A Comparative Risk Assessment of Active Transportation in the United States

Samuel Younkin[1], Jason Vargo[1, 2], ..., James Woodcock[3], Jonathan Patz[1, 2]

1 Global Health Institute, University of Wisconsin–Madison, Madison, WI, USA
2 Nelson Institute for Environmental Studies Center for Sustainability and the Global Environment, University of Wisconsin–Madison, Madison, WI, USA
3 CEDAR

Abstract

Comparative risk assessments to quantify the health potential of active transportation, e.g., walking and cycling, have been performed in the United States on a limited basis. Current methodologies, which rely on extensive data collection and user expertise, have estimated health benefits for select metropolitan areas under a few alternative scenarios. These analyses have largely been conducted using the Integrated Transport Health Impacts Model (ITHIM) which was developed in the UK and has been adapted for use in the US, primarily California. Here we present an implementation of the active transportation component of ITHIM to model the health impacts of modified active transportation levels. Using the newly developed software package *ITHIM* written in R, along with publicly available national datasets, we quantify the change in disability adjusted life years over a wide range of scenarios. The R package implementation of ITHIM's physical activity component provides a robust environment for exploring the relationship between active transportation and positive health outcomes. A web-based user-interface is available for users who want a first estimate for the impact active transportation has on their state/region.

Introduction

Regular exercise has been shown to provide health benefits in terms of chronic disease prevention [1], cognitive function [2], and overall well being and satisfaction with life [3]. Still only half of Americans obtain the recommended levels of aerobic activity of 2.5 hours per week. There is considerable variability in the

attainment of this goal by location within the country. Supportive local built and policy environments are important determinants of population-level activity patterns. Particularly, as they facilitate and foster routine and moderate physical activity, such as walking, local decisions on design and policy can play a huge role in improving the nation's health.

Given the constraints and complexity of local governance and infrastructure investments, land use, and human behavior there is a need for high quality and reliable data and tools for decision-making related to urban transportaiton systems. The lack of accurate and availbale estimates of the health impacts of such decisions has limited the effectiveness of efforts to encourage non-motorized transport in the US and in other countries. In the last decade new tools, applying comparative risk assessment methodologies, have made progress toward quantifying the multple health impacts of changes to the transportation system and behaviors.

ITHIM is one such statistical model that integrates data on active transport, physical activity, fine particulate matter and greenhouse gas emissions to provide estimates for the proportional change in mortality and morbidity for given baseline and alternate travel scenarios. The model has been used to calculate the health impacts of walking and bicycling short distances usually traveled by car or driving low-emission automobiles [4, 5].

ITHIM uses a comparative risk assessment framework for the active transport component of the model. We improve on the existing implementation of the active transport component by including the distribution of non-travel-related activity, improving numerical precision when computing the population attributable fraction and creating a simple user-interface for the custom-built *R* package *ITHIM*. We present this implementation here by using it to estimate the relationship between overall, nationwide disease burden and active

transportation time in the United States.

Materials and Methods

As with the original ITHIM implementation we, employed a comparative risk assessment (CRA) across scenarios defined by mean active transportation time, i.e., combined walking and cycling. To do so we investigate the proportional change in national disease burden between the baseline model and a given alternative scenario, i.e., the population attributable fraction, PAF. The population attributable fraction is defined below and computed for each disease (ICD-10) included in the model, namely Breast Cancer (C50), Colon Cancer (C18-C21), Dementia (G30), Diabetes (E10-E14) and Cardiovascular disease (I10, I12, I15, I20-I25, I30-I31, I33, I40). We also include the disability impacts of depression, though it is not listed as a cause of death in CDC's data. For readability we omit an index indicating disease.

$$\mathrm{PAF}_j = \frac{\int R_j(x)P(x)\,\mathrm{d}x - \int R_j(x)Q(x)\,\mathrm{d}x}{\int R_j(x)P(x)\,\mathrm{d}x} \approx 1 - \frac{\sum_{i=1}^{n} R_j(x_i')}{\sum_{i=1}^{n} R_j(x_i)} = 1 - \delta_j \qquad (1)$$

Here $R(x)$ represents the risk when the individual has exposure x, physical activity (MET-hrs./week). $P(x)$ and $Q(x)$ denote the population density of individuals with exposure x, in the baseline and alternate scenarios, respectively. The exposure variable, x, is modeled as the sum of two independent random variables, travel-related and non-travel-related physical activity. $\mathbf{x}$ and $\mathbf{x}'$ represent the quantiles for the exposure distribution in baseline and alternative scenarios, respectively. δ_j is the proportional burden for disease j in the alternate scenario.

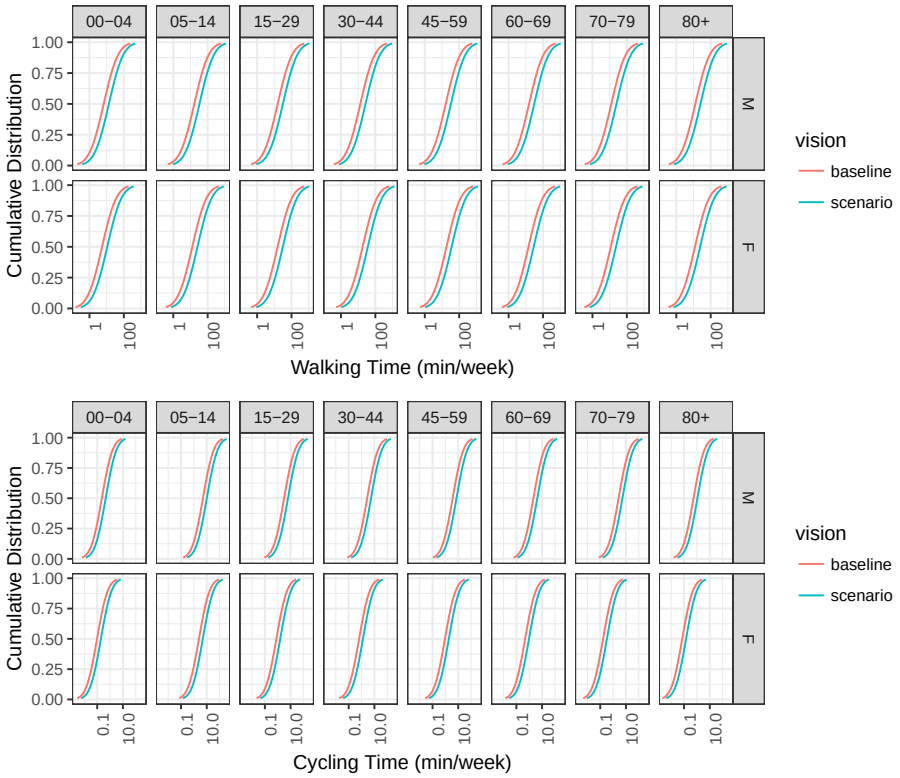

Figure 1: Distributions of walking and cycling transport time for baseline and scenario models, stratified by age and sex. In this example the scenario was created by doubling the overall walking and cycling means.

Travel-Related Physical Activity The distribution for travel-related exposure is estimated from the time spent walking or cycling, i.e., active transport time, which as in the original ITHIM model is assigned a log-normal distribution with constant coefficient of variation across age-sex classes. Travel-related exposure is estimated from active transport time using assumptions about how much physical activity is required for cycling and walking.

Non-Travel-Related Physical Activity With a simulated distribution for total physical activity in hand we compute the empirical percentiles in the baseline and alternate scenarios, $\mathbf{x}$ and $\mathbf{x}'$. The proportional change in disease bur-

den, δ, may be approximated as follows, where R is the risk function for total activity.

$$X_{ij}^{\text{non-travel}} \sim \log \mathcal{N}\left(\mu_{ij}, \gamma \mu_{ij}\right) \tag{2}$$

Within each age-sex class the non-travel-related exposure is modeled with a log-normal distribution. The age-sex specific mean ratios, $r_{ij} = \frac{\mu_{ij}}{\mu_0}$ and γ are estimated from the American Time Use Survey data [6].

Estimation of Disease Burden The disease burden, in this case DALYs, is found for each of the quantiles using equation 3 using B_0, the overall nationwide burden for the baseline. The overall burden is then computed as the sum across quintiles, $B = \sum_{j=1}^{n} B_j$.

$$B_j = \delta B_0 \frac{\text{RR}_j}{\sum_{k=1}^{n} \text{RR}_j} \tag{3}$$

Risk Functions We assign risk based on dose-response curves for each of the diseases included in the model.

$$\text{RR}_j(x) = \text{e}^{-\alpha_j \tilde{x}} \tag{4}$$

where $\tilde{x} = x^k$, $k = \frac{1}{2}$ and r varies by age-sex class and disease. For a table of r values see supplementary material.

Nationwide Health Benefits to Active Transport For the purposes of this implementation we limited our analyses to metropolitan regions of the United States. To idenfity urban populations we used the National Center for Health Statistics 2013 Urban-Rural Classification Scheme for Counties [7]. We eliminated counties in the US with the classifcation of 'nonmetropolitan' (non-core or micropolitan). We used this and similar classification schemes as well as

county and state identifiers in other datsets (NHTS, ATUS, CDC WONDER) to ensure consistency in our estimates for input parameters.

To test the new implementation of ITHIM's active transportation component, we used the US national metro population estimates for active travel time (walking and cycling) as the baseline. We obtained active travel time input parameters for alternate scenarios using states with large sample sizes. From these examples we were able to approximate low and high active travel scenarios. Finally, we applied these active travel times to the US metro population to perform the comparative risk assessment against the baseline US numbers. We then examined the sensitivity of disease burden estimates to input parameters including the active travel time and non-travel activity.

Value of Information

Insert analysis here.

Results

We report total change (averted or increased) in diability adjusted life years (DALYs) as an estimate of disease burden change with alternative scenarios.

To estimate DALYs we use a methodology adopted from previous US implementations of ITHIM which scale Global Budern of Disease Estimates for the US to smaller populations using mortality rate ratios [5]. Age, sex, and casuse specific mortalities were obtained form the CDC WONDER database for US metro counties for the years 2010-2014. Cells with mortality counts less than 10 deaths were suppressed for privacy and imputed (0 to 9) using R's random integer generator.

Our subset of the US population to only metropolitan counties consisted of 1,090 of 3,141 (34.7%) counties, or approximately 272M of 312M (87%) of the

total US population. The age and gender distribution of our subset population was slightly younger (???making this up???) than the national population. Baseline active transport times (walking and cycling) for the US across age and sex groupings was found to be approximately 43 min/week (???). Baseline non-travel activity time was found to be ???X??? (min/week) from the ATUS for US metropolitan populations.

Adapting ITHIM's methods to the R platform enables the rapid and thorough assessment of parameter influence within the model. Another important improvement enabled by the shift to a more computationaly robust platform was to examine methodological assumptions of former implementations. One critical assumption invovles the approximation of ??? used to calculate the attirbutable fraction (see Figure **??**). Previous implementations used quitiles to approximate this value, essentailly the area under the ???? curve. We improved upon those methods to use percentiles or 100 points to estimate the same values. We find our estimates to be much more stable and reliable at all values of active travel and non-travel activity parameters.

As exhibited in Figure **??**, the variability in disease burden benefits varies to a much greater degree when using fewer points in the attirbutable fraction approximation. Importantly, we see greater stability when using the package and percentiles to approximate attirbutable fraction.

Comparing baseline to scenarios with increased active travel produced estimates of up to 360,000??? fewer DALYs (with 300% increase in mean active travel time) for our US population subset. We then varied the mean non-travel activity time and active travel time to observe how this chaged results (**??**). We observe the most dramatic responses in disease burden estimates to changes in active travel time when non-travel activity times are lowest. This result is consistent with the non-linear dose-response curve for physical activity and several

health outcomes.

References

[1] D. E. Warburton, C. W. Nicol, and S. S. Bredin, "Health benefits of physical activity: the evidence," *Canadian medical association journal*, vol. 174, no. 6, pp. 801–809, 2006.

[2] C. H. Hillman, K. I. Erickson, and A. F. Kramer, "Be smart, exercise your heart: exercise effects on brain and cognition," *Nature reviews neuroscience*, vol. 9, no. 1, pp. 58–65, 2008.

[3] J. P. Maher, S. E. Doerksen, S. Elavsky, A. L. Hyde, A. L. Pincus, N. Ram, and D. E. Conroy, "A daily analysis of physical activity and satisfaction with life in emerging adults.," *Health Psychology*, vol. 32, no. 6, p. 647, 2013.

[4] J. Woodcock, M. Givoni, and A. Morgan, "Health impact modelling of active travel visions for england and wales using an integrated transport and health impact modelling tool (ithim)," *PLoS One*, vol. 8, no. 1, 2013.

[5] N. Maizlish, J. Woodcock, S. Co, B. Ostro, A. Fanai, and D. Farley, "Health cobenefits and transportation-related reductions in greenhouse gas emissions in the san francisco bay area.," *American Journal of Public Health*, vol. 103, no. 4, pp. 703–709, 2013.

[6] "American time use survey." `http://www.bls.gov/tus/`. U.S. Bureau of Labor Statistics, 2011.

[7] D. D. Ingram and S. J. Franco, "2013 nchs urban-rural classification scheme for counties.," *Vital and health statistics. Series 2, Data evaluation and methods research*, no. 166, p. 1, 2014.