**METHODS**

**Study Design**

This study was designed to determine what variables in children predict the cadence at which a walk transitions to a run. To address this question we used data collected at the University of Massachusetts, Amherst’s Physical Activity and Health Laboratory to develop models that predict the cadence at the walk to run transition from an easily measurable set of demographic and metabolic variables.

**Participants**

Some questions

1. When was the data collected and what was the primary outcome?
2. Has the data already been described elsewhere?
3. How were the participant’s recruited?
4. How many minutes were the children asked to walk at a constant pace?
5. How was “run” versus “walk” defined?
6. What is run cadence?

The original dataset contained 122 participants between the ages of 7 and 20. Participants were guided to walk at a constant pace for \_\_\_ minutes. At the end of this interval, the speed was increased by 0.5 mph. The increase in speed was continued until the individual transitioned from walk to run. The activity concluded at the walk to run transition interval. Of the 122 participants, 69 were able to attain the walk to run interval and only these data were used for model development.

**Data preparation**

Using the ‘findLinearCombos’ command from the “CARET” package in R (**R** Core Team 2012) and the ‘cor’ command we eliminated linear dependencies and highly correlated variables from the consideration. DO WE NEED TO LIST WHAT THEY WERE?? DO YOU REMEMBER WHAT THEY WERE? I’M TRYING TO EASILY FIGURE IT OUT FROM THE CODE, BUT I’M SLEEPY AND I HAD A CONCUSSION THE OTHER DAY, YA KNOW?! ☺☺☺ The dependent variable for all models was the participant walk to run transition cadence.

**Model Development**

Regression models were developed using a set of independent variables after linear dependencies were removed. The best subsets method of identifying the optimal model was applied using the “leaps” package in **R** (**R** Core Team 2012). The leaps algorithm was used to compare regression models for every possible subset of factors, selecting the model with the lowest Bayesian information criterion (BIC). A k-means clustering approach feeding a Gaussian mixture model and regularization methods were also considered were also developed.

**R Shiny App Development**

An R Shiny application was created in R (R Core Team 2012) to provide users with the expected pace at which the minor will transition from walking to running.

**RESULTS**

**Participant characteristics**

The 69 participants are a mix of males and females which are treated identically in model development. Participant characteristics in the reference database appear in **Table 1**.

**Model Variables**

The list of independent variables considered for model development after accounting for linear dependencies appear in **Table 2**. Two sets of highly correlated (greater than 90%) independent variables were identified; waist circumference strongly correlates with weight while BMI percentile correlates with BMI z-score. Waist circumference and BMI percentile were removed from consideration because weight is an easier and more practical measure for an individual to attain and because BMI z-score is a more rigorous representation of a person’s BMI by accounting for age and gender.

**Mathematical Models**

After comparing regression models for every possible subset of factors, the leaps algorithm arrived at an optimal regression model:

**Other Considered Models**

The k-means clustering approach feeding a Gaussian mixture model did not provide the predictive power of the regression model based on BIC. Regularization methods also did not outperform multiple linear regression.

**App Development**

Figure 1 depicts the user interface of the developed R Shiny app available at <https://dustyturner.shinyapps.io/KidsStep/>. After the user inputs age, gender, height, and weight the application returns the expected walk-to-run transition cadence with a 95% confidence interval. The application also calculates an individual’s BMI and subsequent BMI z-score.

**FIGURE LEGENDS:**

**Figure 1:** Screen shot of the R Shiny App. After user input, BMI (green) and BMI z-score (blue) are output with expected walk-to-run transition cadence (orange). The graphs on the right displays walk-to-run transition cadence in response to changes in user inputs.



**Table 1:** Subject characteristics. Results are presented in mean ± SD.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | N | Age | BMI | Walk-to-Run Transition Cadence |
| Males | 37 | 15.0 ± 3.88 | 22.6 ± 6.01 | 156 ± 10.2 |
| Females | 32 | 14.9 ± 3.70 | 21.9 ± 4.66 | 158 ± 8.92 |

**Table 2** List of independent variables used to develop models..

|  |  |
| --- | --- |
| **Independent Variable** | **Explanation** |
| Sex | Male or Female |
| Age (years) | Age of participant in years |
| Height (cm) | Height of participant in cm |
| Weight (kg) | Body Weight |
| Waist (cm) | Waist circumference |
| BMI (kg/m2) | Body Mass Index |
| BMI percentile | The percentile of the participant’s BMI |
| BMI z-score | The standardized BMI score for the participant based on age and gender |
| Classification of Obesity Status | 85th percentile BMI designated overweight, 95th percentile BMI designated as having obesity |
| % body fat | Measured using bioimpedence |
| Walk VO2 | Oxygen uptake in L/min during walking phase. |
| Run VO2 | Oxygen uptake in L/min during running phase. |
| Run METS Youth1 |  |
| Run METS Youth 3 |  |
| Walk METS Youth 3 |  |