberg et al., 2010). One of the primary benefits of PA for diabetes control is through improved insulin action and reducing blood glucose levels (Abushamat et al., 2019; Colberg et al., 2010). Mechanisms for PA reducing CVD risk also include improving lipid profile, reducing blood pressure, and improving cardiac muscle strength (Mora et al., 2007). So while increasing weekly EE expenditure by ~2000 kcals/month, if all other influences are kept equal, *should* result in approximately .25kg/month in weight reduction[5], it may be that various CVD and diabetes related risk factors may change differently depending on whether that same amount of energy is spent in shorter bouts of VPA or longer bouts of MPA.

6.4 Strengths & Limitations

A main limitation of this dissertation is the use of accelerometry as a reference measure by which the IPAQ was compared to. Accelerometry is more objective than report-based approaches to measuring MBs in that it directly measures movement rather than relying on recall or subjective judgement of PA intensity. As discussed more objective, direct, methods of measurement than accelerometry are available to measure activity levels, but lack feasibility and usually measure energy expenditure rather than time use; when time use rather than energy expenditure is the measurement of interest, the reference measurement should reflect that (Freedson et al., 2012). Therefore, while accelerometry may not be a "gold standard" measurement method, it was the optimal compromise among available approaches to measure waking day MB time use in free-living individuals with schizophrenia.

[5] Based on 7700kcal/kg of body fat, and assuming a purely energy-in:energy-out model that does not account for changes in basal metabolic rate.

However, classifying raw accelerometer data into different MBs relies on post processing, and while device-based measures do record movement directly, many subjective decisions are made by researchers during post processing; notably the selection of methods to classify time use. In the studies presented in this dissertation, fixed cut points were used to classify time as SB, LPA, MPA or VPA based on the counts per minute recorded. These intensity thresholds have been widely adopted following initial use to analyze National Health and Nutrition Examination Survey accelerometry data (Troiano et al., 2008). The thresholds themselves were based on the weighted averages of four studies using treadmill or track walking at the target intensity to identify thresholds. These intensity thresholds assume that the mechanical efficiency of activities are similar across populations (i.e. they are equally physiologically demanding) (Strath et al., 2012). Individuals with schizophrenia tend to have gate deficits, such as slower and shorter strides, that decrease walking efficiency, and result in increased energy expenditure (Heggelund et al., 2012). As a result, intensity thresholds used for postprocessing accelerometry data based on individuals without such deficits may not be accurately classifying movement intensity in individuals with disability (Strath et al., 2012). If such movement goes undetected, the accelerometer may be underestimating the intensity of activity over a given epoch; alternatively, if excess movement due to gait inefficiencies is captured by the accelerometer, activities viewed as LPA by an outside observer may be more appropriately classified as MPA. The latter possibility may explain why average MPA reported in this sample appears relatively high, as walking accrued through daily life may be meeting criteria for MPA (and was simultaneously not captured by the IPAQ-SF MPA items). Furthermore, fixed intensity cut points classify MBs in absolute terms rather than relative to an individual's capacity (Strath et al., 2012), such that certain individuals may be approaching their

maximum capacity for PA but accelerometer post processing may still classify activity as MPA. It has been suggested that rather than continuing to refine fixed intensity thresholds mathematical modelling and pattern recognition approaches using machine learning approaches should be developed to classify activity (Freedson et al., 2012; Strath et al., 2012). Evaluation of such techniques in a population with heterogeneous physical capabilities like individuals with schizophrenia is therefore warranted.

An additional limitation is that data for these studies was collected as a supplement to an assessment of determinants of meeting Canadian Physical Activity Guidelines MVPA guidelines in individuals with schizophrenia (see Arbour-Nicitopoulos et al., 2017). As a result, data collection methods focused on capturing MVPA rather than complete daily MBs as a whole, which has emerged as a focus of interest. Participants were not instructed to keep sleep logs and at the time of analysis sleep detection algorithms were not available for waist worn accelerometers in adult populations. This decision presented two problems: 1) SB was not distinguishable from sleep and 2) sleep could not be controlled for when performing isotemporal analysis. The former issue was partially corrected for by eliminating SB from 0h00 to 5h59 as no activity above the SB threshold was detected during this period, and most likely represented sleep. However, it remains likely that time classified as SB outside of this sleep filter would be better categorized as sleep, as such SB may be overestimated in this sample. Focusing data collection on capturing MVPA also influenced administration of the IPAQ-SF as the items assessing time spent in walking were not included to prevent confusion with brisk walking which Canadian Physical Activity Guidelines and the American College of Sports Medicine (American College of Sports Medicine, 2014) identify as MPA. In doing so, there was no self-reported measure of LPA to compare against accelerometry data (although by only asking about walking

as a form of LPA, it was unlikely that scores generated by the IPAQ-SF would be similar to accelerometry). Any isotemporal analysis performed on IPAQ MB data would not have adjusted for LPA as has been recommended (Pedišić, 2014; van der Ploeg & Hillsdon, 2017). Hence, no analyses were performed to determine if IPAQ MB data had similar associations with external variables (such as health or demographic factors) as accelerometry did in Chapter 5. With the adoption of 24h guidelines in Canada imminent (Ross et al., in press), future studies will need to avoid these pitfalls in data collection methods in order to align with the prevailing paradigm.

6.5 Future directions.

Despite the limitations of the IPAQ that have been demonstrated in this dissertation, self-report based measurement will likely remain necessary to perform large scale epidemiological research. For individuals with severe mental illness, the SIMPAQ appears to reduce some of the uncertainty in MVPA and SB scores due to smaller LoA, however in individuals with schizophrenia the correlational data appears low, suggesting that scores obtained for this population may not be particularly accurate. Additionally, much like the IPAQ-SF, the SIMPAQ may not be capturing LPA well. The SIMPAQ includes items for sleep, SB, walking, exercise/sport and incidental activity (accrued through work or household chores), however, SIMPAQ scoring methodology used in the primary validation study (Rosenbaum et al., 2020) combines walking and sport/exercise item scores to estimate time spent in MVPA while the incidental activity questions were not compared against the accelerometer criterion. It is unclear if incidental activity measured by the SIMPAQ is intended to represent time spent in LPA, however the American College of Sports Medicine (American College of Sports Medicine, 2014) includes various household chores and physical labour as examples of MPA and VPA,

suggesting that equating incidental activity of daily life with LPA would not align with definitions of the LPA construct. As discussed, the measurement of LPA is essential to determine if replacing SB with LPA provides additional health benefits over increasing MVPA (Pedišić, 2014; van der Ploeg & Hillsdon, 2017). Both the IPAQ, as the most common measurement tool for MVPA and SB in individuals with schizophrenia, and the SIMPAQ which was designed to measure MVPA and SB in individuals with severe mental illness, are challenged in assessing LPA.

It is evident that self-report questionnaires to measure time spent in MBs for people with schizophrenia needs improvement to produce scores that can be considered valid estimates of time use. Foremost, tools should measure MB constructs necessary for analysis, namely SB, LPA, MPA and VPA[6]. Depending on the research question at hand, such as for mental health outcomes, differentiating MPA and VPA accrued through labour, chores, and commuting from leisure time PA may be important (Cerin et al., 2009; Jonsdottir et al., 2010; Tamminen et al., 2020; White et al., 2017). As an alternative to questionnaires, activity diaries reduce the challenges of recall (or at least reduce the length of time between activities and reporting if protocols are adhered to) and thus may improve accuracy and can be used to collect data on PA context and type when desired (Sylvia et al., 2014). There has also been recent interest in differentiating mentally active (e.g. reading or writing a dissertation) and passive (e.g. watching TV) SBs in the context of mental health outcomes, and whether substituting one for the other (or PA) is beneficial for mental health outcomes (e.g. Hallgren et al., 2020); such analyses currently require at least some form of self-report in order to classify subtypes of SB appropriately.

[6] As note elsewhere sleep should also be considered, however analyses were of course focused on the measurement of waking day MBs.

Combining diaries with mobile technologies that can provide reminder prompts based on time or changes in movement may also improve adherence and recall (Reichert et al., 2020).

Furthermore, while accuracy is of prime importance for measuring time spent in MBs, exploring other aspects of construct validity as outlined by Messick (1995) may contribute to better accuracy. In particular, trying to understand response processes through cognitive interviews may help questionnaire designers understand how individuals with schizophrenia approach responding to items. Questionnaire design tends to be a top down approach led by scientists and health professionals (e.g. SIMPAQ, IPAQ). Cognitive interviews attempt to collect qualitative data on how participants comprehend questions and response options, how participants retrieve relevant information from memory (e.g. estimation versus counting) and whether participants make an effort to tell the truth and answer the question thoughtfully (Willis, 1999). Such bottom up information may help researchers refine existing questions or develop new approaches to data collection that facilitate more accurate responses.

The data presented and conclusions drawn in this dissertation may be disheartening for health practitioners looking for a quick, inexpensive, and accurate way to track the activity levels of their patients. However, the needs of clinicians differ from researchers, the data presented suggest that the IPAQ may not be suitable for comparing scores between individuals and drawing conclusions derived from those comparisons, however the IPAQ may be acceptable for measuring change within individuals. As noted, while the IPAQ has not been directly tested for sensitivity to change within individuals with schizophrenia, the stability of IPAQ scores appears relatively good between administrations, especially when the time between administrations is short; so long as patients consistently over or underestimate their PA levels, the IPAQ may be adequate for tracking at a population level. Given that this remains untested, intervention studies

should rely on more robust approaches to MB data collection such as device-based measurement. As well, device-based measurement may be feasible for clinicians to use as an indicator of activity levels. While high quality accelerometry protocols may not be feasible to implement in a clinical setting due to device expense and person power required to disseminate and analyze data, pedometer apps available on smartphones may be viable alternative to detect movement, even if the output does not distinguish between SB and various intensities of PA, studies have repeatedly shown that step counts recorded on various smartphones tend to agree with step counts measured by research grade accelerometers (Duncan, Wunderlich, et al., 2017; Evenson et al., 2015; Major & Alford, 2016; Nolan et al., 2014). An online survey of individuals with schizophrenia found that over half of respondents had access to a smartphone, with younger individuals more likely to have one (Gay et al., 2016). For those individuals who have a smartphone or other consumer grade fitness trackers and are willing to share activity data with their health care providers, step counts from these devices may be useful for clinicians to determine if their patients have begun to move more or less as a supplement to self-report measures.

6.6 Summary

The four studies performed for this dissertation make several novel contributions to measurement of MBs in individuals with schizophrenia. Studies 1 and 2 are the first studies to quantify the agreement between the IPAQ-SF scores as a measure of time spent in MBs and accelerometry based measurements of MVPA and SB in individuals with schizophrenia, with Study 2 being the first validation study to assess sitting scores as an indicator of SB despite widely being used for this purpose. Study 3 builds on these novel results by assessing factors

associated with discrepancy between self-reported and device-based measures. While Study 4 makes several contributions by being the first study in individuals with schizophrenia to examine factors associated with accelerometer protocol compliance, assessing daily MB behaviours for influences of bias due to accelerometer wear behaviours, and performing an isotemporal analysis of MB that evaluates the effects associated with exchanging SB for varying levels of PA and *vice versa*.

In conclusion, these studies suggest that the IPAQ should not continue to be used as research tool for measuring time spent in MVPA and SB in individuals with schizophrenia. There is substantial uncertainty in the scores observed between participants, and calibrating scores based on correlates of error obscures low levels of MVPA and SB. On the other hand, accelerometry provides a feasible and well adhered to approach to MB measurement in individuals with schizophrenia with a variety of MB outcomes available through post processing. Device based measurement tools are continually decreasing in cost and increasingly ubiquitous in consumer grade electronics and thus provide an opportunity for higher quality measurement. The accelerometry data analyzed also suggests that focusing on encouraging more VPA may be necessary to address the physical health concerns of individuals with schizophrenia, though more research is necessary to replicate findings while considering pragmatic challenges of behavior change in this population.

6.7 Tables

Table 6.1 *Summary of Study Findings, Strengths and Limitations by Chapter*

	Chapter 2	Chapter 3	Chapter 4	Chapter 5
Purpose	Compare IPAQ-SF MVPA scores to accelerometer criterion for the same observation period[1,2] Assess reliability of IPAQ-SF MVPA 4-weeks apart[1,2]	Compare IPAQ-SF SB scores to accelerometer criterion for the same observation period[1,2] Assess reliability of IPAQ-SF sitting 4-weeks apart[1,2]	Identify individual factors associated with difference between IPAQ-SF and accelerometer for MVPA & SB[3] Test adjustments to IPAQ scores to better align with accelerometry derived estimates[4]	Assess factors associated with accelerometer wear adherence[5] Compare daily MB time use derived from accelerometry across 7-day observation period[5,6] Assess associations between time use exchange (e.g. replacing SB for LPA) and participant descriptors and health factors[6]
Methods	[1] Bland-Altman plots to assess mean difference & LoA [2] Correlation between measurements	[1] Bland-Altman plots to assess mean difference & LoA [2] Correlation between measurements	[3] Regression analysis of difference between measurement tools [4] 5-fold repeated resampling to test regression calibration equations against holdout sample	[5] Mixed model regression [6] Isotemporal substitution analysis

Novel Contributions	IPAQ-SF MVPA scores as a measure of time use have not been evaluated for agreement (previous data correlational)	IPAQ-SF sitting item as an indicator of time spent in SB had not been validated	Evaluates whether IPAQ-SF is more accurate for some individuals Adjustments based on these factors would improve IPAQ-SF scores as an estimate of time use	Factors associated with accelerometer compliance has not been reported in individuals with schizophrenia Assess daily MB behaviour for influences of bias Effects of MB exchange has also not been performed
Findings	IPAQ-SF underestimates MVPA by an average of -9.4 (LoA: -113.7 to 95.0) min/day Correlation between IPAQ and accelerometry ($r_{Spearman}$ = 0.25, p <0.001) similar to general population IPAQ classified meeting guidelines (≥150 min) correctly 56% of the time	IPAQ-SF performs best when ≥10-min bout length is used to define SB. IPAQ-SF overestimates SB by an average of 28.7 (LoA: -453.8 to 511.1) min/day; Similar correlation between measures similar to MVPA (ρ = 0.30, p = 0.009) Retest correlation (0.49, p < .001) similar to MVPA.	Individuals with greater BMI, above waist circumference cut points, and taking a higher medication dose tend to overestimate MVPA; being below waist circumference and higher medication dose associated with more absolute error Higher apathy and affective symptoms tend to overestimate SB; older age associated with lower absolute SB error	Accelerometer protocol adhered to by 89.4% of participants; higher education, lower levels of positive symptoms and better subjective physical health associated with better adherence. Days of the weekend had better adherence and more total wear accelerometer time. The first observation day tended to have lowest amount of wear even when meeting threshold for inclusion (≥600 min/day)

	Retest correlation (0.47, p<.001) similar to previous studies.		Regression-calibration equations applied to IPAQ scores can reduce LoAs but inflate low levels of SB and MVPA.	More daily wear time associated with greater proportion of SB. After adjusting for wear time, days of the weekend had a greater proportion of LPA than weekdays. Age related to lower proportion VPA; African Origin/Black individuals had lower SB than White/Caucasian individuals. Exchanging SB or LPA for VPA associated with lower likelihood of being overweight; replacing MPA with VPA may also contribute to lower BMI. Exchanging SB with MPA or VPA associated with lower waist circumference.
Strengths	Measures agreement/discrepancy compared to criterion accelerometry as	Measures agreement/discrepancy compared to criterion	Sources of error may be useful for improving or developing questionnaires or other	Accelerometer protocol adherence consistent across heterogeneous sample; identifies

	opposed to just monotonic correlation.	accelerometry as well as monotonic correlation. Tested varying minimum bout lengths to determine if short bouts of SB were accounted for in sitting scores.	measurement tools. Identifies potential sources of systematic bias. Regression-calibration is commonly used in nutritional sciences to improve self-report accuracy, but infrequently applied to PA measurement.	participants who may need additional supports to ensure data is representative of population. Considers the exchange of SB, LPA, MPA and VPA, not just SB and MVPA. Indicates VPA may be important for managing weight in this population.
Limitations	Accelerometry, while optimal for the methodology, still requires subjective judgment by researcher (e.g. intensity cut points), thus only agreement between measures and not "accuracy" can be determined. Accelerometers may not capture some types of MVPA due to device position (cycling, seated arm-based exercise) or if the device was removed (swimming)	No inclinometer data to classify accelerometer wear as SB to fit definition of sitting/lying down. Much IPAQ-SF data unusable due to "Don't know/Not sure" response option. No sleep log or suitable detection algorithm available; plausible sleep excluded from SB totals by simple filter (0h00-5h59). 4-week retest reliability assumes sitting time	Same limitations as Chapters 2 & 3. Participant descriptor factors measured include other self-report tools which inherently include measurement error as well. Effects of regression-calibration at population level estimated using 5-fold holdout resampling; but no explicit large external sample to test on.	No measure of sleep; plausible sleep excluded from SB totals by simple filter (0h00-5h59). Accelerometer wear time not standardized, had to be controlled for statistically Small effect sizes suggest large sample sizes needed to confirm findings related to health. Multiple comparisons increase risk of Type I error.

	4-week retest reliability assumes MVPA was not different across measured weeks.[4]	was not different across measured weeks.[4]

Note: Superscripts correspond between purpose and methods

References

Abushamat, L. A., McClatchey, P. M., Scalzo, R. L., & Reusch, J. E. B. (2019). *The Role of Exercise in Diabetes.* (K. R. Feingold, B. Anawalt, A. Boyce, G. Chrousos, K. Dungan, A. Grossman, J. M. Hershman, G. Kaltsas, C. Koch, P. Kopp, M. Korbonits, R. McLachlan, J. E. Morley, M. New, L. Perreault, J. Purnell, R. Rebar, F. Singer, D. L. Trence, … D. P. Wilson (eds.)).

Agogo, G. O., Van Der Voet, H., Van'T Veer, P., Ferrari, P., Leenders, M., Muller, D. C., Sánchez-Cantalejo, E., Bamia, C., Braaten, T., Knüppel, S., Johansson, I., Van Eeuwijk, F. A., & Boshuizen, H. (2014). Use of two-part regression calibration model to correct for measurement error in episodically consumed foods in a single-replicate study design: EPIC case study. *PLoS ONE*, *9*(11). https://doi.org/10.1371/journal.pone.0113160

Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, D. R., Tudor-Locke, C., Greer, J. L., Vezina, J., Whitt-Glover, M. C., & Leon, A. S. (2011). 2011 Compendium of Physical Activities. *Medicine & Science in Sports & Exercise*, *43*(8), 1575–1581. https://doi.org/10.1249/MSS.0b013e31821ece12

Allison, D. B., Mentore, J. L., Heo, M., Chandler, L. P., Cappelleri, J. C., Infante, M. C., & Weiden, P. J. (1999). Antipsychotic-Induced Weight Gain: A Comprehensive Research Synthesis. *American Journal of Psychiatry*, *156*(11), 1686–1696. https://doi.org/10.1176/ajp.156.11.1686

American College of Sports Medicine. (2014). ACSM Guidelines for Exercise Testing and Prescription 10th Edition. In *American College of Sports Medicine*. https://doi.org/10.1007/s13398-014-0173-7.2

American Educational Research Association, American Psychological Association, National

Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders. In *Arlington*. https://doi.org/10.1176/appi.books.9780890425596.744053

Annamalai, A., Kosir, U., & Tek, C. (2017). Prevalence of obesity and diabetes in patients with schizophrenia. *World Journal of Diabetes*, *8*(8), 390. https://doi.org/10.4239/wjd.v8.i8.390

Appelbaum, P. S., & Grisso, T. (2001). MacCAT-CR: MacArthur competence assessment tool for clinical research. In *MacCAT-CR: MacArthur Competence Assessment Tool for Clinical Research*. Professional Resource Press.

Arbour-Nicitopoulos, K. P., Duncan, M. J., Remington, G., Cairney, J., & Faulkner, G. E. (2017). The Utility of the Health Action Process Approach Model for Predicting Physical Activity Intentions and Behavior in Schizophrenia. *Frontiers in Psychiatry*, *8*(August), 1–8. https://doi.org/10.3389/fpsyt.2017.00135

Arnott, W., Sali, L., & Copland, D. (2016). Impaired reading comprehension in schizophrenia: Evidence for underlying phonological processing deficits. *Psychiatry Research*, *187*(1), 6–10. https://doi.org/10.1016/j.psychres.2010.11.025

Aschbrenner, K. A., Naslund, J. A., Gorin, A. A., Mueser, K. T., Scherer, E. A., Viron, M., Kinney, A., & Bartels, S. J. (2018). Peer support and mobile health technology targeting obesity-related cardiovascular risk in young adults with serious mental illness: Protocol for a randomized controlled trial. *Contemporary Clinical Trials*, *74*(July), 97–106. https://doi.org/10.1016/j.cct.2018.10.005

Australian Government, & Department of Health. (2019). *Australia's Physical Activity and*

Sedentary Behaviour Guidelines and the Australian 24-Hour Movement Guidelines. Education and Prevention.

Ayabe, M., Yahiro, T., Yoshioka, M., Higuchi, H., Higaki, Y., & Tanaka, H. (2009). Objectively measured age-related changes in the intensity distribution of daily physical activity in adults. *Journal of Physical Activity and Health*, *6*(4), 419–425. https://doi.org/10.1123/jpah.6.4.419

Bak, M., Fransen, A., Janssen, J., Van Os, J., & Drukker, M. (2014). Almost all antipsychotics result in weight gain: A meta-analysis. *PLoS ONE*, *9*(4), 10–12. https://doi.org/10.1371/journal.pone.0094112

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2019). *lme4: Linear Mixed-Effects Models using "Eigen" and S4*. https://cran.r-project.org/package=lme4

Bauman, A., Ainsworth, B. E., Bull, F., Craig, C. L., Hagströmer, M., Sallis, J. F., Pratt, M., & Sjöström, M. (2009). Progress and pitfalls in the use of the International Physical Activity Questionnaire (IPAQ) for adult physical activity surveillance. *Journal of Physical Activity and Health*, *6*(s1), S5–S8. https://doi.org/10.1123/jpah.6.s1.s5

Bauman, A., Bull, F., Chey, T., Craig, C. L., Ainsworth, B. E., Sallis, J. F., Bowles, H. R., Hagstromer, M., Sjostrom, M., Pratt, M., Díaz, C. G., Bazan, N., Kunic, H., Bauman, A., Merom, D., Smith, B., De Bourdeaudhuij, I., Lefevre, J., Philippaerts, R., … Hipp, D. (2009). The international prevalence study on physical activity: Results from 20 countries. *International Journal of Behavioral Nutrition and Physical Activity*, *6*, 1–11. https://doi.org/10.1186/1479-5868-6-21

Beauchamp, M. R., & McEwan, D. (2017). *Response Processes and Measurement Validity in Health Psychology BT - Understanding and Investigating Response Processes in*

Validation Research (B. D. Zumbo & A. M. Hubley (eds.); pp. 13–30). Springer International Publishing. https://doi.org/10.1007/978-3-319-56129-5_2

Beebe, L. H., & Harris, R. F. (2012). Using pedometers to document physical activity in patients with schizophrenia spectrum disorders: a feasibility study. *Journal of Psychosocial Nursing and Mental Health Services*, *50*(2), 44–49. http://myaccess.library.utoronto.ca/login?url=

Bennett, D. A., Landry, D., Little, J., & Minelli, C. (2017). Systematic review of statistical approaches to quantify, or correct for, measurement error in a continuous exposure in nutritional epidemiology. *BMC Medical Research Methodology*, *17*(1), 1–22. https://doi.org/10.1186/s12874-017-0421-6

Bentall, R. P., Baker, G. A., & Havers, S. (1991). Reality monitoring and psychotic hallucinations. *British Journal of Clinical Psychology*. https://doi.org/10.1111/j.2044-8260.1991.tb00939.x

Biddle, G. J. H., Edwardson, C. L., Henson, J., Davies, M. J., Khunti, K., Rowlands, A. V., & Yates, T. (2018). Associations of physical behaviours and behavioural reallocations with markers of metabolic health: A compositional data analysis. *International Journal of Environmental Research and Public Health*, *15*(10), 1–14. https://doi.org/10.3390/ijerph15102280

Biddle, G. J. H., Edwardson, C. L., Henson, J., Rowlands, A. V., & Yates, T. (2019). Reply to mekary, R.A.; Ding, E.L. isotemporal substitution as the gold standard model for physical activity epidemiology: Why it is the most appropriate for activity time research. Int. j. environ. res. public health 2019, 16, 797. *International Journal of Environmental Research and Public Health*, *16*(16), 6–8. https://doi.org/10.3390/ijerph16162885

Bishop, P. A. (2008). Measuring Exercise, Physical Activity, and Health. In P. A. Bishop (Ed.),

Measurement and Evaluation in Physical Activity Applications: Exercise Science, Physical Education, Coaching, Athletic Training, and Health. Holcomb Hathaway. https://books.google.ca/books?id=33L3LQAACAAJ

Biswas, A., Oh, P. I., Faulkner, G. E., Bajaj, R. R., Silver, M. A., Mitchell, M. S., & Alter, D. a. (2015). Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults a systematic review and meta-analysis. *Annals of Internal Medicine*, *162*(2), 123–132. https://doi.org/10.7326/M14-1651

Bland, J. M., & Altman, D. G. (1986). STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*, *327*(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

Bland, J. M., & Altman, D. G. (2003). Applying the right statistics: analyses of measurement studies. *Ultrasound in Obstetrics and Gynecology*, *22*(1), 85–93. https://doi.org/10.1002/uog.122

Børsheim, E., & Bahr, R. (2003). Effect of Exercise Intensity, Duration and Mode on Post-Exercise Oxygen Consumption. In *Sports Medicine*. https://doi.org/10.2165/00007256-200333140-00002

Bresee, L. C., Majumdar, S. R., Patten, S. B., & Johnson, J. A. (2010). Prevalence of cardiovascular risk factors and disease in people with schizophrenia: A population-based study. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2009.12.016

Bueno-Antequera, J., Oviedo-Caro, M. Á., & Munguía-Izquierdo, D. (2017). Relationship between objectively measured sedentary behavior and health outcomes in schizophrenia patients: The PsychiActive project. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2017.11.022

Bushe, C. J., Taylor, M., & Haukka, J. (2010). Review: Mortality in schizophrenia: a measurable clinical endpoint. *Journal of Psychopharmacology*, *24*(4_suppl), 17–25. https://doi.org/10.1177/1359786810382468

Caemmerer, J., Correll, C. U., & Maayan, L. (2012). Acute and maintenance effects of non-pharmacologic interventions for antipsychotic associated weight gain and metabolic abnormalities: a meta-analytic comparison of randomized controlled trials. *Schizophrenia Research*, *140*(1–3), 159–168. https://doi.org/10.1016/j.schres.2012.03.017

Canadian Society of Exercise Physiology. (2012). *Canadian Physical Activity and Sedentary Behaviour Guidelines*. 1–29. https://doi.org/978-1-896900-30-8

Carney, C. P., Jones, L., & Woolson, R. F. (2006). Medical comorbidity in women and men with schizophrenia: A population-based controlled study. *Journal of General Internal Medicine*, *21*(11), 1133–1137. https://doi.org/10.1111/j.1525-1497.2006.00563.x

Castle, D. J., Wessely, S., & Murray, R. M. (1993). Sex and schizophrenia: Effects of diagnostic stringency, and associations with premorbid variables. *British Journal of Psychiatry*, *162*(MAY), 658–664. https://doi.org/10.1192/bjp.162.5.658

Cather, C., Pachas, G. N., Cieslak, K. M., & Evins, A. E. (2017). Achieving Smoking Cessation in Individuals with Schizophrenia: Special Considerations. *CNS Drugs*, *31*(6), 471–481. https://doi.org/10.1007/s40263-017-0438-8

Cerin, E., Leslie, E., Sugiyama, T., & Owen, N. (2009). Associations of multiple physical activity domains with mental well-being. *Mental Health and Physical Activity*, *2*(2), 55–64. https://doi.org/10.1016/j.mhpa.2009.09.004

Chan, E. K. H., Zumbo, B. D., Zhang, W., Chen, M. Y., Darmawanti, I., & Mulyana, O. P. (2014). Validity and Validation in Social, Behavioral, and Health Sciences. In B. D. Zumbo

& E. K. H. Chan (Eds.), *Validity and Validation in Social, Behavioral, and Health Sciences* (Vol. 54). Springer International Publishing. https://doi.org/10.1007/978-3-319-07794-9

Chastin, S. F. M., Palarea-Albaladejo, J., Dontje, M. L., & Skelton, D. A. (2015). Combined Effects of Time Spent in Physical Activity, Sedentary Behaviors and Sleep on Obesity and Cardio-Metabolic Health Markers: A Novel Compositional Data Analysis Approach. *Plos One*, *10*(10), e0139984. https://doi.org/10.1371/journal.pone.0139984

Chesney, E., Goodwin, G. M., & Fazel, S. (2014). Risks of all-cause and suicide mortality in mental disorders: A meta-review. *World Psychiatry*, *13*(2), 153–160. https://doi.org/10.1002/wps.20128

Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011a). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and Science in Sports and Exercise*, *43*(2), 357–364. https://doi.org/10.1249/MSS.0b013e3181ed61a3

Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011b). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and Science in Sports and Exercise*, *43*(2), 357–364. https://doi.org/10.1249/MSS.0b013e3181ed61a3

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cohn, T., Prud'homme, D., Streiner, D., Kameh, H., & Remington, G. (2004). Characterizing coronary heart disease risk in chronic schizophrenia: high prevalence of the metabolic syndrome. *Can J Psychiatry*, *49*(11), 753–760.

Colberg, S. R., Sigal, R. J., Fernhall, B., Regensteiner, J. G., Blissmer, B. J., Rubin, R. R., Chasan-Taber, L., Albright, A. L., Braun, B., Medicine, A. C. of S., & Association, A. D.

(2010). Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement. *Diabetes Care*, *33*(12), e147–e167. https://doi.org/10.2337/dc10-9990

Cole, R. J., Kripke, D. F., Gruen, W., Mullaney, D. J., & Gillin, J. C. (1992). Automatic sleep/wake identification from wrist activity. *Sleep*, *15*(5), 461–469.

Comas, M., & Thio-Henestrosa, S. (2011). CoDaPack 2.0: a stand-alone, multi-platform compositional software. *4th International Workshop on Compositional Data Analysis*.

Coodin, S. (2001). Body mass index in persons with schizophrenia. In *Canadian Journal of Psychiatry*. https://doi.org/10.1177/070674370104600610

Correll, C. U., Manu, P., Olshanskiy, V., Napolitano, B., Kane, J. M., & Malhotra, A. K. (2009). Cardiometabolic risk of second-generation antipsychotic medications during first-time use in children and adolescents. *JAMA - Journal of the American Medical Association*. https://doi.org/10.1001/jama.2009.1549

Correll, C. U., Solmi, M., Veronese, N., Bortolato, B., Rosson, S., Santonastaso, P., Thapa-Chhetri, N., Fornaro, M., Gallicchio, D., Collantoni, E., Pigato, G., Favaro, A., Monaco, F., Kohler, C., Vancampfort, D., Ward, P. B., Gaughran, F., Carvalho, A. F., & Stubbs, B. (2017). Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry*, *16*(2), 163–180. https://doi.org/10.1002/wps.20420

Costa, R., Bastos, T., Probst, M., Seabra, A., Abreu, S., Vilhena, E., Rosenbaum, S., Ward, P. B., & Corredeira, R. (2018). Association of lifestyle-related factors and psychological factors on quality of life in people with schizophrenia. *Psychiatry Research*, *267*, 382–393.

https://doi.org/https://doi.org/10.1016/j.psychres.2018.06.022

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-Country reliability and validity. *Medicine and Science in Sports and Exercise*, *35*(8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB

Dale, D., Welk, G. J., & Matthews, C. E. (2002). Methods for Assessing Physical Activity and Challenges for Research. In G. J. Welk (Ed.), *Physical Activity Assessments for Health-related Research* (pp. 19–34). Human Kinetics. https://books.google.com/books?id=O9-vt1CZJp8C&pgis=1

Daumit, G. L., Goldberg, R. W., Anthony, C., Dickerson, F., Brown, C. H., Kreyenbuhl, J., Wohlheiter, K., & Dixon, L. B. (2005). Physical activity patterns in adults with severe mental illness. *The Journal of Nervous and Mental Disease*, *193*(10), 641–646.

Dayabandara, M., Hanwella, R., Ratnatunga, S., Seneviratne, S., Suraweera, C., & de Silva, V. A. (2017). Antipsychotic-associated weight gain: Management strategies and impact on treatment adherence. *Neuropsychiatric Disease and Treatment*, *13*, 2231–2241. https://doi.org/10.2147/NDT.S113099

Dixon, L., Weiden, P., Delahanty, J., Goldberg, R., Postrado, L., Lucksted, A., & Lehman, A. (2000). Prevalence and Correlates of Diabetes in National Schizophrenia Samples. *Schizophrenia Bulletin*, *26*(4), 903–912. https://doi.org/10.1093/oxfordjournals.schbul.a033504

Dumuid, D., Pedišić, Ž., Stanford, T. E., Martín-Fernández, J. A., Hron, K., Maher, C. A., Lewis, L. K., & Olds, T. (2019). The compositional isotemporal substitution model: A method for estimating changes in a health outcome for reallocation of time between sleep, physical

activity and sedentary behaviour. *Statistical Methods in Medical Research*, *28*(3), 846–857. https://doi.org/10.1177/0962280217737805

Dumuid, D., Stanford, T. E., Martin-Fernández, J. A., Pedišić, Ž., Maher, C. A., Lewis, L. K., Hron, K., Katzmarzyk, P. T., Chaput, J. P., Fogelholm, M., Hu, G., Lambert, E. V., Maia, J., Sarmiento, O. L., Standage, M., Barreira, T. V., Broyles, S. T., Tudor-Locke, C., Tremblay, M. S., & Olds, T. (2018). Compositional data analysis for physical activity, sedentary time and sleep research. *Statistical Methods in Medical Research*, *27*(12), 3726–3738. https://doi.org/10.1177/0962280217710835

Duncan, M. J., Arbour-Nicitopoulos, K. P., Subramaniapillai, M., Remington, G., & Faulkner, G. E. (2019). Revisiting the International Physical Activity Questionnaire (IPAQ): Assessing sitting time among individuals with schizophrenia. *Psychiatry Research*, *271*(November 2018), 311–318. https://doi.org/10.1016/j.psychres.2018.11.063

Duncan, M. J., Arbour-Nicitopoulos, K. P., Subramanieapillai, M., Remington, G., & Faulkner, G. E. (2017). Revisiting the International Physical Activity Questionnaire (IPAQ): Assessing physical activity among individuals with schizophrenia. *Schizophrenia Research*, *179*, 2–7. https://doi.org/10.1016/j.schres.2016.09.010

Duncan, M. J., Wunderlich, K., Zhao, Y., & Faulkner, G. E. (2017). Walk this way: validity evidence of iphone health application step count in laboratory and free-living conditions. *Journal of Sports Sciences*, 1–10. https://doi.org/10.1080/02640414.2017.1409855

Ekkekakis, P., Parfitt, G., & Petruzzello, S. J. (2011). The pleasure and displeasure people feel when they exercise at different intensities: decennial update and progress towards a tripartite rationale for exercise intensity prescription. *Sports Medicine*, *41*(8), 641–671. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=2

1780850

Engh, J. A., Egeland, J., Andreassen, O. A., Bang-Kittilsen, G., Bigseth, T. T., Holmen, T. L., Martinsen, E. W., Mordal, J., & Andersen, E. (2019). Objectively assessed daily steps—not light intensity physical activity, moderate-to-vigorous physical activity and sedentary time—is associated with cardiorespiratory fitness in patients with schizophrenia. *Frontiers in Psychiatry*, *10*(FEB), 1–5. https://doi.org/10.3389/fpsyt.2019.00082

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, and Computers*. https://doi.org/10.3758/BF03203630

Eskelinen, S., Sailas, E., Joutsenniemi, K., Holi, M., Koskela, T. H., & Suvisaari, J. (2017). Multiple physical healthcare needs among outpatients with schizophrenia: findings from a health examination study. *Nordic Journal of Psychiatry*, *71*(6), 448–454. https://doi.org/10.1080/08039488.2017.1319497

Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *The International Journal of Behavioral Nutrition and Physical Activity*, *12*(1), 159. https://doi.org/10.1186/s12966-015-0314-1

Faulkner, G. E., Cohn, T., & Remington, G. (2006). Validation of a physical activity assessment tool for individuals with schizophrenia. *Schizophrenia Research*, *82*(2–3), 225–231. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed7&NEWS=N&AN=2006095871

Faulkner, G. E., Cohn, T., & Remington, G. (2007). Interventions to reduce weight gain in schizophrenia. *Schizophrenia Bulletin*, *33*(3), 654–656. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed8&NEWS=N&AN=

2007428678

Faulkner, G. E., Cohn, T., Remington, G., & Irving, H. (2007). Body mass index, waist circumference and quality of life in individuals with schizophrenia. *Schizophrenia Research*, *90*(1–3), 174–178. https://doi.org/10.1016/j.schres.2006.10.009

Fervaha, G., Hill, C., Agid, O., Takeuchi, H., Foussias, G., Siddiqui, I., Kern, R. S., & Remington, G. (2014). Examination of the validity of the Brief Neurocognitive Assessment (BNA) for schizophrenia. *Schizophrenia Research*, *166*(1–3), 304–309. https://doi.org/10.1016/j.schres.2015.05.015

Firth, J., Stubbs, B., Rosenbaum, S., Vancampfort, D., Malchow, B., Schuch, F., Elliott, R., Nuechterlein, K. H., & Yung, A. R. (2016). Aerobic Exercise Improves Cognitive Functioning in People With Schizophrenia: A Systematic Review and Meta-Analysis. *Schizophrenia Bulletin*, *43*(3), 546–556. https://doi.org/10.1093/schbul/sbw115

Firth, J., Stubbs, B., Vancampfort, D., Schuch, F. B., Rosenbaum, S., Ward, P. B., Firth, J. A., Sarris, J., & Yung, A. R. (2018). The validity and value of self-reported physical activity and accelerometry in people with schizophrenia: A population-scale study of the UK Biobank. *Schizophrenia Bulletin*, *44*(6), 1293–1300. https://doi.org/10.1093/schbul/sbx149

Foldemo, A., Wärdig, R., Bachrach-Lindström, M., Edman, G., Holmberg, T., Lindström, T., Valter, L., & Osby, U. (2014). Health-related quality of life and metabolic risk in patients with psychosis. *Schizophrenia Research*, *152*(1), 295–299. https://doi.org/10.1016/j.schres.2013.11.029

Foussias, G., & Remington, G. (2010). Negative symptoms in schizophrenia: avolition and Occam's razor. *Schizophrenia Bulletin*, *36*(2), 359–369.

Fox, J., Weisberg, S., & Price, B. (2019). *car: Companion to Applied Regression*. https://cran.r-

project.org/package=car

Freedson, P., Bowles, H. R., Troiano, R., & Haskell, W. (2012). Assessment of physical activity using wearable monitors: Recommendations for monitor calibration and use in the field. *Medicine and Science in Sports and Exercise*, *44*(SUPPL. 1), 1–4. https://doi.org/10.1249/MSS.0b013e3182399b7e

Garcin, M., Mille-Hamard, L., & Billat, V. (2004). Influence of aerobic fitness level on measured and estimated perceived exertion during exhausting runs. *International Journal of Sports Medicine*, *25*(4), 270–277. https://doi.org/10.1055/s-2004-819939

Gardner, D. M., Pharm, D., Murphy, A. L., Pharm, D., Donnell, H. O., Pharm, B. S., Centorrino, F., Baldessarini, R. J., Donnell, O., & Al, E. T. (2010). *International Consensus Study of Antipsychotic Dosing*. *June*, 686–693. https://doi.org/10.1176/appi.ajp.2009.09060802

Gay, K., Torous, J., Joseph, A., Pandya, A., & Duckworth, K. (2016). Digital Technology Use Among Individuals with Schizophrenia: Results of an Online Survey. *JMIR Mental Health*, *3*(2), e15–e15. https://doi.org/10.2196/mental.5379

Gomes, E., Bastos, T., Probst, M., Ribeiro, J. C., Silva, G., & Corredeira, R. (2016). Reliability and validity of 6MWT for outpatients with schizophrenia: A preliminary study. *Psychiatry Research*, *237*, 37–42. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medc&NEWS=N&AN=26921049

Gorczynski, P., & Faulkner, G. E. (2010). Exercise therapy for schizophrenia. *The Cochrane Database of Systematic Reviews*, *5*, CD004412–CD004412. https://doi.org/10.1002/14651858.CD004412.pub2

Gore, C. J., & Withers, R. T. (1990). Effect of exercise intensity and duration on postexercise

metabolism. *Journal of Applied Physiology*. https://doi.org/10.1152/jappl.1990.68.6.2362

Guo, X., Zhang, Z., Wei, Q., Lv, H., Wu, R., & Zhao, J. (2013). The relationship between obesity and neurocognitive function in Chinese patients with schizophrenia. *BMC Psychiatry*, *13*(March 2016), 109. https://doi.org/10.1186/1471-244X-13-109

Guolo, A. (2008). Robust techniques for measurement error correction: A review. *Statistical Methods in Medical Research*, *17*(6), 555–580. https://doi.org/10.1177/0962280207081318

Gupta, N., Mathiassen, S. E., Mateu-Figueras, G., Heiden, M., Hallman, D. M., Jørgensen, M. B., & Holtermann, A. (2018). A comparison of standard and compositional data analysis in studies addressing group differences in sedentary behavior and physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, *15*(1), 1–12. https://doi.org/10.1186/s12966-018-0685-1

Gurpegui, M., Martínez-Ortega, J. M., Gutiérrez-Rojas, L., Rivero, J., Rojas, C., & Jurado, D. (2012). Overweight and obesity in patients with bipolar disorder or schizophrenia compared with a non-psychiatric sample. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*. https://doi.org/10.1016/j.pnpbp.2012.01.014

Gurusamy, J., Gandhi, S., Damodharan, D., Ganesan, V., & Palaniappan, M. (2018). Exercise, diet and educational interventions for metabolic syndrome in persons with schizophrenia: A systematic review. *Asian Journal of Psychiatry*, *36*, 73–85. https://doi.org/https://doi.org/10.1016/j.ajp.2018.06.018

Guy, W. (1976a). Clinical global impression scale. *The ECDEU Assessment Manual for Psychopharmacology-Revised. Volume DHEW Publ No ADM 76*, *338*, 218–222.

Guy, W. (1976b). Clinincal global impressions. In *In: ECDEU Assessment Manual for Psychopharmacology-Revised* (pp. 218–222). National Institute of Mental Health.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v033.i02

Hagstromer, M., Ainsworth, B. E., Oja, P., & Sjostrom, M. (2010). Comparison of a subjective and an objective measure of physical activity in a population sample. *Journal of Physical Activity & Health*, *7*(4), 541–550.

Hallgren, M., Nguyen, T.-T.-D., Owen, N., Stubbs, B., Vancampfort, D., Lundin, A., Dunstan, D., Bellocco, R., & Lagerros, Y. T. (2020). Cross-sectional and prospective relationships of passive and mentally active sedentary behaviours and physical activity with depression. *The British Journal of Psychiatry*, *217*(2), 413–419. https://doi.org/10.1192/bjp.2019.60

Harvey, P. D., Koren, D., Reichenberg, A., & Bowie, C. R. (2006). Negative Symptoms and Cognitive Deficits: What Is the Nature of Their Relationship? *Schizophrenia Bulletin*, *32*(2), 250–258. http://dx.doi.org/10.1093/schbul/sbj011

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Model Assessment and Selection. In *The elements of statistical learning,* (2nd ed.). Springer.

Hayes, R. L., & O'Grady, B. M. (2003). Do people with schizophrenia comprehend what they read? *Schizophrenia Bulletin*, *29*(3), 499–507. http://www.ncbi.nlm.nih.gov/pubmed/14609243

Healy, G. N., Clark, B. K., Winkler, E. A. H., Gardiner, P. A., Brown, W. J., & Matthews, C. E. (2011). Measurement of adults' sedentary time in population-based studies. *American Journal of Preventive Medicine*, *41*(2), 216–227. https://doi.org/10.1016/j.amepre.2011.05.005

Heggelund, J., Morken, G., Helgerud, J., Nilsberg, G., & Hoff, J. (2012). Therapeutic effects of

maximal strength training on walking efficiency in patients with schizophrenia - a pilot study. *BMC Research Notes*, *5*(1), 344. http://www.biomedcentral.com/content/5/1/344

Hennekens, C. H., Hennekens, A. R., Hollar, D., & Casey, D. E. (2005). Schizophrenia and increased risks of cardiovascular disease. In *American Heart Journal* (Vol. 150, Issue 6, pp. 1115–1121). https://doi.org/10.1016/j.ahj.2005.02.007

Herrmann, S. D., Barreira, T. V, Kang, M., & Ainsworth, B. E. (2013). How many hours are enough? Accelerometer wear time may provide bias in daily activity estimates. *Journal of Physical Activity & Health*, *10*(5), 742–749. https://doi.org/2010-0233 [pii]

Hjorth, P., Davidsen, A. S., Kilian, R., & Skrubbeltrang, C. (2014). A systematic review of controlled interventions to reduce overweight and obesity in people with schizophrenia. *Acta Psychiatrica Scandinavica*, *130*(4), 279–289. https://doi.org/10.1111/acps.12245

Hjorthøj, C., Stürup, A. E., McGrath, J. J., & Nordentoft, M. (2017). Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *The Lancet Psychiatry*, *4*(4), 295–301. https://doi.org/10.1016/S2215-0366(17)30078-0

Homel, P., Casey, D., & Allison, D. B. (2002). Changes in body mass index for individuals with and without schizophrenia, 1987-1996. *Schizophrenia Research*, *55*(3), 277–284. https://doi.org/10.1016/S0920-9964(01)00256-0

IDF. (2006). The IDF Consensus Worldwide Definition of The Metabolic Syndrome. In *Obesity and metabolism*.

IPAQ. (2005). *Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) – Short and Long Forms*. http://www.ipaq.ki.se/scoring.pdf

Jakobsen, A. S., Speyer, H., Nørgaard, H. C. B., Karlsen, M., Hjorthøj, C., Krogh, J., Mors, O., Nordentoft, M., & Toft, U. (2018). Dietary patterns and physical activity in people with

schizophrenia and increased waist circumference. *Schizophrenia Research*, *199*, 109–115. https://doi.org/https://doi.org/10.1016/j.schres.2018.03.016

Janney, C. A., Ganguli, R., Tang, G., Cauley, J. A., Holleman, R. G., Richardson, C. R., & Kriska, A. M. (2015). Physical Activity and Sedentary Behavior Measured Objectively and Subjectively in Overweight and Obese Adults With Schizophrenia or Schizoaffective Disorders. *The Journal of Clinical Psychiatry*, *76*(10), e1277–e1284. https://doi.org/10.4088/JCP.14m09330

Jin, H., Meyer, J. M., Mudaliar, S., & Jeste, D. V. (2008). Impact of atypical antipsychotic therapy on leptin, ghrelin, and adiponectin. In *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2007.11.026

Jonsdottir, I. H., Rödjer, L., Hadzibajramovic, E., Börjesson, M., & Ahlborg, G. (2010). A prospective study of leisure-time physical activity and mental health in Swedish health care workers and social insurance officers. *Preventive Medicine*, *51*(5), 373–377. https://doi.org/10.1016/j.ypmed.2010.07.019

Keefe, R. S. E., & Fenton, W. S. (2007). How should DSM-V criteria for schizophrenia include cognitive impairment? *Schizophrenia Bulletin*, *33*(4), 912–920. https://doi.org/10.1093/schbul/sbm046

Kelly, D. L., McMahon, R. P., Wehring, H. J., Liu, F., MacKowick, K. M., Boggs, D. L., Warren, K. R., Feldman, S., Shim, J. C., Love, R. C., & Dixon, L. (2011). Cigarette smoking and mortality risk in people with schizophrenia. *Schizophrenia Bulletin*. https://doi.org/10.1093/schbul/sbp152

Kelly, L. A., McMillan, D. G. E., Anderson, A., Fippinger, M., Fillerup, G., & Rider, J. (2013). Validity of actigraphs uniaxial and triaxial accelerometers for assessment of physical

activity in adults in laboratory conditions. *BMC Medical Physics*, *13*(1), 5. https://doi.org/10.1186/1756-6649-13-5

Kiang, M., Christensen, B. K., Remington, G., & Kapur, S. (2003). Apathy in schizophrenia: Clinical correlates and association with functional outcome. *Schizophrenia Research*. https://doi.org/10.1016/S0920-9964(02)00433-4

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, *28*(5). https://doi.org/10.18637/jss.v028.i05

LaPorte, R. E., Montoye, H. J., & Caspersen, C. J. (1985). Assessment of physical activity in epidemiologic research: problems and prospects. *Public Health Reports (Washington, D.C. : 1974)*, *100*(2), 131–146. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1424723&tool=pmcentrez&rendertype=abstract

Laursen, T. M., Munk-Olsen, T., & Vestergaard, M. (2012). Life expectancy and cardiovascular mortality in persons with schizophrenia. *Current Opinion in Psychiatry*, *25*(2), 83–88. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed10&NEWS=N&AN=2012089356

Laursen, T. M., Nordentoft, M., & Mortensen, P. B. (2014). Excess early mortality in schizophrenia. *Annual Review of Clinical Psychology*, *10*(November). https://doi.org/10.1146/annurev-clinpsy-032813-153657

Lenth, R. (2019). *emmeans: Estimated Marginal Means, aka Least-Squares Means*.

https://cran.r-project.org/package=emmeans

Li, R., Ma, X., Wang, G., Yang, J., & Wang, C. (2016). Why sex differences in schizophrenia? *Journal of Translational Neuroscience*, *1*(1), 37–42. http://www.ncbi.nlm.nih.gov/pubmed/29152382%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5688947

Lindamer, L. A., McKibbin, C., Norman, G. J., Jordan, L., Harrison, K., Abeyesinhe, S., & Patrick, K. (2008). Assessment of physical activity in middle-aged and older adults with schizophrenia. *Schizophrenia Research*, *104*(1–3), 294–301. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed8&NEWS=N&AN=2008408097

Major, M. J., & Alford, M. (2016). Validity of the iPhone M7 motion co-processor as a pedometer for able-bodied ambulation. *Journal of Sports Sciences*, *34*(23), 2160–2164. https://doi.org/10.1080/02640414.2016.1189086

Malaspina, D., Owen, M. J., Heckers, S., Tandon, R., Bustillo, J., Schultz, S., Barch, D. M., Gaebel, W., Gur, R. E., Tsuang, M., Van Os, J., & Carpenter, W. (2013). Schizoaffective Disorder in the DSM-5. *Schizophrenia Research*, *150*(1), 21–25. https://doi.org/https://doi.org/10.1016/j.schres.2013.04.026

Mangiafico, S. (2020). *rcompanion: Functions to Support Extension Education Program Evaluation*. https://cran.r-project.org/package=rcompanion

Manu, P., Dima, L., Shulman, M., Vancampfort, D., De Hert, M., & Correll, C. U. (2015). Weight gain and obesity in schizophrenia: Epidemiology, pathobiology, and management. *Acta Psychiatrica Scandinavica*, 97–108. https://doi.org/10.1111/acps.12445

Marin, R. S., Biedrzycki, R. C., & Firinciogullari, S. (1991). Reliability and validity of the

Apathy Evaluation Scale. *Psychiatry Research*, *38*(2), 143–162.

McCleery, A., & Nuechterlein, K. H. (2019). Cognitive impairment in psychotic illness: Prevalence, profile of impairment, developmental course, and treatment considerations. *Dialogues in Clinical Neuroscience*, *21*(3), 239–248. https://doi.org/10.31887/DCNS.2019.21.3/amccleery

McCreadie, R. G. (2003). Diet, smoking and cardiovascular risk in people with schizophrenia. *British Journal of Psychiatry*, *183*(06), 534–539. https://doi.org/10.1192/bjp.183.6.534

McCreadie, R. G., Macdonald, E., Blacklock, C., Tilak-Singh, D., Wiles, D., Halliday, J., & Paterson, J. (1998). Dietary intake of schizophrenic patients in Nithsdale, Scotland: case-control study. *BMJ*, *317*(7161), 784–785. https://doi.org/10.1136/bmj.317.7161.784

McGrath, J., Saha, S., Chant, D., & Welham, J. (2008). Schizophrenia: A Concise Overview of Incidence, Prevalence, and Mortality. *Epidemiologic Reviews*, *30*(1), 67–76. https://doi.org/10.1093/epirev/mxn001

McGregor, D. E., Chastin, S., Dall, P., & Palarea-Albaladejo, J. (2017). *Open Science Compositional Analysis Research in Physical Activity*. OpenCoDa. https://opencoda.net/

McGregor, D. E., Palarea-Albaladejo, J., Dall, P. M., Stamatakis, E., & Chastin, S. F. M. (2019). Differences in physical activity time-use composition associated with cardiometabolic risks. *Preventive Medicine Reports*, *13*(October 2018), 23–29. https://doi.org/10.1016/j.pmedr.2018.11.006

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. In *Biochemia Medica* (Vol. 22, Issue 3, pp. 276–282).

McIver, S., O'Halloran, P., & McGartland, M. (2009). Yoga as a treatment for binge eating disorder: A preliminary study. *Complementary Therapies in Medicine*, *17*(4), 196–202.

https://doi.org/10.1016/j.ctim.2009.05.002

McNamee, L., Mead, G., MacGillivray, S., & Lawrie, S. M. (2013). Schizophrenia, poor physical health and physical activity: Evidence-based interventions are required to reduce major health inequalities. In *British Journal of Psychiatry* (Vol. 203, Issue 4, pp. 239–241). https://doi.org/10.1192/bjp.bp.112.125070

Mekary, R. A., & Ding, E. L. (2019). Isotemporal substitution as the gold standard model for physical activity epidemiology: Why it is the most appropriate for activity time research. *International Journal of Environmental Research and Public Health*, *16*(5), 11–13. https://doi.org/10.3390/ijerph16050797

Mekary, R. A., Willett, W. C., Hu, F. B., & Ding, E. L. (2009). Isotemporal substitution paradigm for physical activity epidemiology and weight change. *American Journal of Epidemiology*, *170*(4), 519–527. https://doi.org/10.1093/aje/kwp163

Messick, S. (1995). Validity of Psychological Assessment. *American Psychologist*, *50*(9), 741–749. https://doi.org/10.1037//0003-066X.50.9.741

Metcalf, K. M., Baquero, B. I., Coronado Garcia, M. L., Francis, S. L., Janz, K. F., Laroche, H. H., & Sewell, D. K. (2018). Calibration of the global physical activity questionnaire to Accelerometry measured physical activity and sedentary behavior. *BMC Public Health*, *18*(1), 1–10. https://doi.org/10.1186/s12889-018-5310-3

Meyer, J. M., Davis, V. G., Goff, D. C., McEvoy, J. P., Nasrallah, H. A., Davis, S. M., Rosenheck, R. A., Daumit, G. L., Hsiao, J., Swartz, M. S., Stroup, T. S., & LIEBERMAN, J. A. (2008). Change in metabolic syndrome parameters with antipsychotic treatment in the CATIE Schizophrenia Trial: prospective data from phase 1. *Schizophrenia Research*, *101*(1–3), 273–286.

http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed8&NEWS=N&AN=2008212976

Miller, D. D. (2004). Atypical antipsychotics: sleep, sedation, and efficacy. *Primary Care Companion to the Journal of Clinical Psychiatry*, *6*(Suppl 2), 3–7. https://www.ncbi.nlm.nih.gov/pubmed/16001094

Mora, S., Cook, N., Buring, J. E., Ridker, P. M., & Lee, I.-M. (2007). Physical activity and reduced risk of cardiovascular events: potential mediating mechanisms. *Circulation*, *116*(19), 2110–2118. https://doi.org/10.1161/CIRCULATIONAHA.107.729939

Morrow, J. R. (2002). Measurement issues for the assessment of Physical activity. In G. J. Welk (Ed.), *Physical Activity Assessments for Health-related Research* (pp. 37–50). Human Kinetics. https://books.google.com/books?id=O9-vt1CZJp8C&pgis=1

Mozaffarian, D. (2017). Foods, obesity, and diabetes-are all calories created equal? *Nutrition Reviews*, *75*, 19–31. https://doi.org/10.1093/nutrit/nuw024

Naslund, J. A., Whiteman, K. L., McHugo, G. J., Aschbrenner, K. A., Marsch, L. A., & Bartels, S. J. (2017). Lifestyle interventions for weight loss among overweight and obese adults with serious mental illness: A systematic review and meta-analysis. In *General Hospital Psychiatry*. https://doi.org/10.1016/j.genhosppsych.2017.04.003

Newcomer, J. W., & Hennekens, C. H. (2007). Severe mental illness and risk of cardiovascular disease. *JAMA : The Journal of the American Medical Association*, *298*(15), 1794–1796. https://doi.org/10.1001/jama.298.15.1794

Nolan, M., Mitchell, J. R., & Doyle-Baker, P. K. (2014). Validity of the Apple iPhone®/iPod Touch® as an Accelerometer-Based Physical Activity Monitor: A Proof-of-Concept Study. *Journal of Physical Activity and Health*, *11*(4), 759–769. https://doi.org/10.1123/jpah.2011-

0336

Olfson, M., Gerhard, T., Huang, C., Crystal, S., & Stroup, T. S. (2015). Premature mortality among adults with schizophrenia in the United States. *JAMA Psychiatry*, *72*(12), 1172–1181. https://doi.org/10.1001/jamapsychiatry.2015.1737

Orozco, L., Buchleitner, A., Gimenez-Perez, G., Roqué, I., Figuls, M., Richter, B., & Mauricio, D. (2008). Exercise or exercise and diet for preventing type 2 diabetes mellitus. *Cochrane Database Syst Rev.*, *3*(3).

Owen, N., Healy, G. N., Matthews, C. E., & Dunstan, D. W. (2010). Too much sitting: the population health science of sedentary behavior. *Exercise and Sport Sciences Reviews*, *38*(3), 105–113. https://doi.org/10.1097/JES.0b013e3181e373a2

Padilla, J.-L., & Benítez, I. (2014). Validity evidence based on response processes. *Psicothema*, *26*(1), 136–144. https://doi.org/10.7334/psicothema2013.259

Palarea-Albaladejo, J., & Martin-Fernandez, J. A. (2015). zCompositions -- R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems*, *143*, 85–96. http://dx.doi.org/10.1016/j.chemolab.2015.02.019

Paley, C. A., & Johnson, M. I. (2018). Abdominal obesity and metabolic syndrome: Exercise as medicine? *BMC Sports Science, Medicine and Rehabilitation*, *10*(1), 1–8. https://doi.org/10.1186/s13102-018-0097-1

Palmer, B. W., Heaton, R. K., Paulsen, J. S., Kuck, J., Braff, D., Harris, M. J., Zisook, S., & Jeste, D. V. (1997). Is it possible to be schizophrenic yet neuropsychologically normal? *Neuropsychology*, *11*(3), 437–446. https://doi.org/10.1037//0894-4105.11.3.437

Partti, K., Vasankari, T., Kanervisto, M., Perälä, J., Saarni, S. I., Jousilahti, P., Lönnqvist, J., &

Suvisaari, J. (2015). Lung function and respiratory diseases in people with psychosis: Population-based study. *British Journal of Psychiatry*, *207*(1), 37–45. https://doi.org/DOI: 10.1192/bjp.bp.113.141937

Pedišić, Ž. (2014). Measurement issues and poor adjustments for physical activity and sleep undermine sedentary behaviour research - The focus should shift to the balance between sleep, sedentary behaviour, standing and activity. *Kinesiology*, *46*(1), 135–146.

Pedišić, Ž., Dumuid, D., & Olds, T. S. (2017). Integrating sleep, sedentary behaviour, and physical activity research in the emerging field of time-use epidemiology: Definitions, concepts, statistical methods, theoretical framework, and future directions. *Kinesiology*, *49*(2), 252–269.

Perez-Iglesias, R., Crespo-Facorro, B., Martinez-Garcia, O., Ramirez-Bonilla, M. L., Alvarez-Jimenez, M., Pelayo-Teran, J. M., Garcia-Unzueta, M. T., Amado, J. A., & Vazquez-Barquero, J. L. (2008). Weight gain induced by haloperidol, risperidone and olanzapine after 1 year: Findings of a randomized clinical trial in a drug-naïve population. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2007.10.022

Peterson, N. E., Sirard, J. R., Kulbok, P. A., Deboer, M. D., & Erickson, J. M. (2015). Validation of accelerometer thresholds and Inclinometry for measurement of sedentary behavior in young adult university students. *Research in Nursing and Health*, *38*(6), 492–499. https://doi.org/10.1002/nur.21694

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Connor Gorber, S., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *5*. https://doi.org/10.1186/1479-5868-5-56

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Reichenberg, A., Harvey, P. D., Bowie, C. R., Mojtabai, R., Rabinowitz, J., Heaton, R. K., & Bromet, E. (2009). Neuropsychological function and dysfunction in schizophrenia and psychotic affective disorders. *Schizophrenia Bulletin*, *35*(5), 1022–1029. https://doi.org/10.1093/schbul/sbn044

Reichenberg, A., Weiser, M., Caspi, A., Knobler, H. Y., Lubin, G., Harvey, P. D., Rabinowitz, J., & Davidson, M. (2006). Premorbid intellectual functioning and risk of schizophrenia and spectrum disorders. *Journal of Clinical and Experimental Neuropsychology*, *28*(2), 193–207. https://doi.org/10.1080/13803390500360372

Reichert, M., Giurgiu, M., Koch, E. D., Wieland, L. M., Lautenbach, S., Neubauer, A. B., von Haaren-Mack, B., Schilling, R., Timm, I., Notthoff, N., Marzi, I., Hill, H., Brüßler, S., Eckert, T., Fiedler, J., Burchartz, A., Anedda, B., Wunsch, K., Gerber, M., … Liao, Y. (2020). Ambulatory assessment for physical activity research: State of the science, best practices and future directions. *Psychology of Sport and Exercise*, *50*(June), 101742. https://doi.org/10.1016/j.psychsport.2020.101742

Revelle, W. R. (2017). *psych: Procedures for personality and psychological research*.

Reynolds, G. P., & Kirk, S. L. (2010). Metabolic side effects of antipsychotic drug treatment - pharmacological mechanisms. In *Pharmacology and Therapeutics*. https://doi.org/10.1016/j.pharmthera.2009.10.010

Rosenbaum, S., Morell, R., Abdel-Baki, A., Ahmadpanah, M., Anilkumar, T. V., Baie, L., Bauman, A., Bender, S., Boyan Han, J., Brand, S., Bratland-Sanda, S., Bueno-Antequera, J., Camaz Deslandes, A., Carneiro, L., Carraro, A., Castañeda, C. P., Castro Monteiro, F.,

Chapman, J., Chau, J. Y., … Ward, P. B. (2020). Assessing physical activity in people with mental illness: 23-country reliability and validity of the simple physical activity questionnaire (SIMPAQ). *BMC Psychiatry*, *20*(1), 1–12. https://doi.org/10.1186/s12888-020-2473-0

Rosenbaum, S., & Ward, P. B. (2016). The Simple Physical Activity Questionnaire. *The Lancet Psychiatry*, *3*(1), e1. https://doi.org/10.1016/S2215-0366(15)00496-4

Ross, R., Chaput, J.-P., Giangregorio, L. M., Janssen, I., Saunders, T. J., Kho, M. E., Poitras, V. J., Tomasone, J. R., El-Kotob, R., McLaughlin, E. C., Duggan, M., Carrier, J., Carson, V., Chastin, S. F., Latimer-Cheung, A. E., Chulak-Bozzer, T., Faulkner, G., Flood, S. M., Gazendam, M. K., … Tremblay, M. S. (2020). Canadian 24-Hour Movement Guidelines for Adults aged 18–64 years and Adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology, Nutrition, and Metabolism*, *45*(10 (Suppl. 2)), S57–S102. https://doi.org/10.1139/apnm-2020-0467

Ryan, M. C. M., Collins, P., & Thakore, J. H. (2003). Impaired Fasting Glucose Tolerance in First-Episode, Drug-Naive Patients With Schizophrenia. *American Journal of Psychiatry*, *160*(2), 284–289. https://doi.org/10.1176/appi.ajp.160.2.284

Sadeh, A., Sharkey, K. M., & Carskadon, M. A. (1994). Activity-based sleep-wake identification: an empirical test of methodological issues. *Sleep*, *17*(3), 201–207.

Sagud, M., Vuksan-Cusa, B., Jakšic, N., Mihaljevic-Peleš, A., Kuzman, M. R., & Pivac, N. (2018). Smoking in schizophrenia: An updated review. *Psychiatria Danubina*, *30*, S216–S223.

Saha, S., Chant, D., Welham, J., & McGrath, J. (2005). A systematic review of the prevalence of schizophrenia. *PLoS Med*, *2*(5), e141. https://doi.org/05-PLME-RA-0069R1

[pii]\r10.1371/journal.pmed.0020141

Saint-Maurice, P. F., Welk, G. J., Beyler, N. K., Bartee, R. T., & Heelan, K. A. (2014). Calibration of self-report tools for physical activity research: The Physical Activity Questionnaire (PAQ). *BMC Public Health*, *14*(1), 1–9. https://doi.org/10.1186/1471-2458-14-461

Sallis, J. F., Owen, N., & Fotheringham, M. J. (2000). Behavioral epidemiology: A systematic framework to classify phases of research on health promotion and disease prevention. *Annals of Behavioral Medicine*, *22*(4), 294–298. https://doi.org/10.1007/BF02895665

Sallis, J. F., & Saelens, B. E. (2000). Assessment of Physical Activity by Self-Report: Status, Limitations, and Future Directions. *Research Quarterly for Exercise and Sport*, *71*(sup2), 1–14. https://doi.org/10.1080/02701367.2000.11082780

Schilling, R., Schärli, E., Fischer, X., Donath, L., Faude, O., Brand, S., Pühse, U., Zahner, L., Rosenbaum, S., Ward, P. B., Carraro, A., & Gerber, M. (2018). The utility of two interview-based physical activity questionnaires in healthy young adults: Comparison with accelerometer data. *PLoS ONE*, *13*(9), 1–12. https://doi.org/10.1371/journal.pone.0203525

Shafer, A. (2005). Meta-analysis of the brief psychiatric rating scale factor structure. *Psychological Assessment*, *17*(3), 324–335. https://doi.org/10.1037/1040-3590.17.3.324

Sharpe, J.-K., Stedman, T. J., Byrne, N. M., & Hills, A. P. (2006). Letter to the Editors: Accelerometry is a valid measure of physical inactivity but not of energy expended on physical activity in people with schizophrenia. In C. Casey Faulkner, Lees, Schutz, Trost (Ed.), *Schizophrenia Research* (Vol. 85, Issues 1–3, pp. 300–301). Elsevier Science. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc5&NEWS=N&AN=2006-08996-033

Shaw, K., Gennat, H. C., O'Rourke, P., & Del Mar, C. (2007). *Exercise for overweight or obesity (Review)*. *1*, 1–68. papers3://publication/uuid/5EC560E9-798F-4DF2-81C5-1F5F180808FF

Sheehan, D. V, Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59 Suppl 2*, 22–57.

Silverstone, T., Smith, G., & Goodall, E. (1988). Prevalence of obesity in patients receiving depot antipsychotics. *British Journal of Psychiatry*. https://doi.org/10.1192/bjp.153.2.214

Simeone, J. C., Ward, A. J., Rotella, P., Collins, J., & Windisch, R. (2015). An evaluation of variation in published estimates of schizophrenia prevalence from 1990-2013: A systematic literature review. *BMC Psychiatry*, *15*(1), 1–14. https://doi.org/10.1186/s12888-015-0578-7

Skender, S., Ose, J., Chang-Claude, J., Paskow, M., Brühmann, B., Siegel, E. M., Steindorf, K., & Ulrich, C. M. (2016). Accelerometry and physical activity questionnaires - a systematic review. *BMC Public Health*, *16*(1), 515. https://doi.org/10.1186/s12889-016-3172-0

Slater, J. A., Botsis, T., Walsh, J., King, S., Straker, L. M., & Eastwood, P. R. (2015). Assessing sleep using hip and wrist actigraphy. *Sleep and Biological Rhythms*, *13*(2), 172–180. https://doi.org/10.1111/sbr.12103

Soundy, A., Roskell, C., Stubbs, B., & Vancampfort, D. (2014). Selection, Use and Psychometric Properties of Physical Activity Measures to Assess Individuals with Severe Mental Illness: A Narrative Synthesis. *Archives of Psychiatric Nursing*, *28*(2), 135–151. https://doi.org/10.1016/j.apnu.2013.12.002

Soundy, A., Taylor, A., Faulkner, G. E., & Rowlands, A. (2007). Psychometric Properties of the

7-Day Physical Activity Recall Questionnaire in Individuals with Severe Mental Illness. *Archives of Psychiatric Nursing*, *21*(6), 309–316. https://doi.org/http://dx.doi.org/10.1016/j.apnu.2007.03.001

Soundy, A., Wampers, M., Probst, M., De Hert, M., Stubbs, B., & Vancampfort, D. (2013). Physical activity and sedentary behaviour in outpatients with schizophrenia: A systematic review and meta-analysis. *International Journal of Therapy and Rehabilitation*, *20*(12), 588–596.

Speakman, J. R., & Westerterp, K. R. (2010). Associations between energy demands, physical activity, and body composition in adult humans between 18 and 96 y of age. *The American Journal of Clinical Nutrition*, *92*(4), 826–834.

Štefelová, N., Dygrýn, J., Hron, K., Gába, A., Rubín, L., & Palarea-Albaladejo, J. (2018). Robust compositional analysis of physical activity and sedentary behaviour data. *International Journal of Environmental Research and Public Health*, *15*(10). https://doi.org/10.3390/ijerph15102248

Strassnig, M., Brar, J. S., & Ganguli, R. (2003). Nutritional Assessment of Patients With Schizophrenia: A Preliminary Study. *Schizophrenia Bulletin*, *29*(2), 393–397. https://doi.org/10.1093/oxfordjournals.schbul.a007013

Strath, S. J., Pfeiffer, K. A., & Whitt-Glover, M. C. (2012). Accelerometer use with children, older adults, and adults with functional limitations. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S77–S85. https://doi.org/10.1249/MSS.0b013e3182399eb1

Stubbs, B., Firth, J., Berry, A., Schuch, F. B., Rosenbaum, S., Gaughran, F., Veronesse, N., Williams, J., Craig, T., Yung, A. R., & Vancampfort, D. (2016). How much physical activity do people with schizophrenia engage in? A systematic review, comparative meta-

analysis and meta-regression. In *Schizophrenia Research* (Vol. 176, Issues 2–3). https://doi.org/10.1016/j.schres.2016.05.017

Stubbs, B., Vancampfort, D., De Hert, M., & Mitchell, a J. (2015). The prevalence and predictors of type two diabetes mellitus in people with schizophrenia: a systematic review and comparative meta-analysis. *Acta Psychiatrica Scandinavica*, *132*(2), 144–157. https://doi.org/10.1111/acps.12439

Stubbs, B., Williams, J., Gaughran, F., & Craig, T. (2016). How sedentary are people with psychosis? A systematic review and meta-analysis. *Schizophrenia Research*, *171*(1–3), 103–109. https://doi.org/10.1016/j.schres.2016.01.034

Suetani, S., Rosenbaum, S., Scott, J. G., Curtis, J., & Ward, P. B. (2016). Bridging the gap: What have we done and what more can we do to reduce the burden of avoidable death in people with psychotic illness? In *Epidemiology and Psychiatric Sciences*. https://doi.org/10.1017/S2045796015001043

Sugawara, N., Yasui-Furukori, N., Sato, Y., Saito, M., Furukori, H., Nakagami, T., Kudo, S., & Kaneko, S. (2013). Body mass index and quality of life among outpatients with schizophrenia in Japan. *BMC Psychiatry*, *13*(1), 1–6. https://doi.org/10.1186/1471-244X-13-108

Suvisaari, J., Perälä, J., Saarni, S. I., Härkänen, T., Pirkola, S., Joukamaa, M., Koskinen, S., Lönnqvist, J., & Reunanen, A. (2008). Type 2 diabetes among persons with schizophrenia and other psychotic disorders in a general population survey. *European Archives of Psychiatry and Clinical Neuroscience*, *258*(3), 129–136. https://doi.org/10.1007/s00406-007-0762-y

Sylvia, L. G., Bernstein, E. E., Hubbard, J. L., Keating, L., & Anderson, E. J. (2014). Practical

guide to measuring physical activity. *Journal of the Academy of Nutrition and Dietetics*, *114*(2), 199–208. https://doi.org/10.1016/j.jand.2013.09.018

Tamminen, N., Reinikainen, J., Appelqvist-Schmidlechner, K., Borodulin, K., Mäki-Opas, T., & Solin, P. (2020). Associations of physical activity with positive mental health: A population-based study. *Mental Health and Physical Activity*, *18*(January), 100319. https://doi.org/10.1016/j.mhpa.2020.100319

Tandon, R., Keshavan, M. S., & Nasrallah, H. A. (2008). Schizophrenia,"just the facts": What we know in 2008: Part 1: Overview. *Schizophrenia Research*, *100*, 4–19. https://doi.org/10.1016/j.schres.2008.01.022

Templ, M., Hron, K., & Filzmoser, P. (2011). robCompositions: An R-package for Robust Statistical Analysis of Compositional Data. In *Compositional Data Analysis: Theory and Applications*. https://doi.org/10.1002/9781119976462.ch25

Thomas, D., Elliott, E. J., & Naughton, G. A. (2009). Exercise for type 2 diabetes mellitus (Review). *The Cochrane Library*, *1*.

Thorp, A. a, Owen, N., Neuhaus, M., & Dunstan, D. W. (2011). Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996-2011. *American Journal of Preventive Medicine*, *41*(2), 207–215. https://doi.org/10.1016/j.amepre.2011.05.004

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S. F. M., Altenburg, T. M., & Chinapaw, M. J. M. (2017). Sedentary Behavior Research Network (SBRN) – Terminology Consensus Project process and outcome. *International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 75. https://doi.org/10.1186/s12966-017-0525-8

Tremblay, M. S., Carson, V., Chaput, J. P., Connor Gorber, S., Dinh, T., Duggan, M., Faulkner, G. E., Gray, C. E., Grube, R., Janson, K., Janssen, I., Katzmarzyk, P. T., Kho, M. E., Latimer-Cheung, A. E., LeBlanc, C., Okely, A. D., Olds, T., Pate, R. R., Phillips, A., … Zehr, L. (2016). Canadian 24-hour movement guidelines for children and youth: An integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology, Nutrition and Metabolism*. https://doi.org/10.1139/apnm-2016-0151

Tremblay, M. S., Chaput, J. P., Adamo, K. B., Aubert, S., Barnes, J. D., Choquette, L., Duggan, M., Faulkner, G. E., Goldfield, G. S., Gray, C. E., Gruber, R., Janson, K., Janssen, I., Janssen, X., Jaramillo Garcia, A., Kuzik, N., LeBlanc, C., MacLean, J., Okely, A. D., … Carson, V. (2017). Canadian 24-Hour Movement Guidelines for the Early Years (0-4 years): An Integration of Physical Activity, Sedentary Behaviour, and Sleep. *BMC Public Health*. https://doi.org/10.1186/s12889-017-4859-6

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical Activity in the United States measured by accelerometer. *Medicine & Science in Sports & Exercise*, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3

Trost, S. G., McIver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine & Science in Sports & Exercise*, *37*(Supplement), S531–S543. https://doi.org/10.1249/01.mss.0000185657.86065.98

van der Ploeg, H. P., & Hillsdon, M. (2017). Is sedentary behaviour just physical inactivity by another name? *International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 1–8. https://doi.org/10.1186/s12966-017-0601-0

Van Hees, V. T., Sabia, S., Anderson, K. N., Denton, S. J., Oliver, J., Catt, M., Abell, J. G., Kivimäki, M., Trenell, M. I., & Singh-Manoux, A. (2015). A novel, open access method to

assess sleep duration using a wrist-worn accelerometer. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0142533

Vancampfort, D., Correll, C. U., Galling, B., Probst, M., De Hert, M., Ward, P. B., Rosenbaum, S., Gaughran, F., Lally, J., & Stubbs, B. (2016). Diabetes mellitus in people with schizophrenia, bipolar disorder and major depressive disorder: a systematic review and large scale meta-analysis. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *15*(2), 166–174. https://doi.org/10.1002/wps.20309

Vancampfort, D., De Hert, M., De Herdt, A., Soundy, A., Stubbs, B., Bernard, P., & Probst, M. (2014). Associations between perceived neighbourhood environmental attributes and self-reported sitting time in patients with schizophrenia: A pilot study. In C. Celis-Morales Cook, Craig, De Bourdeaudhuij, Duvivier, Faulkner, Gorczynski, Humpel, Janney, Lackey, Lin, McGinn, Nathans-Barel, Owen, Pate, Prins, Sallis, Scheewe, Spittaels, Strassnig, Van Dyck, Vancampfort, Vancampfort, Vancampfort, Vancampfort, Vancam (Ed.), *Psychiatry Research* (Vol. 215, Issue 1, pp. 33–38). Elsevier Science. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc10&NEWS=N&AN=2013-42650-001

Vancampfort, D., De Hert, M., Myin-Germeys, I., Rosenbaum, S., Stubbs, B., Van Damme, T., & Probst, M. (2017). Validity and correlates of the International Physical Activity Questionnaire in first-episode psychosis. *Early Intervention in Psychiatry*, *February 2017*, 562–567. https://doi.org/10.1111/eip.12521

Vancampfort, D., Firth, J., Schuch, F. B., Rosenbaum, S., Mugisha, J., Hallgren, M., Probst, M., Ward, P. B., Gaughran, F., De Hert, M., Carvalho, A. F., & Stubbs, B. (2017). Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and

major depressive disorder: a global systematic review and meta-analysis. *World Psychiatry*, *16*(3), 308–315. https://doi.org/10.1002/wps.20458

Vancampfort, D., Guelinckx, H., De Hert, M., Stubbs, B., Soundy, A., Rosenbaum, S., De Schepper, E., & Probst, M. (2014). Reliability and clinical correlates of the Astrand-Rhyming sub-maximal exercise test in patients with schizophrenia or schizoaffective disorder. *Psychiatry Research*, *220*(3), 778–783. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=25246409

Vancampfort, D., Probst, M., Knapen, J., Carraro, A., & De Hert, M. (2012). Associations between sedentary behaviour and metabolic parameters in patients with schizophrenia. *Psychiatry Research*, *200*(2–3), 73–78. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed11&NEWS=N&AN=2013042615

Vancampfort, D., Probst, M., Sweers, K., Maurissen, K., Knapen, J., Willems, J. B., Heip, T., & De Hert, M. (2012). Eurofit test battery in patients with schizophrenia or schizoaffective disorder: Reliability and clinical correlates. *European Psychiatry : The Journal of the Association of European Psychiatrists*, *27*(6), 416–421. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed10&NEWS=N&AN=2012415458

Vancampfort, D., Rosenbaum, S., Probst, M., Connaughton, J., du Plessis, C., Yamamoto, T., & Stubbs, B. (2016). What are the top 10 physical activity research questions in schizophrenia? *Disability and Rehabilitation*, *8288*(June), 1–9. https://doi.org/10.3109/09638288.2015.1116622

Vancampfort, D., Stubbs, B., Mitchell, A. J., De Hert, M., Wampers, M., Ward, P. B., Rosenbaum, S., & Correll, C. U. (2015). Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic disorders, bipolar disorder and major depressive disorder: a systematic review and meta-analysis. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *14*(3), 339–347. https://doi.org/10.1002/wps.20252

Warburton, D. E. R., Charlesworth, S., Ivey, A., Nettlefold, L., & Bredin, S. S. D. (2010). A systematic review of the evidence for Canada's Physical Activity Guidelines for Adults. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(1), 39. http://www.ijbnpa.org/content/7/1/39

Ware, J. (2002). How to score version 2 of the SF-12 health survey. In *Health Assessment Lab*.

Warren, T. Y., Barry, V., Hooker, S. P., Sui, X., Church, T. S., & Blair, S. N. (2010). Sedentary behaviors increase risk of cardiovascular disease mortality in men. *Medicine and Science in Sports and Exercise*, *42*(5), 879–885. https://doi.org/10.1249/MSS.0b013e3181c3aa7e

Welk, G. J., Beyler, N. K., Kim, Y., & Matthews, C. E. (2017). Calibration of Self-Report Measures of Physical Activity and Sedentary Behavior. *Medicine & Science in Sports & Exercise*, *49*(7), 1473–1481. https://doi.org/10.1249/MSS.0000000000001237

White, R. L., Babic, M. J., Parker, P. D., Lubans, D. R., Astell-Burt, T., & Lonsdale, C. (2017). Domain-Specific Physical Activity and Mental Health: A Meta-analysis. *American Journal of Preventive Medicine*, *52*(5), 653–666. https://doi.org/10.1016/j.amepre.2016.12.008

Willis, G. B. (1999). Cognitive Interviewing. A "how to" guide. *Evaluation*, 1–37. https://doi.org/10.1525/jer.2006.1.1.9

Woerner, M. G., Mannuzza, S., & Kane, J. M. (1988). Anchoring the BPRS: an aid to improved

reliability. *Psychopharmacology Bulletin*, *24*(1), 112–117.

World Health Organisation. (2004). Global strategy on diet, physical activity and health. *World Health*, *2002*, 57–57. https://doi.org/10.1080/11026480410034349

World Health Organization. (2014). *WHO | Physical activity*. WHO.

Yamamoto, H., Yamamoto, K., Miyaji, S., Yukawa-Inui, M., Hori, T., Tatematsu, S., Yutani, M., Tanaka, K., & Miyaoka, H. (2011). Daily physical activity in patients with schizophrenia. *Kitasato Medical Journal*, *41*, 145–153.

Appendices

Appendix A Questionnaires Used for Data Collection

A.1 International Physical Activity Questionnaire Short Form

We are interested in finding out about the kinds of physical activities that people do as part of their everyday lives. The questions will ask you about the time you spent being physically active in the **last 7 days**. Please answer each question even if you do not consider yourself to be an active person. Please think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

Think about all the **moderate** activities that you did in the **last 7 days**. **Moderate** activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

1. During the **last 7 days**, on how many days did you do **moderate** physical activities like carrying light loads, ***brisk*** walking, bicycling at a regular pace, or doubles tennis? Do not include walking casually.

   _____ **days per week**

   ☐ No moderate physical activities → ***Skip to question 5***

2. How much time did you usually spend doing **moderate** physical activities on one of those days?

   _____ **minutes per day**

   ☐ Don't know/Not sure

Think about all the **vigorous** activities that you did in the **last 7 days**. **Vigorous** physical activities refer to activities that take hard physical effort and make you breathe much harder than normal. Think *only* about those physical activities that you did for at least 10 minutes at a time.

3. During the **last 7 days**, on how many days did you do **vigorous** physical activities like heavy lifting, digging, aerobics, or fast bicycling?

   _____ **days per week**

   ☐ No vigorous physical activities → ***Skip to question 3***

4. How much time did you usually spend doing **vigorous** physical activities on one of those days?

   _____ **minutes per day**

   ☐ Don't know/Not sure

The last question is about the time you spent **sitting** on weekdays during the **last 7 days**. Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television.

5. During the **last 7 days**, how much time did you spend **sitting** on a **week day**?

   _____ **hours and** _____ **minutes per day**

   ☐ Don't know/Not sure

A.2 Demographics Data Collection Form

Date: _____ ID: _____

Age (yrs): _____

Sex: Male ☐ Female ☐

Ethnicity:

African ☐ Aboriginal ☐ Asian ☐

S.Asian ☐ Hispanic ☐ White ☐

Living Arrangements:

Independent ☐ Group (meals provided) ☐ With Family ☐

Group (no meals provided) ☐

Employment Status:

Full-time ☐ Part-time ☐ Not employed ☐

Student ☐ Retired ☐ Other: _____ ☐

Educational Attainment:

High school (no diploma) ☐ High school (diploma) ☐ Postsecondary ☐

Other: _____ ☐

Marital Status: Single ☐ Married ☐ Separated ☐ Divorced ☐

Diagnosis (MINI)
Plus chart review:

Schizophrenia ☐

Schizoaffective ☐

Psychosis NOS ☐

Other ☐ Specify:_____

Substance Abuse: Yes ☐ No ☐

Psychiatric History:
Date of first hospitalization (dd/mm/yy): _____

Date of first antipsychotic rx (dd/mm/yy): _____

Smoking: Number of cigarettes/day _____ Years as a regular smoker _____

Medication:

I. Current antipsychotic:

Medication:	**Dose (mg):**	**Duration: (mth)**
1_____	_____	_____
2_____	_____	_____
3_____	_____	_____

II. Concomitant Psychiatric:

Medication:	**Dose (mg):**	**Duration: (mth)**
1_____	_____	_____
2_____	_____	_____
3_____	_____	_____

III. Concomitant Medical:

Medication:	**Dose (mg):**	**Duration: (mth)**
1_____	_____	_____
2_____	_____	_____
3_____	_____	_____

Measurements:

Height (cms): _____

Weight (kgs): _____

Waist circumference (cms): _____

2

Appendix B Compositional Isotemporal Substitution Effects

Appendix B. Compositional Isotemporal Substitution Effects

Displaced Behaviour (-10-min):	Predicted Mean	Predicted Change		
	None	Sedentary Behaviour (SB)	Light Physical Activity (LPA)	Moderate Physical Activity (MPA)
Replacement Behaviour (+10-min)				
Body Mass Index	29.26			
SB			0.05	-0.05
LPA		-0.05		-0.11
MPA		0.04	0.10	
VPA		-1.37	-1.32	-1.43
SF-12 Health Survey - Physical Composite Score	29.80			
SB			0.03	0.19
LPA		-0.03		0.15
MPA		-0.17	-0.13	
VPA		-0.18	-0.15	0.00
SF-12 Health Survey - Mental Composite Score	51.92			
SB			0.17	-0.91
LPA		-0.17		-1.07
MPA		0.80	0.97	
VPA		-0.44	-0.26	-1.34
Brief Psychiatric Rating Scale (BPRS) - Total	34.21			
SB			-0.03	0.11
LPA		0.03		0.14
MPA		-0.10	-0.13	
VPA		0.99	0.96	1.10
Clinical Global Impression - Severity Scale	3.54			
SB			0.02	0.00
LPA		-0.02		-0.02
MPA		0.00	0.02	
VPA		0.19	0.21	0.19
Apathy Evaluation Scale	30.79			
SB			-0.17	0.36

LPA		0.17		0.53
MPA		-0.31	-0.49	
VPA		-0.05	-0.22	0.31
BPRS-Activation	4.46			
SB			-0.01	-0.05
LPA		0.01		-0.04
MPA		0.05	0.04	
VPA		0.16	0.15	0.11
BPRS-Affective	8.93			
SB			0.00	-0.01
LPA		0.00		-0.02
MPA		0.01	0.01	
VPA		0.01	0.02	0.00
BPRS-Positive	7.67			
SB			0.00	0.04
LPA		0.00		0.04
MPA		-0.04	-0.04	
VPA		0.54	0.54	0.58
BPRS-Negative	7.65			
SB			0.03	0.09
LPA		-0.03		0.06
MPA		-0.08	-0.05	
VPA		0.03	0.06	0.12
BPRS-Resistance	5.53			
SB			-0.05	0.04
LPA		0.05		0.09
MPA		-0.03	-0.08	
VPA		0.22	0.18	0.26
Symbol Coding Task	45.04			
SB			0.18	-0.23
LPA		-0.18		-0.41
MPA		0.20	0.38	
VPA		-1.91	-1.74	-2.15
Letter Number Span Task	13.44			
SB			0.00	0.08
LPA		0.00		0.08
MPA		-0.07	-0.07	
VPA		0.48	0.47	0.55

Cognitive Total (sum of *z*-scores)	0.19			
SB			0.01	0.01
LPA		-0.01		-0.01
MPA		-0.01	0.01	
VPA		-0.03	-0.01	-0.02

Note: The Predicted Mean represents the value predicted based on regression analysis for a White/Caucasian Male individual aged 41 years (the mean/modal characteristics of the sample) with an average daily waking movement behaviour profile of: 482 minutes of SB, 334 minutes of LPA, 79 minutes of MPA and 1 minute of VPA (the geometric mean converted to 24h equivalents and rounded to full minutes, see Table 5.6). Predicted Change is the difference between the Predicted Mean and a 10-minute exchange of movement behaviours between the column (displaced/reduced) and the row (replacement/increased) while maintaining all other participant characteristics. VPA was not included as a displaceable behaviour as the geometric mean was <10-min/day per participant.