

RESEARCH

Recursive Partitioning Methods to Examine the Effects of Physical Activity on Physical and Mental Health: A Case Study at One Public University in the United States

Samantha Robinson and Jessica Gilmore

University of Arkansas, US

Corresponding author: Samantha Robinson (sewrob@uark.edu)

The purpose of this study was to examine the effects of physical activity (PA) on physical and mental health within an adult sample, while controlling for demographic characteristics. Data were analyzed from the 2018–2019 Exercise is Medicine® On Campus (EIM-OC) assessment of students, faculty, and staff at one public university in the United States. Participants completed a survey including a demographic questionnaire, the 12-Item Short Form Health Survey (SF-12), and the International Physical Activity Questionnaire (IPAQ). Classification and Regression Trees, a supervised learning technique based upon recursive partitioning, were utilized to explore the relationships between PA and health. Variable-importance was assessed via the permutation method and confirmed with ensemble learning methods. Results indicated that PA, especially in leisure, was predictive of both physical and mental health. Further, leisure PA was the single most important variable contributing to mental health. Greater levels of physical health were observed in adults achieving at least 672 MET minutes of leisure PA per week, while greater levels of mental health were observed in adults achieving at least 950 MET minutes of leisure PA per week. Despite current recommendations for PA (irrespective of domain), more PA (especially in leisure) improves physical and mental health.

Keywords: exercise is medicine on campus; international physical activity questionnaire; short form health survey; regression trees; variable importance

Introduction

Numerous studies have cited both an overall decline in physical activity (PA) over the last several decades as well as increasing rates of physical inactivity and sedentary behavior, especially in high-income countries such as the United States (US) (Brownson et al., 2005; Guthold et al., 2018; Katzmarzyk & Mason, 2009; US Department of Health and Human Services [DHHS], 2018). Released in 2018, the second PA guidelines in the US recommend that adults should participate in at least 150 minutes of moderate-intensity aerobic activity, 75 minutes of vigorous-intensity aerobic activity, or an equivalent combination of both per week i.e., approximately 600 metabolic equivalent (MET) minutes per week (DHHS, 2018; Riebe et al., 2021). Despite recommendations, only 50.3% of adults in the US met the PA guidelines in 2017 and the most recent US National Health and Nutrition Examination Survey (NHANES) indicated that both children and adults spend approximately 7.7 hours per day being sedentary (DHHS, 2018; Centers for Disease Control and Prevention, n.d.). This observed decrease in PA along with increases in physical inactivity and sedentary behavior is not restricted to the US or even to high-income countries however. The World Health Organization (WHO) also reports that at the global level, more than 1 in 4 adults is insufficiently physically active (WHO, 2016). The current prevalence of insufficient physical activity is even higher in adolescents aged 11–17 years, with over 80% insufficiently active (Guthold et al., 2020).

Research has consistently shown that PA correlates with improved physical health outcomes. Moderate-to-vigorous PA is associated with increased life expectancy and decreased risk for Non-Communicable Diseases (NCDs) such as cardiovascular disease (including heart disease and stroke), cancers (e.g., bladder, breast, colon, endometrium, esophagus, kidney, lung, and stomach), respiratory diseases, and diabetes (DHHS, 2018; Miles, 2007; Moore et al., 2012; WHO, 2018). PA is known to improve physical health and quality of life, while reducing all-cause mortality risk, in all persons regardless of age, race, ethnicity, or disability (DHHS, 2018).

While a majority of US adults and adolescents were already insufficiently active prior to the COVID-19 pandemic, PA has declined even further due to the subsequently mandated stay-at-home orders, additional restrictions on movement, increased unemployment levels, and forced transitions to remote working and learning (Alomari et al., 2020; Cheikh, et al., 2020; Constant et al., 2020). This decline in reported PA and increase in sedentary behavior has been observed in both adults and in adolescents who are, potentially, at greater risk for mental health issues such as anxiety and depression (Alomari et al., 2020; Huber et al., 2020).

Even before the pandemic, mental health was an increasingly prevalent public health issue in the US. The National Institute of Mental Health (NIMH) estimated that 51.5 million adults (over 1/5 of all US adults) had a mental illness in 2019, with major depression listed as one of the most common mental disorders (DHHS, 2018; NIMH, n.d.). A growing body of literature indicates that engaging in regular PA improves overall mental health and wellbeing while reducing the effects of mental health disorders such as anxiety and depression (DHHS, 2018; Fox, 1999; Paluska & Schwenk, 2000; Penedo & Dahn, 2005; Saxena et al., 2005). Benefits of PA on mental health have been observed in both the adult population and among adolescents (Biddle & Asare, 2011; Tyson et al., 2010).

Overall health, both physical and mental, is shown to improve with routine moderate-to-vigorous PA. However, many physical and mental health benefits appear to follow a dose-response pattern such that PA that occurs more consistently, for longer periods of time, and at higher levels of intensity results in additional physical and mental health benefits (DHHS, 2018; Hamer et al., 2008; Miles, 2007; Tyson et al., 2010). Moreover, while PA recommendations tend to focus on total PA, a growing amount of research suggests that the domain in which PA occurs (active/transport, domestic, leisure, or work) may result in differential health benefits (White et al., 2017). In fact, rather than compensating for the lack of PA in one domain, PA that primarily occurs in another domain may even be detrimental to health (Coenen et al., 2018). Given that physical and mental health benefits of PA appear to be domain-specific and to follow a dose-response pattern, current recommendations for PA might not be best practices.

To the authors' knowledge, no study has sought to examine associations between health and domain-specific physical activity (PA) using recursive partitioning statistical methods e.g., Classification and Regression Trees (CART) methods. Utilizing such methods could inform the creation of public health protocols for PA, guiding decision-making and individualized PA recommendations specific to both physical and mental health outcomes. Consequently, the purpose of the current study is twofold. The primary purpose of this study is to explore the relationships that exist between PA and health, while controlling for demographics. The secondary purpose of this study is to illustrate the benefits of CART methods, which facilitate visualization of existent relationships and contextualize these relationships for use in individualized PA recommendations and protocol creation. Given the exploratory nature of the current study there no hypotheses were established a priori.

Method

Participants and Recruitment

Exercise is Medicine® On Campus (EIM-OC) is an initiative grounded in the EIM global mission (EIM, 2019). With over 280 campuses registered worldwide, EIM-OC calls upon university and college campuses to promote PA as a vital indicator of health and to create a campus culture that encourages PA and physical movement among faculty, staff, and students (EIM, 2019).

Participants were recruited from one registered, silver-level recognized EIM-OC public university campus community through physical advertisements such as flyers and posters, through electronic advertisements such as the daily emailed campus news and university social media posts, and through resulting snowball sampling in the form of word-of-mouth recruitment. All participants were members of the campus community i.e., faculty, staff, or students. Membership in the campus community was the only inclusion criteria for the study. Participants were asked to complete surveys regarding their health, PA, and demographic characteristics. While there was no formal compensation for survey completion, participants that completed the survey were entered into a drawing for the chance to win one of five 50 USD cash prizes available. The study had Institutional Review Board (IRB) approval.

Instruments and Measures

SF-12

The SF-12 Health Survey is a shortened version of the SF-36 survey. Designed as an alternative instrument to measure health, the 12-item instrument was constructed using a subset of the original 36-item survey. Reliability and validity are adequate across a number of populations for the two-component model of health (Gandek et al., 1998; Jenkinson et al., 1997; Lundberg et al., 1999; Ware et al., 1996).

Utilizing a mix of yes/no and Likert-type response questions that ask participants to self-report health and functioning during the past seven days, the SF-12 models both physical and mental health components, resulting in two summary scores i.e., the Physical Component Summary (PCS) and Mental Component Summary (MCS) scores.

The PCS and MCS scores range from 0 to 100, with lower values being indicative of lower levels of health and functioning and higher values being indicative of higher levels of health and functioning. Both scores are t-scores with a mean of 50 and a standard deviation of 10 in the US general population (Ware et al., 1996).

Given that the SF-12 is not an age-specific measure and given that physical and mental health component scores are thought to vary throughout an individual's lifespan, the use of age-specific mean difference scores is suggested for interpretation (Utah Department of Health [UDH], 2001). The difference between each observed SF-12 component score and the mean score relative to the participant's age group resulted in age-specific mean difference scores for the PCS and MCS measures. These age-specific mean difference scores can theoretically range from -100 to $+100$ and express both the direction and the distance an individual's score is from the mean score of their age group. For example, an individual who has a difference score of -10 scored ten points lower than the mean component score for their specific age group. This indicates that the individual has below average self-reported health and functionality relative to their age group. The use of age-specific mean difference scores allows us to assess individuals across different age groups, because the difference score has the same meaning of health relative to an individual's specific age group i.e., being above or below average health and functionality.

IPAQ

The International Physical Activity Questionnaire (IPAQ) is an instrument used to assess PA in adults aged 15–96 years old (IPAQ, n.d.). The IPAQ is a reliable and valid self-report instrument of PA and displays adequate measurement properties across nations (Craig et al., 2003). Participants are asked about walking, moderate-intensity activity, and vigorous-intensity activity in the PA domains of work, transportation, domestic activity, and leisure time. Individuals are asked to provide both the duration and frequency of physical activity in the four PA domains using two alternate reference periods, either during the “last 7 days” or during a “usual week”. Answers are then converted into MET minutes per week, with higher values indicating more PA in that specific domain.

Analytic Approach

Data Pre-processing

Preliminary exploration of the 2018–2019 EIM-OC data revealed approximately 9% of observations contained missing values. Missing values for the variables were imputed using Multivariate Imputation by Chained Equations (MICE), also called full conditional specification (Buuren & Groothuis-Oudshoorn, 2011). MICE was conducted with a non-parametric, Random Forest (RF) algorithm, as developed by Stekhoven et al. (2012). This approach to missing value imputation results in less biased parameter estimates (Shah et al., 2014).

Diagnostic checks were conducted for the missing value imputation procedure. MICE implements an iterative Markov Chain Monte Carlo (MCMC) type of algorithm during missing value imputation. Convergence of the algorithm was checked with diagnostic trace plots using 100 iterations. **Figure 1** displays the trace plots for a selection of the variables. There was no apparent trend, and the trace lines mix well in both mean and standard deviation, suggesting algorithm convergence. Additionally, the plausibility of imputed values was checked by comparing the univariate distributions of imputed values to those of observed values (see **Figure 2**). Imputed values were observed to be within the range of the observed data and the distributions of imputed and observed values appeared similar suggesting that the imputations were plausible, and that the data was missing completely at random (MCAR).

In addition to missing value imputation, health component age-specific mean difference scores were calculated by subtracting the US age group average health component summary scores from every observed health component summary score in the current sample, for PCS and MCS scores. This was done to control for the intrinsic age effects that are commonly observed in physical and mental health as measured by the

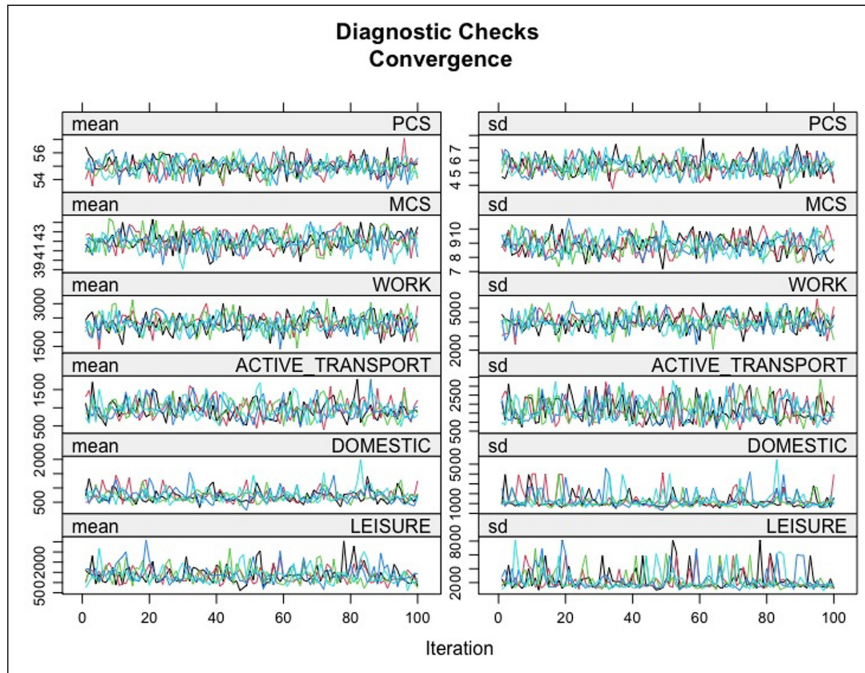


Figure 1: Diagnostic trace plots for the convergence of the Multivariate Imputation by Chained Equations (MICE) iterative Markov Chain Monte Carlo (MCMC) algorithm.

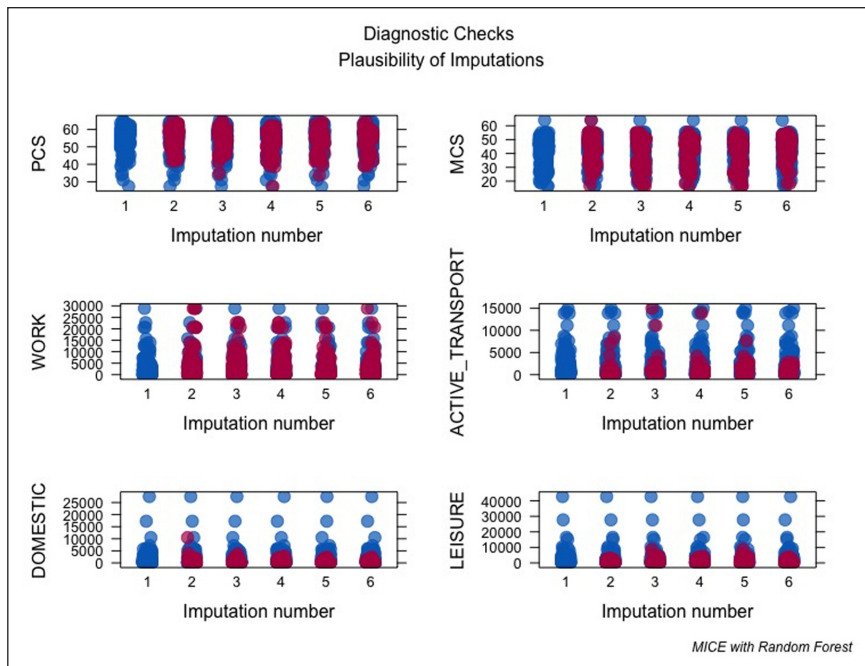


Figure 2: Diagnostic plots for the plausibility of Multivariate Imputation by Chained Equations (MICE) imputations.

SF-12. PCS scores and MCS scores often differ during an individual’s life span as well as across age groups, with PCS scores typically declining and MCS scores typically improving with age (UDH, 2001).

Statistical Analysis

Classification and regression trees (CART) methods are a type of nonparametric, supervised learning technique used to predict a dependent variable using several independent variables. The resulting prediction can be visualized in the form of a decision tree. If predicting a categorical variable, a classification tree is created. If predicting a continuous variable, as in the current study, a regression tree is created. Recursive partitioning is the statistical method for multivariable analysis that is utilized to create such a tree and refers to the

process of recursively splitting the feature space into several binary, if-then splits until a certain stopping criterion is met (Breiman, 1984). According to Strobl, Malley, and Tutz (2009), the main characteristic of this nonparametric regression approach is that the recursive partitions are made such that similar observations become grouped together.

Regression trees were constructed to separately predict the two dependent variables i.e., age-specific mean difference PCS and MCS scores using domain-specific PA and participant demographics as predictor and control variables of interest. Recursive partitioning splits in the regression trees were made to minimize the Residual Sum of Squares (RSS) and each tree was pruned so that tree formation stopped when splitting groups in terminal nodes (i.e., the last groupings created by the splits) had 25 or fewer observations.

Variable-importance of the predictor variables was then assessed via the permutation method using the Root Mean Squared Error (RMSE) loss-function (Breiman, 2001). The robustness of the variable-importance results was confirmed using ensemble learning methods, which use multiple learning algorithms in aggregate e.g., a set of several regression trees. One of the most commonly used ensemble methods i.e., bootstrap aggregating was utilized in the current study. Bootstrap aggregating, also known as bagging, uses multiple bootstrap samples to create multiple regression trees, which are then combined for prediction and can be utilized to ascertain the importance of each predictor variable. In the current study, bagging of 500 regression trees was utilized to assess the robustness of the permutation-based variable-importance results.

All data pre-processing and all statistical analysis were conducted in R, which is a freely available software for statistical computing developed by the R Core Team and maintained by the R Foundation for Statistical Computing (R Core Team, 2020).

Results

Participant Demographics

Demographic features of the sample participants are displayed in **Table 1**. Of the 398 participants, the vast majority were female (76.2%) and white (81.1%), with 94.4% of the individuals identifying as non-Hispanic.

Table 1: Demographic features representative of the sample participants.

| Demographic | N = 398* | | |
|---------------------------|-----------------|----------------------------------|------------|
| <i>Sex/Gender</i> | | <i>Education</i> | |
| <i>Female</i> | 301 (76.2) | <i>High School</i> | 34 (15.5) |
| <i>Male</i> | 94 (23.8) | <i>College/University</i> | 86 (39.3) |
| <i>Ethnicity</i> | | <i>Graduate School</i> | 99 (45.2) |
| <i>Hispanic</i> | 22 (5.6) | <i>Mother's Education</i> | |
| <i>Non-Hispanic</i> | 370 (94.4) | <i>Below High School</i> | 10 (2.8) |
| <i>Race</i> | | <i>High School</i> | 118 (32.9) |
| <i>Asian</i> | 25 (6.3) | <i>College/University</i> | 141 (39.3) |
| <i>Black</i> | 13 (3.3) | <i>Graduate School</i> | 90 (25.1) |
| <i>Multiple</i> | 18 (4.5) | <i>Father's Education</i> | |
| <i>White</i> | 321 (81.1) | <i>Below High School</i> | 18 (5.1) |
| <i>Other</i> | 19 (4.8) | <i>High School</i> | 114 (32.4) |
| <i>Role</i> | | <i>College/University</i> | 131 (37.2) |
| <i>Admin or Faculty</i> | 39 (9.9) | <i>Graduate School</i> | 89 (25.3) |
| <i>Graduate Student</i> | 63 (16.0) | <i>Greek Life</i> | |
| <i>Staff</i> | 125 (31.7) | <i>Current</i> | 69 (39.9) |
| <i>Undergraduate</i> | 167 (42.4) | <i>Former</i> | 8 (4.6) |
| <i>Age (years)</i> | 32.3 (14.3) | <i>Never</i> | 96 (55.5) |

* For Categorical Variables, table values are Count (%); For Continuous Variables, table values are Mean (SD); Nonresponses/NAs are removed from individual descriptive statistic calculations.

While many of the participants were undergraduates (42.4%), the mean age was 32.3 years with a standard deviation of 14.3 years. The majority of the subjects' parents had a high school or college education. A majority of the participants (55.5%) reported never being associated with Greek Life but 39.9% were associated with Greek Life at the time of the study.

Health and PA variable descriptive statistics are displayed in **Table 2**. The mean physical health component score among the participants was 55.3, higher than what is expected in the US population. The mean mental health component score was 42.1, which is lower than what is expected in the US population. All domains of PA had a mean duration of MET-minutes per week over 1,000, with the most PA reported in the work domain (M = 2227 MET-min/wk) and the least PA reported in the domestic domain (M = 1030 MET-min/wk). All domains of PA appear to be highly variable and are right-skewed, with many participants reporting very little or even no PA in that domain while a small number of participants report very high levels of PA.

Physical Health

A pruned and simplified regression tree predicting the PCS age-specific mean difference score from PA and demographic variables is displayed in **Figure 3**. It can be seen that the PCS age-specific mean difference score for an individual can be adequately predicted by variables such as leisure PA, age, description, and active transport PA. Having more than 672 self-reported leisure PA MET-minutes per week results in a higher predicted PCS age-specific mean difference score, especially for individuals over the age of 45 years, regardless of any other demographic features.

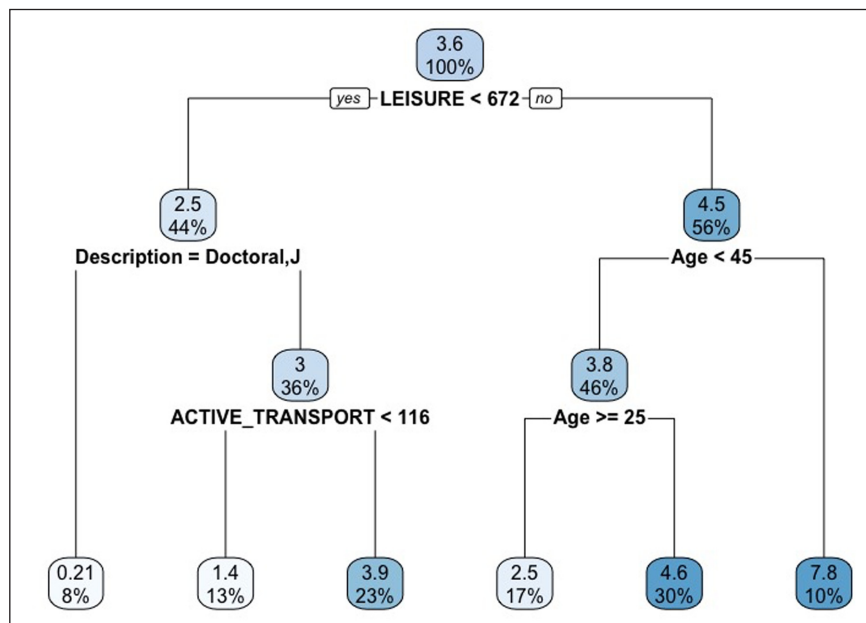


Figure 3: Regression tree predicting the age-specific mean difference score for physical health.

Table 2: Select physical activity measures and reported health of the sample participants.

| Demographic | N = 398* |
|----------------------------------------------|-------------|
| <i>Physical Health Composite Score</i> | 55.3 (5.6) |
| <i>Mental Health Composite Score</i> | 42.1 (8.9) |
| <i>Active/Transport MET Minutes per Week</i> | 1216 (2102) |
| <i>Domestic MET Minutes per Week</i> | 1030 (2129) |
| <i>Leisure MET Minutes per Week</i> | 1704 (3306) |
| <i>Work MET Minutes per Week</i> | 2227 (4056) |

* For Categorical Variables, table values are Count (%); For Continuous Variables, table values are Mean (SD); Nonresponses/NAs are removed from individual descriptive statistic calculations.

Figure 4 displays the mean variable-importance using the RMSE loss-function for predicting age-specific mean differences in physical health based upon 50 permutations. Box plots reveal the distribution of the mean variable-importance across all 50 permutations. As displayed in **Figure 4**, age and PA attributed to leisure were the most important variables when predicting the physical health of an individual, as measured by the PCS age-specific mean difference score.

Bagging implemented with 500 regression trees confirmed the overall variable-importance of age and leisure PA.

Mental Health

A pruned and simplified regression tree predicting the MCS age-specific mean difference score from PA and demographic variables is displayed in **Figure 5**. It can be seen that the MCS age-specific mean difference score for an individual can be adequately predicted by variables such as leisure PA, age, a father’s education

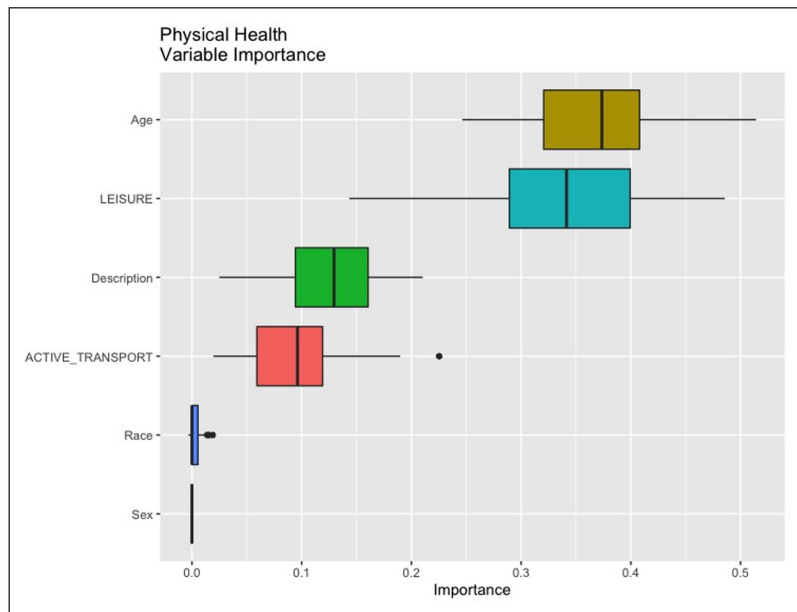


Figure 4: Mean variable-importance using the Root Mean Squared Error (RMSE) loss-function for predicting age-specific mean differences in physical health.

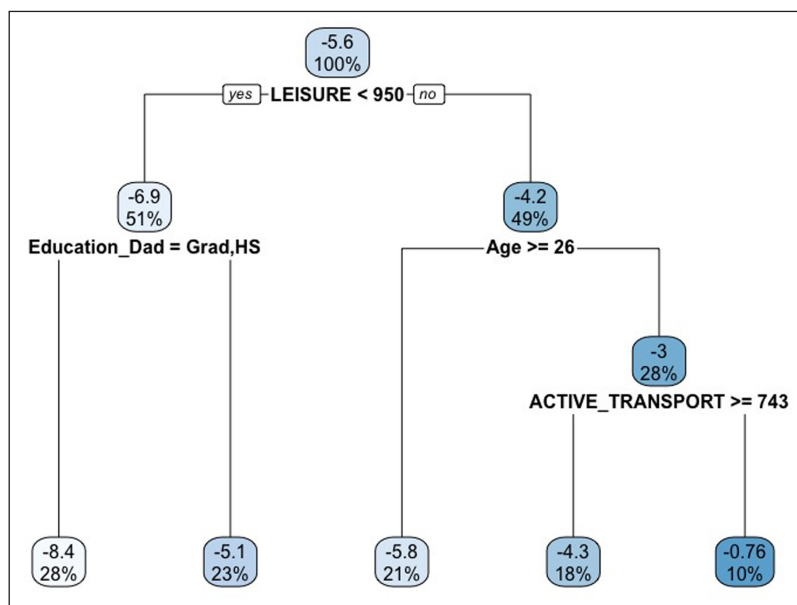


Figure 5: Regression tree predicting the age-specific mean difference score for mental health.

level, and active transport PA. Having more than 950 self-reported leisure PA MET-minutes per week results in a higher predicted MCS age-specific mean difference score, especially for individuals less than 26 years of age and those reporting less active transport PA, regardless of any other demographic features.

Figure 6 displays the mean variable-importance using the RMSE loss-function for predicting age-specific mean differences in mental health based upon 50 permutations. As displayed in **Figure 5**, PA attributed to leisure was the single most important variable when predicting the mental health of an individual, as measured by the MCS age-specific mean difference score, followed by age, father's education level, and active transport PA.

Bagging implemented with 500 regression trees confirmed the overall variable-importance of leisure PA.

Discussion

The purpose of this study was to examine the effects of domain-specific PA on physical and mental health, while illustrating the potential value in CART methods, which are easy to visualize and mirror human decision-making. Results from the regression tree model for physical health indicated that this component of health was most influenced by leisure PA and age, with individuals reporting higher than recommended leisure PA i.e., ≥ 672 MET minutes per week estimated to have greater levels of physical health especially in adults aged 45 years or above. The mental health component of health was most influenced by leisure PA, with improved mental health levels estimated for those that are highly active in this PA domain i.e., ≥ 950 MET minutes per week.

Previous research has consistently shown that higher levels of PA are associated with better physical health outcomes including but not limited to increased life expectancy, improved quality of life, decreased risk for NCDs, and lessened risk for all-cause mortality (DHHS, 2018; Miles, 2007; Moore et al., 2012; WHO, 2018). Moreover, results from previous work suggest that engaging in regular PA improves overall mental health in both adults and adolescents (Biddle & Asare, 2011; DHHS, 2018; Fox, 1999; Paluska & Schwenk, 2000; Penedo & Dahn, 2005; Saxena et al., 2005; Tyson et al., 2010). However, the effects of PA on overall health and wellbeing likely follow a dose-response pattern (White et al., 2017). This dose-response PA curve indicates the health benefits potentially differ owing to population subgroup membership, PA domain, and health component of interest i.e., physical or mental. For instance, a recent systematic review with meta-analysis suggested that, rather than compensating for physical inactivity in one PA domain, higher levels of work PA had a detrimental impact on mortality (Coenen et al., 2018). PA guidelines should likely evolve to provide more targeted recommendations that take into consideration age and PA domain while also setting clear criteria for PA levels needed to achieve desired physical and mental health outcomes.

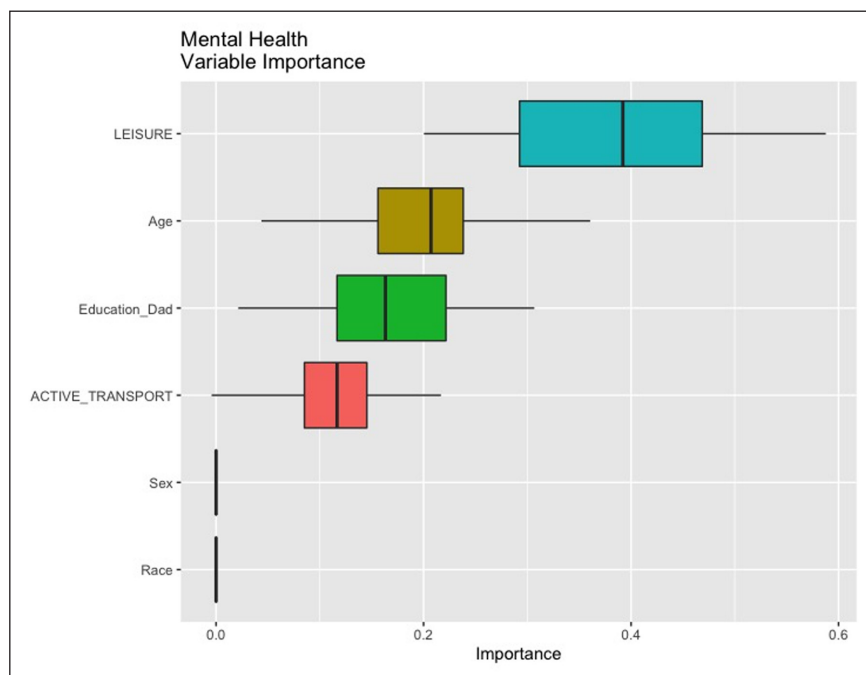


Figure 6: Mean variable-importance using the Root Mean Squared Error (RMSE) loss-function for predicting age-specific mean differences in mental health.

The study is limited currently in scope, as all participants were located at one public university in the United States. Additionally, the measure of PA utilized in this study is a self-report measure, rather than a more objective measure of PA such as might be obtained using accelerometers. While there are limitations, the current study is novel in that it attempts to utilize CART methods to refine PA recommendations while considering domain-specific PA and exploring physical health and mental health outcomes separately. In future work, the use of CART methods to refine PA recommendations could be replicated with a larger, nationally representative sample utilizing objective measures to create a data-informed PA protocol for mental health benefits as well as physical health benefits.

Current results support prior research in physical health, indicating that achieving a modest amount of leisure PA above standard recommendations for overall PA resulted in improved physical health. The results also indicate that age played an important role in the estimation of physical health even when using age-specific mean difference scores to assess physical health. This finding is in alignment with previous results indicating that even a low-dose of moderate-to-vigorous PA improves physical health in older adults (Hupin et al., 2015). The CART results for physical health also support the hypothesis that health-enhancing physical activity may occur in any PA domain (at least for certain population subgroups), as it appears that lower levels of leisure PA can be somewhat compensated for with higher levels of reported transportation PA (see **Figure 3**). This is not the case for mental health, however. Although the results for mental health also indicate that leisure PA plays an important role in the estimation of mental health, this variable emerges as the single most determining factor in mental wellbeing.

Conclusion

Physical activity, especially in leisure, has the potential to greatly improve the physical and mental health of individuals. Especially given the potential increase in the prevalence of mental health disorders stemming from the COVID-19 pandemic, leisure PA should be encouraged further once individuals begin to resume more pre-pandemic behaviours in school, work, and community settings. Additionally, given the substantive improvements in both physical and mental health, public university campuses should continue to encourage leisure PA while also providing greater support for community members to engage in such PA. Additional campus resources should be allocated to support leisure PA in the university setting. For instance, exercise areas could be offered in multiple locations across campus to facilitate accessibility. Additional programming could be provided that connects students, faculty, and staff with accountability groups and/or partners on campus to ensure a community network for PA. These types of investments in supporting and normalizing leisure PA in university settings should be prioritized to better the physical and mental health of students, faculty, and staff.

Competing Interests

The authors have no competing interests to declare.

References

- Alomari, M. A., Khabour, O. F., & Alzoubi, K. H.** (2020). Changes in Physical Activity and Sedentary Behavior Amid Confinement: The BKSQ-COVID-19 Project. *Risk Management and Healthcare Policy, 13*, 1757–1764. DOI: <https://doi.org/10.2147/RMHP.S268320>
- Biddle, S. J., & Asare, M.** (2011). Physical activity and mental health in children and adolescents: a review of reviews. *British Journal of Sports Medicine, 45*(11), 886–895. DOI: <https://doi.org/10.1136/bjsports-2011-090185>
- Breiman, L.** (1984). *Classification and regression trees*. Wadsworth International Group.
- Breiman, L.** (2001). Random forests. *Machine Learning, 45*(1), 5–32. DOI: <https://doi.org/10.1023/A:1010933404324>
- Brownson, R. C., Boehmer, T. K., & Luke, D. A.** (2005). Declining Rates of Physical Activity in the United States: What Are the Contributors? *Annual Review of Public Health, 26*(1), 421–443. DOI: <https://doi.org/10.1146/annurev.publhealth.26.021304.144437>
- Buuren, S. van, & Groothuis-Oudshoorn, K.** (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, 45*(3). DOI: <https://doi.org/10.18637/jss.v045.i03>
- Centers for Disease Control and Prevention.** (n.d.). *DNPAO Data, Trends and Maps: Explore by Location*. Centers for Disease Control and Prevention. https://nccd.cdc.gov/dnpao_dtm/rdPage.aspx?rdReport=DNPAO_DTM.ExploreByLocation


- Cheikh Ismail, L., Osaili, T. M., Mohamad, M. N., Al Marzouqi, A., Jarrar, A. H., Abu Jamous, D. O., ... Al Dhaheri, A. S.** (2020). Eating Habits and Lifestyle during COVID-19 Lockdown in the United Arab Emirates: A Cross-Sectional Study. *Nutrients*, *12*(11), 3314. DOI: <https://doi.org/10.3390/nu12113314>
- Coenen, P., Huysmans, M. A., Holtermann, A., Krause, N., van Mechelen, W., Straker, L. M., & van der Beek, A. J.** (2018). Do highly physically active workers die early? A systematic review with meta-analysis of data from 193 696 participants. *British Journal of Sports Medicine*, *52*(20), 1320–1326. DOI: <https://doi.org/10.1136/bjsports-2017-098540>
- Constant, A., Conserve, D. F., Gallopel-Morvan, K., & Raude, J.** (2020). Socio-Cognitive Factors Associated With Lifestyle Changes in Response to the COVID-19 Epidemic in the General Population: Results From a Cross-Sectional Study in France. *Frontiers in Psychology*, *11*. DOI: <https://doi.org/10.3389/fpsyg.2020.579460>
- Craig, C., Marshall, A. L., Jöström, M., Bauman, A. E., Booth, M. L., Alnsworth, B. E., ... Oja, P.** (2003). International Physical Activity Questionnaire: 12-Country Reliability and Validity. *Medicine & Science in Sports & Exercise*, *35*(8), 1381–1395. DOI: <https://doi.org/10.1249/01.MSS.0000078924.61453.FB>
- Exercise is Medicine (EIM).** (2019). *Exercise is Medicine On Campus Action Guide*. https://www.exerciseismedicine.org/assets/page_documents/EIM%20on%20Campus%20action%20guide.pdf
- Fox, K. R.** (1999). The influence of physical activity on mental well-being. *Public Health Nutrition*, *2*(3a), 411–418. DOI: <https://doi.org/10.1017/S1368980099000567>
- Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., ... Sullivan, M.** (1998). Cross-Validation of Item Selection and Scoring for the SF-12 Health Survey in Nine Countries. *Journal of Clinical Epidemiology*, *51*(11), 1171–1178. DOI: [https://doi.org/10.1016/S0895-4356\(98\)00109-7](https://doi.org/10.1016/S0895-4356(98)00109-7)
- Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C.** (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million participants. *The Lancet Global Health*, *6*(10). DOI: [https://doi.org/10.1016/S2214-109X\(18\)30357-7](https://doi.org/10.1016/S2214-109X(18)30357-7)
- Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C.** (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health*, *4*(1), 23–35. DOI: [https://doi.org/10.1016/S2352-4642\(19\)30323-2](https://doi.org/10.1016/S2352-4642(19)30323-2)
- Hamer, M., Stamatakis, E., & Steptoe, A.** (2008). Dose-response relationship between physical activity and mental health: the Scottish Health Survey. *British Journal of Sports Medicine*, *43*(14), 1111–1114. DOI: <https://doi.org/10.1136/bjism.2008.046243>
- Huber, B. C., Steffen, J., Schlichtiger, J., Graupe, T., Deuster, E., Strouvelle, V. P., ... Brunner, S.** (2020). Alteration of physical activity during COVID-19 pandemic lockdown in young adults. *Journal of Translational Medicine*, *18*(1). DOI: <https://doi.org/10.1186/s12967-020-02591-7>
- Hupin, D., Roche, F., Gremeaux, V., Chatard, J. C., Oriol, M., Gaspoz, J. M., Barthélémy, J. C., & Edouard, P.** (2015). Even a low-dose of moderate-to-vigorous physical activity reduces mortality by 22% in adults aged ≥60 years: a systematic review and meta-analysis. *British journal of sports medicine*, *49*(19), 1262–1267. DOI: <https://doi.org/10.1136/bjsports-2014-094306>
- IPAQ.** (n.d.). *IPAQ scoring protocol – International Physical Activity Questionnaire*. <https://sites.google.com/site/theipaq/scoring-protocol>
- Jenkinson, C., Layte, R., Jenkinson, D., Lawrence, K., Petersen, S., Paice, C., & Stradling, J.** (1997). A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *Journal of Public Health*, *19*(2), 179–186. DOI: <https://doi.org/10.1093/oxfordjournals.pubmed.a024606>
- Katzmarzyk, P. T., & Mason, C.** (2009). The Physical Activity Transition. *Journal of Physical Activity and Health*, *6*(3), 269–280. DOI: <https://doi.org/10.1123/jpah.6.3.269>
- Lundberg, L., Johannesson, M., Isacson, D. G. L., & Borgquist, L.** (1999). The Relationship between Health-state Utilities and the SF-12 in a General Population. *Medical Decision Making*, *19*(2), 128–140. DOI: <https://doi.org/10.1177/0272989X9901900203>
- Miles, L.** (2007). Physical activity and health. *Nutrition Bulletin*, *32*(4), 314–363. DOI: <https://doi.org/10.1111/j.1467-3010.2007.00668.x>
- Moore, S. C., Patel, A. V., Matthews, C. E., Berrington de Gonzalez, A., Park, Y., Katki, H. A., ... Lee, I.-M.** (2012). Leisure Time Physical Activity of Moderate to Vigorous Intensity and Mortality: A Large Pooled Cohort Analysis. *PLoS Medicine*, *9*(11). DOI: <https://doi.org/10.1371/journal.pmed.1001335>
- National Institute of Mental Health.** (n.d.). *Mental Illness*. <https://www.nimh.nih.gov/health/statistics/mental-illness.shtml>
- Paluska, S. A., & Schwenk, T. L.** (2000). Physical Activity and Mental Health. *Sports Medicine*, *29*(3), 167–180. DOI: <https://doi.org/10.2165/00007256-200029030-00003>

- Penedo, F. J., & Dahn, J. R.** (2005). Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Current Opinion in Psychiatry*, *18*(2), 189–193. DOI: <https://doi.org/10.1097/00001504-200503000-00013>
- R Core Team.** (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. URL <http://www.R-project.org/>
- Riebe, D., Ehrman, J. K., Liguori, G., & Magal, M.** (2021). *Acsm's guidelines for exercise testing and prescription*. Langara College.
- Saxena, S., Van Ommeren, M., Tang, K. C., & Armstrong, T. P.** (2005). Mental health benefits of physical activity. *Journal of Mental Health*, *14*(5), 445–451. DOI: <https://doi.org/10.1080/09638230500270776>
- Shah, A. D., Bartlett, J. W., Carpenter, J., Nicholas, O., & Hemingway, H.** (2014). Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. *American Journal of Epidemiology*, *179*(6), 764–774. DOI: <https://doi.org/10.1093/aje/kwt312>
- Stekhoven, D. J., & Buhlmann, P.** (2012). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, *28*(1), 112–118. DOI: <https://doi.org/10.1093/bioinformatics/btr597>
- Strobl, C., Malley, J., & Tutz, G.** (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, *14*(4), 323–348. DOI: <https://doi.org/10.1037/a0016973>
- Tyson, P., Wilson, K., Crone, D., Brailsford, R., & Laws, K.** (2010). Physical activity and mental health in a student population. *Journal of Mental Health*, *19*(6), 492–499. DOI: <https://doi.org/10.3109/09638230902968308>
- U.S. Department of Health and Human Services (DHHS).** (2018). *Physical Activity Guidelines for Americans*, 2nd Edition. https://health.gov/sites/default/files/2019_09/Physical_Activity_Guidelines_2nd_edition.pdf
- Utah Department of Health (UDH).** (2001). *Interpreting the SF-12*. http://health.utah.gov/opha/publications/2001hss/sf12/SF12_Interpreting.pdf
- Ware, J. E., Kosinski, M., & Keller, S. D.** (1996). A 12-Item Short-Form Health Survey. *Medical Care*, *34*(3), 220–233. DOI: <https://doi.org/10.1097/00005650-199603000-00003>
- 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. DOI: <https://doi.org/10.1016/j.amepre.2016.12.008>
- World Health Organization (WHO).** (2016). *Physical activity*. <https://www.who.int/en/news-room/fact-sheets/detail/physical-activity>
- World Health Organization (WHO).** (2018). *Noncommunicable diseases*. <https://www.who.int/en/news-room/fact-sheets/detail/noncommunicable-diseases>

How to cite this article: Robinson, S., & Gilmore, J. (2021). Recursive Partitioning Methods to Examine the Effects of Physical Activity on Physical and Mental Health: A Case Study at One Public University in the United States. *Physical Activity and Health*, *5*(1), pp. 87–97. DOI: <https://doi.org/10.5334/paah.98>

Submitted: 19 April 2021 **Accepted:** 29 April 2021 **Published:** 21 May 2021

Copyright: © 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Physical Activity and Health* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 