BMJ Open Sport & Exercise Medicine

Changes in physical fitness in transgender adolescents receiving gender-affirming hormone therapy: protocol for a prospective cohort pilot study

Sigrid F Wolf ,[1,2] Lisa M Kuhns,[1,2] Raina V Voss,[1,2] Delanie L Rybacki,[1] Rebecca L Carl ,[1,2] Jamie K Burgess,[1,2] Summer Baptiste,[1] Robert Garofalo[1,2]

To cite: Wolf SF, Kuhns LM, Voss RV, *et al*. Changes in physical fitness in transgender adolescents receiving gender-affirming hormone therapy: protocol for a prospective cohort pilot study. *BMJ Open Sport & Exercise Medicine* 2025;**11**:e002477. doi:10.1136/bmjsem-2025-002477

Accepted 17 February 2025

ABSTRACT

There is very little data about the physical fitness of transgender and gender-diverse (TGD) youth. This protocol paper describes the methods for a pilot study that aims to measure changes in fitness levels in TGD adolescents undergoing gender-affirming hormone therapy (GAHT) using FitnessGram, a reliable, validated tool used to measure physical fitness. A secondary aim of the study is to compare the physical fitness of TGD adolescents with normative data for cisgender individuals. This will be a prospective cohort study. We will recruit 25 TGD participants who are 13–18 years old and patients in our organisation's Gender Development Clinic, planning to initiate GAHT. Participants will complete FitnessGram before starting GAHT and undergo repeat assessments 6 months and 12 months after GAHT initiation. Participants will also complete the Physical Activity Questionnaire for Adolescents. We plan to report changes in the participant's ability to achieve the FitnessGram standards over 12 months. Secondary outcomes will include a comparison of the achievement of the FitnessGram standards in the cohort with the reported achievement of standards for cisgender adolescents.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Transgender adolescents are less physically active than cisgender adolescents. Transgender youth taking gender-affirming hormone therapy (GAHT) have body composition and strength changes over time, but little is known about overall physical fitness changes related to hormone use.

WHAT THIS STUDY ADDS

⇒ This longitudinal assessment using validated measures of physical fitness will provide insights into the impact of GAHT on the physical fitness of transgender adolescents.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ We hope these findings will inform larger, multisite studies to further delineate the effects of hormone therapy on physical fitness in transgender adolescents and inform policy and regulation for TGD youth participation in sports that aim to be both inclusive and fair.

INTRODUCTION

There is scarce data about the physical fitness of transgender and gender-diverse (TGD) youth. Cisgender men (CM) are more physically active than cisgender women (CW) at all ages, but overall physical activity levels decline in all youth as children become adolescents.[1] The available literature suggests that TGD youth are even less physically active than their cisgender peers.[2–5]

There are known body composition differences between TGD people receiving gender-affirming hormone therapy (GAHT) and cisgender people. While all children have similar body composition until about 9 years of age, by 18 years of age, CW have a median percentage body fat of 27.8% and CM have a median percentage body fat of 17.0%.[6] During the first year of testosterone treatment, adolescent transgender men (TM) increase their lean body mass and decrease in total fat mass, and these changes largely remain stable after 3 years with a total fat mass comparable with CM.[7 8] Adolescent transgender women (TW) have a decrease in lean mass over the first year of oestrogen treatment, and the total fat mass remains relatively stable or increased after 3 years of GAHT. The total fat mass in TW is still less than that in CW after 3 years of GAHT.[8]

Data suggest that strength may also change in TGD adolescents undergoing GAHT, along with body composition changes.[9] Grip strength, which has been shown to correlate with total muscle strength in adolescents,

[1]Ann & Robert H Lurie Children's Hospital of Chicago, Chicago, Illinois, USA
[2]Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA

Correspondence to
Dr Sigrid F Wolf;
sfwolf@luriechildrens.org

significantly increases in adolescent TMs and decreases in adolescent TWs over the first year of GAHT.[7 9] However, little is known about how body composition and strength changes relate to overall fitness in TGD adolescents initiating GAHT.

To date, there are only a small number of prospective studies of fitness changes in TGD individuals receiving GAHT; none focus on adolescents. Hamilton *et al* found that adult TW had lower relative VO2 max, lower relative jump height and decreased lung function but higher absolute grip strength than those of CW; adult TM had lower grip absolute grip strength and lower VO2 max than those of CM.[10] Chiccarelli *et al* studied transgender adults' Air Force physical fitness testing. They found that TM performed significantly worse than CM on maximum sit-ups in 1 min, push-ups in 1 min and 1.5 mile run time during the first 3 years of receiving GAHT. TW had significantly better run times than CW for the first 2 years of receiving GAHT and better sit-up scores for 4 years. TW push-up scores remained significantly above CW scores after 4 years of receiving GAHT.[11] Alvares *et al* studied adult, non-athlete TW and found that their VO2 max and grip strength were significantly higher than non-athlete CWs but lower than non-athlete CMs. However, there was no significant difference in VO2 max between CWs and TWs when adjusted for fat-free mass.[12] A review by Hilton and Lundberg in 2021 found small reductions in muscle mass, size and strength in TW receiving GAHT, which did not negate the larger baseline differences between cisgender males and females.[13] A 2021 review by Harper *et al* found that although adult TWs' strength, lean body mass and muscle area decreased with hormone therapy, they remained above those of CWs after 36 months of receiving GAHT.[14]

Our pilot study aims to measure changes in fitness levels in TGD adolescents over their first year of receiving GAHT using FitnessGram. FitnessGram is a reliable, validated physical fitness assessment routinely used in US schools and is the most widely used health-related youth fitness assessment tool globally.[15 16] It has been used in prior research to track changes in fitness levels over time. It requires only a small amount of equipment and a modest yearly licensing fee. It can be conducted in clinics or other settings using a long hallway, an exam room or a small gymnasium.

FitnessGram measures body composition, aerobic capacity, muscle strength, endurance and flexibility. It includes six Healthy Fitness Zone (HFZ) performance standards for males and females ages 10–17+ years old. The HFZ standards denote health-related standards that indicate whether the individual is fit enough for overall good health.[17] The HFZ standards were developed using a nationally representative dataset (National Health and Nutrition Examination Survey) based on the risk of developing metabolic syndrome.[6 17 18] Overall, 23–32% of CW and 27–35% of CM meet at least 5 out of 6 HFZ performance standards.[16]

A key component of FitnessGram is the Progressive Aerobic Cardiovascular Endurance Run (PACER). This is an assessment of aerobic fitness, which consists of a 15 m or 20 m shuttle run at an increasingly fast pace until the participant can no longer complete the laps at the required pace. Their PACER performance is categorised into three categories: HFZ, Needs Improvement Zone (NIZ) and Needs Improvement: Health Risk (NIHR). These three categories correlate to a validated VO2 max (aerobic capacity) estimate in 10- to 18-year-old cisgender youth.[19] VO2 max assesses millilitres of oxygen consumed per minute per kilogram of body weight (mL/kg/min) with higher VO2 max, indicating that the person is more fit because they can consume more oxygen during intense exercise. The PACER HFZ standards are based on gender—with a higher estimated VO2 max for CM—and also based on age—with a lower estimated VO2 max for younger age. The PACER HFZ lap standards correlate to estimated VO2 max ≥38.6–44.3 mL/kg/min; NIZ standards correlate to estimated VO2 max ≥35.4–37.4 mL/kg/min and NIHR standards correlate to estimated VO2 max ≤35.4–37.4 mL/kg/min.[15 19 20] Older studies showed that roughly two-thirds of US adolescents met the HFZ standards for aerobic capacity. In more recent studies of US adolescents, 26.1–49% of CWs and 37.6–62.1% of CM met the HFZ standards for aerobic capacity.[21] It should be noted that these studies did not ask about gender identity, so participants were likely assumed to be cisgender.

FitnessGram also evaluates body mass index (BMI) using age and gender-based criterion standards, correlating with the risk of developing metabolic syndrome.[6 17 22 23] It currently uses the Centers for Disease Control and Prevention BMI standards: HFZ=normal weight (BMI <85th percentile), NI=overweight (≥85th and<95th percentile) or NIHR=obese (≥95th percentile). Maintaining a healthy BMI during adolescence is important for health because youth who are obese have a 19–35% prevalence of metabolic syndrome compared with <2% of normal-weight youth.[22] Approximately 52.7–65.0% of cisgender youth meet the criteria for BMI HFZ.[21]

Importantly, no FitnessGram norms for TGD youth consider the changes in strength or physical activity that may occur after initiating GAHT. Therefore, we plan to assess how the adolescent's physical fitness changes over 12 months based on the FitnessGram performance. We also plan to compare the adolescents' physical fitness to CM and CW HFZ standards and determine whether their fitness will more closely align with their affirmed or gender assigned at birth.

OBJECTIVES

The primary objective of our pilot study is to assess changes in physical fitness measured on FitnessGram in TGD adolescents during their first year of receiving GAHT. Secondarily, we aim to evaluate how TGD adolescent FitnessGram performance is compared with reported cisgender adolescent performance both pre-GAHT and after 1 year of receiving GAHT. Additionally,

we will evaluate the correlation between the reported recent level of physical activity and performance on FitnessGram in TGD youth. Finally, we aim to evaluate the correlation between BMI and FitnessGram performance in TGD youth.

METHODS

Study design

This will be a prospective cohort pilot study.

Recruitment

Our pilot study will be conducted at Lurie Children's Hospital's Gender Development Clinic in the Adolescent & Young Adult Medicine division in Chicago, Illinois, USA. This study will be conducted using a consecutive, nonrandom sampling technique. Eligible patients will be identified via medical record review and invited to participate onsite at the clinic. If patients are not recruited during their clinic appointment but are eligible, they will be recruited via email or phone call within 7 days of their appointment. As part of the recruitment, subjects will be provided with a flyer about the study with a QR code linking to a 1 min 49 s professionally produced video demonstrating a TGD adolescent undergoing the FitnessGram testing that will be conducted during the study.

Inclusion criteria

We will enrol TM and TW participants who are patients in our organisation's Gender Development Clinic. Subjects must be 13–18 years old at the time of enrolment and be planning to initiate GAHT (testosterone or oestrogen) within the next 6 months. The subject and their parent/guardian must be able to read and understand English at a sixth-grade level.

Exclusion criteria

We will exclude potential participants who are not patients in our academic centre's Gender Development Clinic, are not within the target age range or cannot read and understand English at a sixth-grade level.

Study proceedings

The patients will complete FitnessGram before starting GAHT and repeat FitnessGram at 6 months and 12 months after starting GAHT. We will administer a demographics questionnaire, a sports specialisation survey[24 25] and The Physical Activity Questionnaire for Adolescents (PAQ-A) at the initial visit.[26 27] We will administer the PAQ-A again at 6 and 12 months post-GAHT initiation. We plan to obtain the most recent height and weight from the medical record (see the study flow chart in figure 1). Data will be collected in REDCap. The FitnessGram testing will take place in the Gender Development Clinic or a private, rented gymnasium located one block

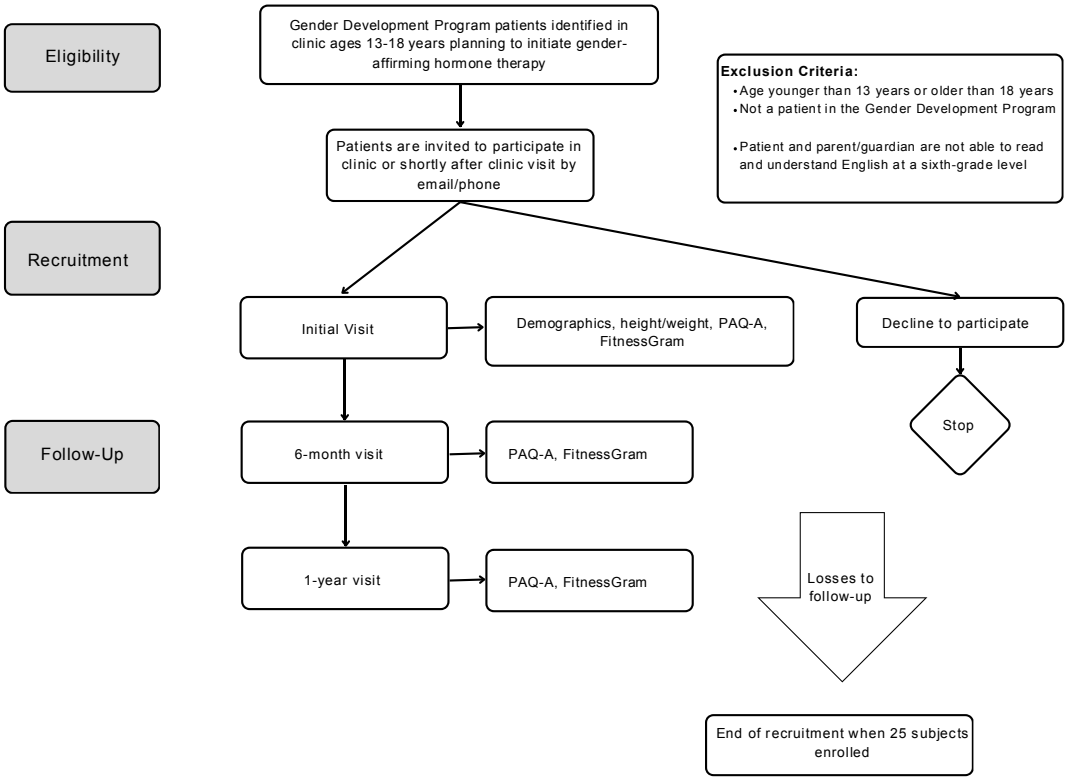

Figure 1 Patient enrolment and study proceedings flow chart. PAQ-A, The Physical Activity Questionnaire for Adolescents.

from the Gender Development Clinic. Data collection will be performed by trained research coordinators and/or study physicians.

A flow chart of patient enrolment and study proceedings is presented in figure 1.

Ethical considerations

The study was approved by the Institutional Review Board of Lurie Children's Hospital (2021-4463). Informed written consent will be obtained from each patient. If the patient is under 18 years of age, they will be assented to, and their legal guardian will be asked to provide parental consent for youth participation. To protect participant confidentiality, we will collect data by identification number and in a password-protected and secure REDCap database.

Funding

This study will be supported by internal philanthropic funding from the Lurie Children's Hospital Division of Adolescent & Young Adult Medicine and grants from the Lurie Children's Hospital Academic Enrichment Fund and the Lurie Children's Hospital Faculty Development Program Fund.

Outcomes

The primary outcome of this pilot study will be the changes in participants' ability to achieve FitnessGram HFZ male and female standards over 12 months. Secondary outcomes will include a comparison of the achievement of the FitnessGram standards in the cohort with the reported achievement of standards for cisgender adolescents. This pilot study will also provide information on the feasibility and acceptability of this type of fitness study in TGD youth.

Statistical analysis

Demographic and physical fitness data collected during the baseline, 6-month and 12-month assessments of participants will be reported using descriptive statistics. Given the small sample sizes for each gender-age group, which limit the ability to reliably assess normality, median and interquartile ranges will be reported for numeric variables referencing physical fitness performance. Discrete variables, such as gender, race, ethnicity, insurance status, whether the participant wore a mask during testing and chest binding status, will be summarised using frequencies and percentages based on the number of participants with available data. This will include the number of participants that met the HFZ standards for each gender-age group. When data are missing, the actual denominator will be stated. On collection of descriptive statistics, we will compare the normative values established for their cisgender counterparts both for the participants' sex assigned at birth and affirmed gender. Due to the small sample size and potential attrition over time, there is insufficient data to develop multivariable models. R V.4.3.1 will be used to conduct the analysis.

Sample size

No power analysis will be performed due to limited potential subjects and a lack of prior data. As a pilot study, we plan to enrol 25 transgender adolescents into our cohort but will not set specific restrictions on gender identity due to the low availability of potential subjects. We may enrol more TMs since our clinic cares for more TMs than TWs and gender-diverse patients. Based on clinic visit flow and anticipated response rate, we anticipate that this study will take 2 years to complete enrolment and 3 years to complete all study visits.

RESULTS

We hypothesise that participants in this prospective cohort pilot study of transgender adolescent patients initiating GAHT will be more likely to meet their affirmed gender's FitnessGram HFZ standards by their 12-month follow-up visit in comparison with the baseline visit. Additionally, we expect to find that a lower percentage of transgender adolescents overall will meet the FitnessGram HFZ standards than has been reported for cisgender youth both before and after receiving GAHT. We anticipate that transgender youth with higher levels of reported recent physical activity at baseline will perform better on FitnessGram than their less active peers. We anticipate that participants with a BMI meeting HFZ standards will have better FitnessGram performance than subjects with BMIs outside the HFZ standards.

DISCUSSION

This prospective pilot study will be an important first step in creating a foundation for understanding the impact of GAHT on TGD adolescents' physical fitness. It will provide valuable data on the acceptability and feasibility of such a study on a larger scale involving multiple sites. It may also bring a call to action for additional research with a more diverse sample and robust outcome measures. These findings may inform policy and future regulations on TGD adolescent sports participation from athletic organising bodies that aim to be inclusive and fair.

Limitations

Our study has several limitations. FitnessGram does not directly measure VO2 max or muscle strength via dynamometry but provides correlates. Our study will have a small sample size. Physical activity and sports participation can be sensitive topics in the TGD community, which may limit the willingness of TGD youth to participate. Our study cohort may be influenced by self-selection bias, with more athletic youth feeling more comfortable participating in a physical fitness study.

Acknowledgements The study team is grateful for the support of the staff, patients and families in the Gender Development Program of Lurie Children's Hospital Division of Adolescent & Young Adult Medicine for undertaking this study.

Contributors SFW, LMK, RVV, RLC, JKB and RG conceived the study design, developed the methodology, provided clinical expertise, authored this manuscript, will oversee recruitment and data collection, and will be responsible for data analysis and future manuscript preparation. DLR will perform data collection

and created the figure for this manuscript. SB will conduct statistical analysis and authored the statistical analysis and sample size portion of this manuscript. Guarantor: SFW accepts full responsibility for the accuracy and integrity of the data presented in this protocol and has access to all study data.

Funding This study will be supported by internal philanthropic funding from the Lurie Children's Hospital Division of Adolescent & Young Adult Medicine, as well as grants from the Lurie Children's Hospital Academic Enrichment Fund (AEF) and the Lurie Children's Hospital Faculty Development Program Fund.

Competing interests RLC is chair for the AAP Council on Sports Medicine and Fitness. She receives travel reimbursement when she travels on their behalf to conferences and has received small honorariums (a total of US$700) for speaking and writing engagements.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Lurie Children's Hospital Institutional Review Board (#2021-4463). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study.

ORCID iDs
Sigrid F Wolf http://orcid.org/0009-0000-9066-3374
Rebecca L Carl http://orcid.org/0000-0001-7642-9872

REFERENCES

1 Trost SG, Pate RR, Sallis JF, *et al*. Age and gender differences in objectively measured physical activity in youth. *Med Sci Sports Exerc* 2002;34:350–5.
2 Voss RV, Kuhns LM, Phillips G 2nd, *et al*. Physical Inactivity and the Role of Bullying Among Gender Minority Youth Participating in the 2017 and 2019 Youth Risk Behavior Survey. *J Adolesc Health* 2023;72:197–206.
3 Lightner JS, Schneider J, Grimes A, *et al*. Physical activity among transgender individuals: A systematic review of quantitative and qualitative studies. *PLoS One* 2024;19:e0297571.
4 Jones BA, Haycraft E, Bouman WP, *et al*. The Levels and Predictors of Physical Activity Engagement Within the Treatment-Seeking Transgender Population: A Matched Control Study. *J Phys Act Health* 2018;15:99–107.
5 VanKim NA, Erickson DJ, Eisenberg ME, *et al*. Weight-related disparities for transgender college students. *Health Behav Policy Rev* 2014;1:161–71.
6 Laurson KR, Eisenmann JC, Welk GJ. Body fat percentile curves for U.S. children and adolescents. *Am J Prev Med* 2011;41:S87–92.
7 Tack LJW, Craen M, Lapauw B, *et al*. Proandrogenic and Antiandrogenic Progestins in Transgender Youth: Differential Effects on Body Composition and Bone Metabolism. *J Clin Endocrinol Metab* 2018;103:2147–56.
8 Boogers LS, Reijtenbagh SJP, Wiepjes CM, *et al*. Time Course of Body Composition Changes in Transgender Adolescents During Puberty Suppression and Sex Hormone Treatment. *J Clin Endocrinol Metab* 2024;109:e1593–601.
9 Wind AE, Takken T, Helders PJM, *et al*. Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? *Eur J Pediatr* 2010;169:281–7.
10 Hamilton B, Brown A, Montagner-Moraes S, *et al*. Strength, power and aerobic capacity of transgender athletes: a cross-sectional study. *Br J Sports Med* 2024;58:586–97.
11 Chiccarelli E, Aden J, Ahrendt D, *et al*. Fit Transitioning: When Can Transgender Airmen Fitness Test in Their Affirmed Gender? *Mil Med* 2023;188:e1588–95.
12 Alvares LAM, Santos MR, Souza FR, *et al*. Cardiopulmonary capacity and muscle strength in transgender women on long-term gender-affirming hormone therapy: a cross-sectional study. *Br J Sports Med* 2022;56:1292–8.
13 Hilton EN, Lundberg TR. Transgender Women in the Female Category of Sport: Perspectives on Testosterone Suppression and Performance Advantage. *Sports Med* 2021;51:199–214.
14 Harper J, O'Donnell E, Sorouri Khorashad B, *et al*. How does hormone transition in transgender women change body composition, muscle strength and haemoglobin? Systematic review with a focus on the implications for sport participation. *Br J Sports Med* 2021;55:865–72.
15 Institute TC. FitnessGram administration manual: the journey to myhealthyzone. In: *Human kinetics*. 5th edn. 2017: 136.
16 Lee EY, Barnes JD, Lang JJ, *et al*. Testing validity of FitnessGram in two samples of US adolescents (12-15 years). *J Exerc Sci Fit* 2020;18:129–35.
17 Welk GJ, De Saint-Maurice Maduro PF, Laurson KR, *et al*. Field evaluation of the new FITNESSGRAM criterion-referenced standards. *Am J Prev Med* 2011;41:S131–42.
18 Eisenmann JC, Laurson KR, Welk GJ. Aerobic Fitness Percentiles for U.S. Adolescents. *Am J Prev Med* 2011;41:S106–10.
19 Mahar MT, Welk GJ, Rowe DA. Estimation of aerobic fitness from PACER performance with and without body mass index. *Meas Phys Educ Exerc Sci* 2018;22:239–49.
20 Mahar MT, Guerieri AM, Hanna MS, *et al*. Estimation of aerobic fitness from 20-m multistage shuttle run test performance. *Am J Prev Med* 2011;41:S117–23.
21 Bai Y, Saint-Maurice PF, Welk GJ, *et al*. Prevalence of Youth Fitness in the United States: Baseline Results from the NFL PLAY 60 FITNESSGRAM Partnership Project. *J Pediatr* 2015;167:662–8.
22 Laurson KR, Welk GJ, Eisenmann JC. Diagnostic performance of BMI percentiles to identify adolescents with metabolic syndrome. *Pediatrics* 2014;133:e330–8.
23 Laurson KR, Welk GJ, Marton O, *et al*. Agreement and Diagnostic Performance of FITNESSGRAM, International Obesity Task Force, and Hungarian National BMI Standards. *Res Q Exerc Sport* 2015;86 Suppl 1:S21–8.
24 Jayanthi NA, LaBella CR, Fischer D, *et al*. Sports-specialized intensive training and the risk of injury in young athletes: a clinical case-control study. *Am J Sports Med* 2015;43:794–801.
25 Miller M, Malekian S, Burgess J, *et al*. Evaluating a Commonly Used Tool for Measuring Sport Specialization in Young Athletes. *J Athl Train* 2019;54:1083–8.
26 Janz KF, Lutuchy EM, Wenthe P, *et al*. Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. *Med Sci Sports Exerc* 2008;40:767–72.
27 Kowalski K, Crocker P, Donen R. The Physical Activity Questionnaire for Older Children (PAQ-C) and Adolescents (PAQ-A) Manual. *Coll Kinesiol Univ Sask* 2004;1–38.