BMJ Open

Development of a multifactorial prediction model for commute mode choice in 10 983 Finnish public sector employees: a cross-sectional study

Anna Makkonen ,[1] Kia Gluschkoff,[2] Jaakko Airaksinen ,[2] Jaana I Halonen ,[3] Paula Salo,[1] Jenni Ervasti [2]

[1]Department of Psychology and Speech-Language Pathology, University of Turku, Turku, Finland
[2]Finnish Institute of Occupational Health, Helsinki, Finland
[3]Department of Public Health, Finnish Institute for Health and Welfare, Helsinki, Finland

Correspondence to
Anna Makkonen;
anna.makkonen@utu.fi

ABSTRACT

Objective The objective of this study is to examine the feasibility of using survey data to identify factors that predict commute mode choice.

Design The study design is cross-sectional.

Setting Survey data from the Finnish Public Sector study (2020) were used.

Participants 42 574 public sector employees, of whom 10 983 were selected for the final sample. These included employees with 5 km or less commuting distances and those working full-time onsite or partly remotely. The mean age was 46 (SD 11) years, and 84% were women.

Primary outcomes Commute by (1) bike or foot (an active mode) during summer and winter weather and (2) by car (a passive mode) during summer and winter weather.

Methods Using logistic Lasso (least-absolute-shrinkage-and-selection-operator) regression, we developed and tested a prediction model for short commutes of 5 km or less to identify the characteristics of employees most likely to commute actively during summer and winter weather and passively during summer and winter weather.

Results All models had a good predictive ability with a C-index of 0.82, 0.77, 0.72 and 0.71. Cycling and walking during summer weather were predicted by shorter commutes, higher physical activity, lower body mass index (BMI), female sex and higher team psychological safety. Predictors of cycling and walking during winter weather were shorter commute length, higher physical activity, lower BMI and higher age. Commuting by car during summer weather was predicted by longer journey length, higher BMI, lower physical activity, male sex and having children 7–18 years old living at home. Predictors of driving during winter weather were almost identical, but the male sex was replaced by having a spouse.

Conclusions We identified the correlates of active and passive commute choice in different weather conditions with eight variables. This information can be used to develop and target interventions to promote sustainable and healthy commuting modes.

INTRODUCTION

Active commute modes, such as cycling or walking, are simple, cost-effective and sustainable alternatives to driving and are

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The strength of this study was the large sample of employees with a high response rate.
⇒ The data used in this study are from a cohort of public sector employees, and the number of women in the sample was predominant.
⇒ No information was available on the commuting environment other than commute distance, for example, information on cycle routes or parking places at the workplace was lacking.
⇒ The cross-sectional design precludes causal conclusions.

increasingly recognised as an essential source of physical activity.[1] Cycling has become a crucial mode of transport in the context of policy objectives to mitigate climate change caused by carbon dioxide emissions from transport. Despite the various benefits, active commuting remains low in many countries, and private car use still accounts for a significant proportion of commutes.[2]

Finland is committed to reducing domestic transport emissions by at least 50% compared with 2005 levels by 2030 and carbon neutrality by 2035. The aim is to increase the number of walking and cycling journeys by 30% from 2018 to 2030.[3] As a comparison, the leading example for commute cycling, Copenhagen, strives for carbon neutrality by 2025, aiming for a 50% modal share for commute cycling and a 75% combined share of walking, cycling and public transport for all trip purposes.[4] However, at the EU level (including the UK), a large study found no indications that cycling would have been substituted for travel on foot, by public transport, or by car from the 1990s to date.[5] Instead, the distances travelled by car increased by about 10% during the study period.[5]

In Finland, nearly 50% of the short 1–3 km trips are made by private car, and only 11%

by bicycle.[6] For journeys of 5–10 km, the proportion for cycling drops to 6%. While the average commuting distance in Finland in 2021 was 6 km, 73% of the trips were made by private car, and only 11% and 8% of commuting trips were made by bicycle or by foot.[6] For commuting distances of 1–10 km, cycling would be a competitive mode of transportation in terms of speed compared with walking, public transportation and private cars, particularly in city centres.

Previous studies have shown that commute mode decisions are based on many different factors of commuters' personal and journey characteristics and active and passive commuting may be predicted by different factors. In general, commuting length and duration seem to be fundamental determinants of commute mode choice.[7–10] Driving is often considered the fastest way to commute. Men commute by car more often than women in several parts of the world, including the Nordic countries like Finland and Sweden.[6 9 11 12]

Sociodemographic factors explain the differences between cyclists and non-cyclists. Particularly in low or emerging cycling cultures like the USA, the UK or Australia, men commute by bicycle more than women. In contrast, in high-cycling countries like the Netherlands, Denmark and Germany, sex and age distribution is more balanced.[13 14] A Danish study indicates that cyclists are more likely to be men when commuting distance increases (>5 km).[8]

Regarding education and income, the connection to commuting mode choice is often present but generally weaker than other explanatory factors. Higher education and income are associated with active commuting and public transportation use, and, in turn, are associated with a lower likelihood of using a car.[9 11 15 16] In contrast, lower education is usually associated with commuting by car.[7 17 18] According to a review, those choosing a carless, proenvironmental lifestyle are often more educated, employed and have higher income.[16] Low-income usually associates with a higher probability of choosing active commuting modes or public transportation,[9 18] though this choice is often driven by necessity rather than environmental consciousness. The propensity for households to select sustainable travel options over cars diminishes as car ownership rises.[6 19 20] Notably, in high-cycling countries like Germany, car ownership levels do not inhibit cycling as much as in Finland.[14]

Although studies on commuting mode choice have steadily increased in the recent decade,[14] there is still a gap in knowledge on how factors other than socioeconomic characteristics, such as health or psychosocial work environment, contribute to commute mode choice. Variables related to work, teamwork, cooperation and leadership in the organisation are essential, as the behaviour of other people influences the behaviour of an individual. In a tight working community, norms can have a stronger influence than in work communities where coworker relationships are not as close. Identifying individuals likely to commute in a particular way could help in designing effective interventions that encourage switching to a more sustainable and healthier commute mode. We examined the feasibility of using survey data on health, health behaviours, psychosocial work environment and sociodemographic, work and employment characteristics to predict commute mode choices. Prediction models were developed for active and passive commuting mode. We aimed to offer some guidelines for policy-makers aiming to encourage sustainable commuting practices.

MATERIALS AND METHODS

Participants

This study was nested within the Finnish Public Sector (FPS) study. The FPS study has been described in detail previously.[21 22] The study population (n=58 971) included employees of four Finnish cities, with survey data collected from employees during September–October 2020. A total of 42 574 employees (72%) responded. The sample comprised approximately 10% of all public sector employees in Finland in 2020. The majority (84%) of participants were women, which is typical in Finnish public sector. In 2020, the most common occupations in the Finnish municipal sector were those related to healthcare, social services and education, representing nearly 50% of all occupational groups. In the FPS study, filling up the questionnaire is voluntary. The voluntariness is clearly stated in the cover letter of the questionnaire (information to study participants). We also inform the participants that filling out the questionnaire is considered consent to participate. Additionally, we ask for informed consent to link survey data with register data.

Analyses were conducted separately for summer weather and winter weather. Short distance, defined as 5 km or less, was chosen because long commuters cannot be expected to switch similarly from driving to walking or cycling. Thus, employees with a commuting distance of 5 km or less are the core group that could be encouraged to adopt an active commuting mode. Full-time remote workers and part-time workers were excluded because they do not commute to the workplace daily. We performed a multiple imputation with one imputed dataset using the R package Mice (multivariate imputation by chained equations) for the missing predictor data. We used only one imputed dataset because using multiple datasets with LASSO would result in multiple sets of predictors that would differ from each other and could not be readily pooled. The level of missing data before imputation is shown in online supplemental figure 1. The final analytical sample consisted of 10 983 participants.

Measures

The outcome variable, commuting mode, was asked with the following question: "How often do you commute in the summer weather by (1) walking; (2) cycling; (3) public transport with 1000 m or more of walking or cycling; (4) public transport with less than 1000 m of walking or cycling; (5) private car use, either as a driver

or a passenger?" The same question was repeated for winter weather. Categories 1–2, walking and cycling, were considered as active commute modes, and five private car use, as passive commute modes. Categories 3 and 4 were omitted from the analysis, as they cannot clearly be classified into active or passive. Response options were as follows: 1=daily or almost daily; 2=a few times a week; 3=once a week; 4=less than once a week; 5 = never.[22] [23] Option 1=daily or almost daily was chosen as the primary outcome variable because it is the most accurate indicator of the main travel mode. For example, choosing option '2=a few times a week' as the outcome for predicting active commuting could mean that the respondent may drive more than 50% of the week to work. However, we ran the additional analysis for active commuting with an outcome of option '2=a few times a week'.

The predictor variables were the 87 other questions in the FPS survey on sociodemographic characteristics, health status, lifestyle habits, psychosocial work environment factors and leadership. Most items had a five-point response scale from 1=strongly agree to 5=strongly disagree. A detailed description of the variables is provided in online supplemental appendix 1, but a short description of key predictors follows.

Commute length was asked with an open question: 'How long is your one-way commute from home?'

The survey included several questions about sociodemographic characteristics: sex, age, type of job contract, occupational title (ISCO-coded=International Standard Classification of Occupations), job tenure, working time (full or part-time; day work or shift work; years in shift work). Sex and occupation were collected from the employers' registers.

The survey included eight items on the team's psychological safety, for example, 'My work is valued at my workplace', 'Our workplace is supportive' and with reverse coded items: for example, 'Bullying occurs in our workplace', 'People on sick leave are easily labeled as truants'. Job demands consisted of five items and job control of nine items.[24] [25] Effort at work was measured with one question: 'How much do you feel you invest your abilities and resources into your work?' and rewards with three items.[26] Seven items measured worktime control.[27] [28] Job insecurities included five items, and changes at work two items. Experiences of organisational justice were measured with seven items of procedural justice, and six items of relational justice.[29] Supervisor support to employees included four items, and work unit support to supervisor four items. *Worktime control* (seven items) was measured using a questionnaire in which the participants were asked to evaluate how much they could influence the following aspects of their working time: length, starting and ending times, breaks and handling of private matters during the workday, scheduling of work shifts, vacations and paid days off, and the taking of unpaid leave.[27] [28] All these measures of psychosocial work environment were scaled from 1 to 5, with 1 indicating 'very much' and 5 indicating 'very little'.

Participants also reported whether they had had a performance appraisals/career development discussion with their supervisor within the preceding 12 months, and whether they perceived the discussion as useful.

Physical activity was measured with questions on average weekly hours of physical activity or exercise during leisure time or commuting within the previous 12 months, with varying intensity corresponding to walking, brisk walking, jogging and running. The response categories were as follows: <30 min, 1 hour, 2–3 hours and >4 hours. We used the following scales for calculations: <30 min = 15 min, 1 hour=45 min, 2–3 hours=2.5 hours and >4 hours = 5 hours. The time spent on activity at each intensity level in hours per week was multiplied by the average energy expenditure of each activity and expressed in MET. Physical activities evaluated to correspond to walking, brisk walking, jogging and running were given MET values 3.5, 5, 8 and 11 MET, respectively.[30] We treated the variable as a continuous variable so that higher MET hours per day indicated higher physical activity and lower hours less activity.

Self-reported body mass index (weight in kg divided by height in m^2) was dichotomised as less than 25 (non-overweight) and 25 or more (overweight).[31] Body mass index is a measure commonly used as an indicator of overweight and obesity and is a risk factor for various health conditions.

Statistical analysis

We built four models: two models for passive commuting in summer and winter weather and two for active commuting in summer and winter weather. The aim was to identify the individual-level predictors of commuting daily by car (passive commuting) and cycling or walking (active commuting). The reason for combining cycling and walking into a single outcome was that the proportion of daily walkers with commute lengths of at least 2 km in the sample was small (5.7%). In Finland, walking reduces to 15% when the journey distance is more than 3 km.[3] The outcomes for the four models were model (1) daily driving (yes, no); and (2) daily cycling or walking to work (yes, no). These four models were analysed separately for summer and winter weather.

We used bootstrap-enhanced least-absolute-shrinkage-and-selection-operator (lasso) with logistic regression for modelling. The main advantage of the lasso model compared with more conventional regression models is its capability to minimise prediction error and to produce a parsimonious (or sparse) model. Basically, the lasso is standard regression with an l₁-norm penalty. Because of the l₁-penalty, the lasso does variable selection and shrinkage. Lasso forces the sum of the absolute value of regression coefficients to be less than a fixed value dependent on a parameter lambda. When the lambda increases, lasso reduces certain regression coefficients to zero, leaving only the most important predictors for the model.[32]

Before developing the parsimonious prediction models, we split the data with a 75/25 split to training and test datasets stratifying for the outcome, and standardised all predictors for lasso. Then, using bootstrap-enhanced lasso with logistic regression and 10-fold cross-validation, we searched for the optimal lambda value for selecting our predictors.[33] The predictors selected by regular lasso may vary depending on the sample and the correlation strength of the candidate predictors. With bootstrap-enhanced lasso, the final selected predictors are present in a set proportion of the bootstrap replications. We used 100 bootstrap replications, set the threshold for predictor selection to 95% and used a lambda value that was one SE from the optimal value instead of selecting the optimal lambda. This widely used approach allowed us to get a more parsimonious prediction model while retaining reasonable accuracy compared with the model using the optimal value.[34 35]

We then fit a model using the predictors retained from the bootstrap-enhanced lasso model to the test dataset to get CIs for our estimates. Using that model, we compared the predictions against the observed cases, plotted a Receiver Operating Characteristic curve (ROC), and computed the area under the curve (AUC). AUC is equal to Harrell's concordance index (C-index)[36] and has a range from 0.5 (no predictive ability) to 1 (maximum predictive ability). C-index under 0.7 represents poor, 0.7–0.8 good and >0.8 strong discrimination ability.

As a sensitivity analysis, we ran an additional six different models. First, four models for active and passive commuting in summer and winter weather without the physical activity (MET) measure. This was done because the measure's wording did not specify whether the physical activity was related to commuting or leisure time. Second, we ran two models with an outcome of cycling or walking a minimum of a few days a week in summer and winter weather to see whether this affected the results. The tables and figures for all the sensitivity analyses are presented in online supplemental tables 1–4 and figures 2–7.

All analyses were performed using R V.4.1.2 (bolasso (0.1.0) and glmnet (4.1-4)).

Patient and public involvement

Patients or members of the public were not involved in the research's design, conduct, reporting or dissemination plans.

RESULTS

Descriptive characteristics of the study population are shown in table 1. Of the participants, 84% were women, corresponding to the sex distribution in the FPS. The mean age was 46 years (SD 11), and 9% worked in a supervisory position.

Bivariate associations between all items with all four outcomes are presented in figure 1 and figure 2. Most of the statistically significant bivariate associations were

Table 1 Characteristics of the study population (n=10 983)

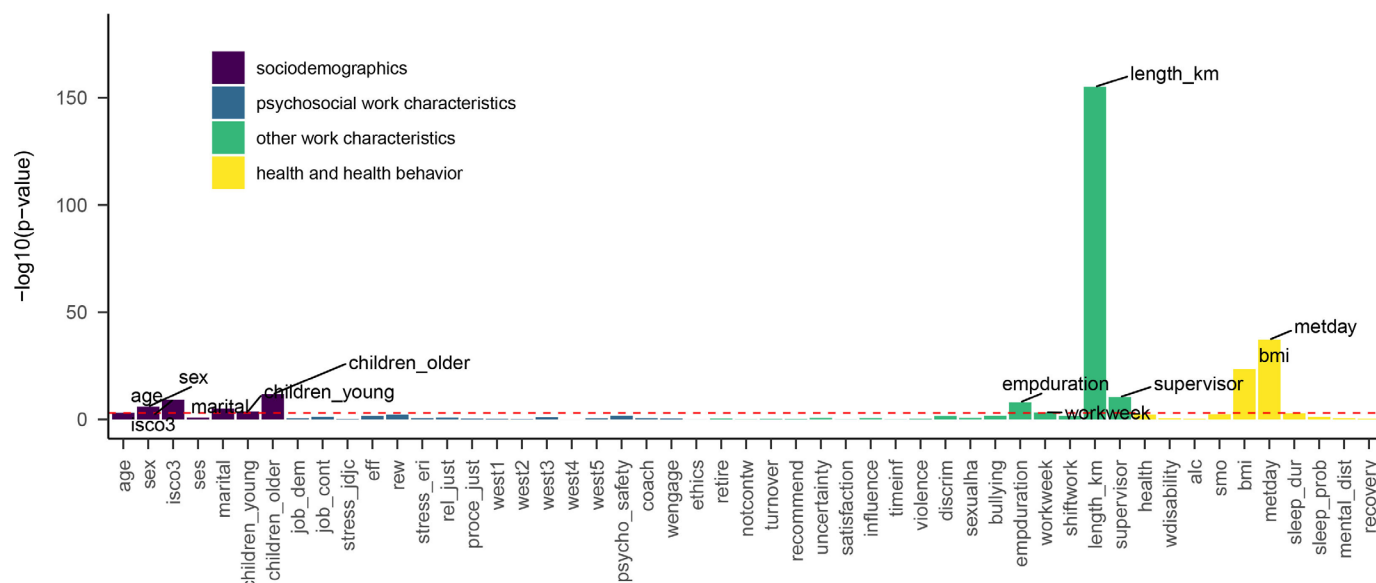| Characteristic	N (%)
Sex %	
Women	9201 (84)
Men	1782 (16)
SES %	
Upper or lower-level non-manual employee	8051 (73)
Manual worker	2932 (27)
Supervisor position, n (%)	
No	9949 (91)
Yes	1034 (9)
Working time n, (%)	
Full-time	8424 (77)
Part-time	2559 (23)
Commute by car, n (%)	
Less frequently	8802 (80)
Daily	2181 (20)
Commute by walk/bike, n (%)	
Less frequently	4844 (44)
Daily	6139 (56)

related to commute length, health and health behaviour, team's psychological safety, efforts put to work and sociodemographic characteristics.

Commuting either passively by car or actively by foot or bicycle both during summer and winter weather was best predicted using 8 variables of the 87 original predictors. These predictors are shown in figure 3 and figure 4. The model for winter weather active commute had a strong discriminative ability, AUC=0.82. A good discriminative ability, AUC=0.71, was observed also for commute by car in winter weather. Also, models for summer weather active and passive commute had good discriminative abilities, AUC=0.72 for commute by car, and 0.77 for commute by bike or on foot. The ROC curves for the four main models are presented in online supplemental figures 8 and 9. The individual-level predictors for commute modes, and their ORs with 95% CIs for all models are shown in table 2 and table 3.

The probability to commute by car in summer weather was higher with longer commute length. Other factors associated with a higher probability of commuting by car were low physical activity, living with children from 7 to 18 years old, male sex and higher body mass index. The probability to walk or cycle to work during summer weather was higher with shorter commute length, higher physical activity, lower body mass index, among women and those reporting higher team psychological safety (table 2).

The probability to commute by car in winter weather was higher with longer commute length, living with children from 7 to 18 years old, lower physical activity, higher body mass index and among participants who were
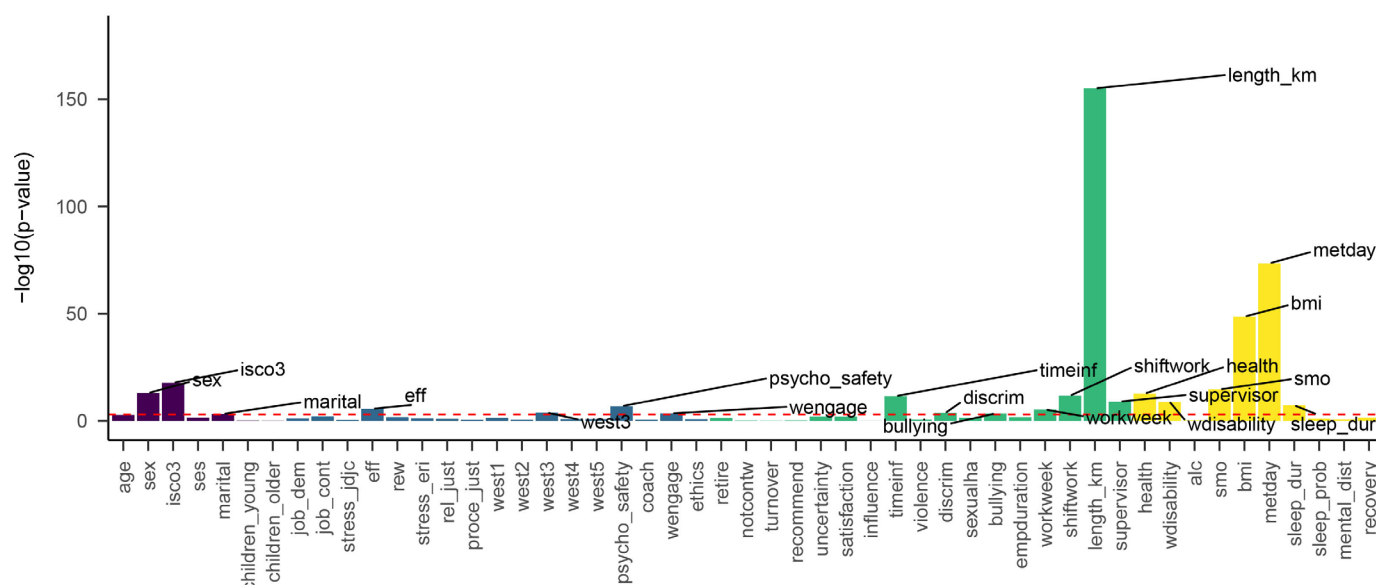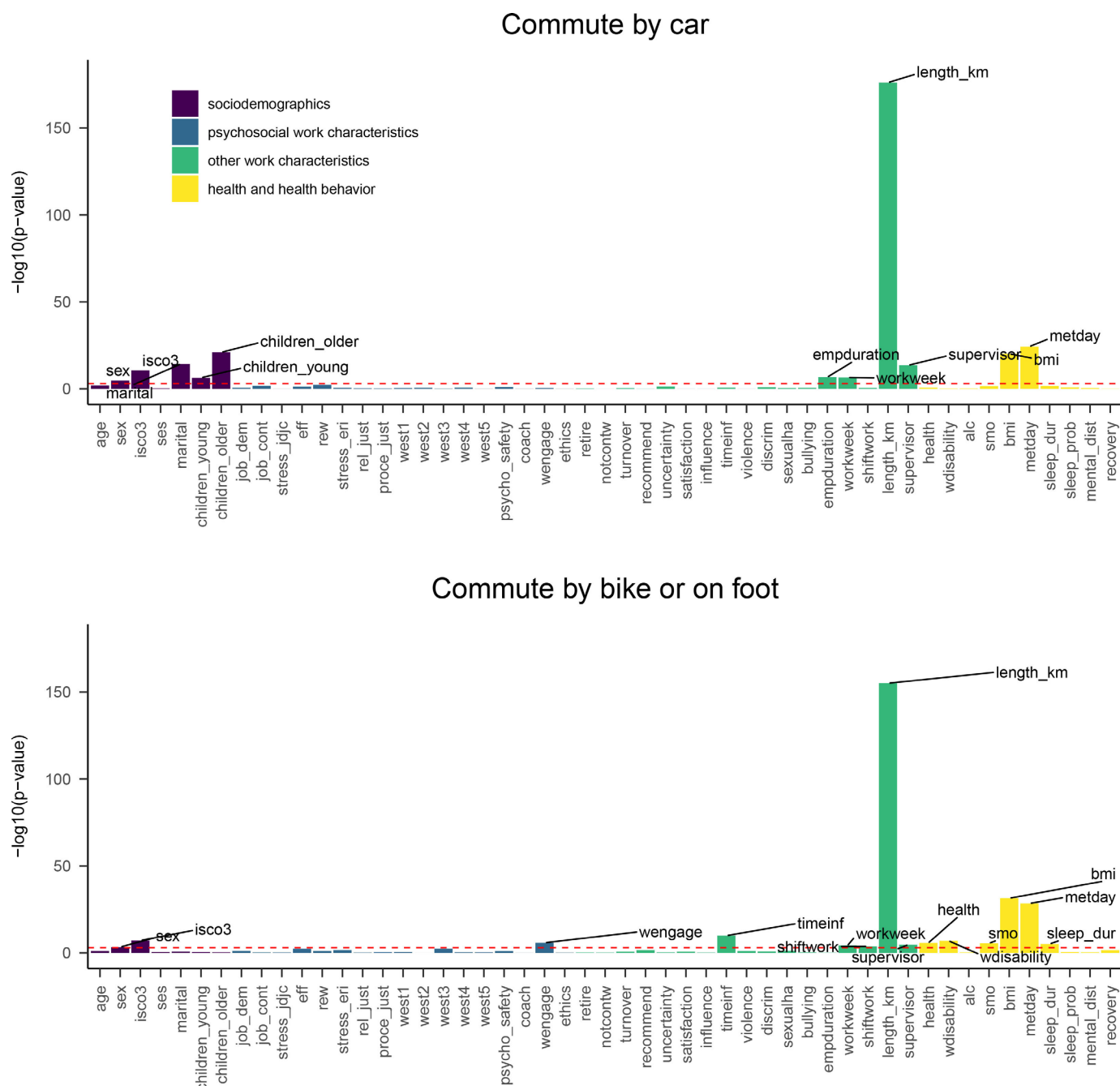
Figure 1 Bivariate associations between predictor items and commute mode choice for passive and active commuting during summer weather. The individual items are grouped with colours to indicate themes. The horizontal red line indicates Bonferroni corrected statistical significance (p=0.05/87=0.0006). Predictors with the highest −log10 (p values) are labelled.

married or cohabiting. The probability to walk or cycle to work during winter weather was higher with shorter commute length, higher physical activity, lower body mass index and higher age (table 3).

A sensitivity analysis for summer weather commuting by car without predictor variable of physical activity resulted in a model with three predictors: longer commute length, higher body mass index and living with children aged 7–18 years. The predictive performance of the more parsimonious model was similar (AUC 0.72). The associations and ROC curves of the sensitivity analyses are presented in the online supplemental table 1 and figure

2. Sensitivity analysis for summer weather commuting by bike or on foot without physical activity variable resulted in a model with six predictors: shorter commute length, lower body mass index, not smoking, being a woman, high effort at work and higher team psychological safety. The model performance was worse (AUC 0.76). The associations and ROC curves are presented in the online supplemental table 1 and figure 3.

Sensitivity analysis with an outcome of a minimum of a few days walking or cycling per week in summer weather produced a model with a similar AUC of 0.77. The model included five predictors: shorter

Commute by car

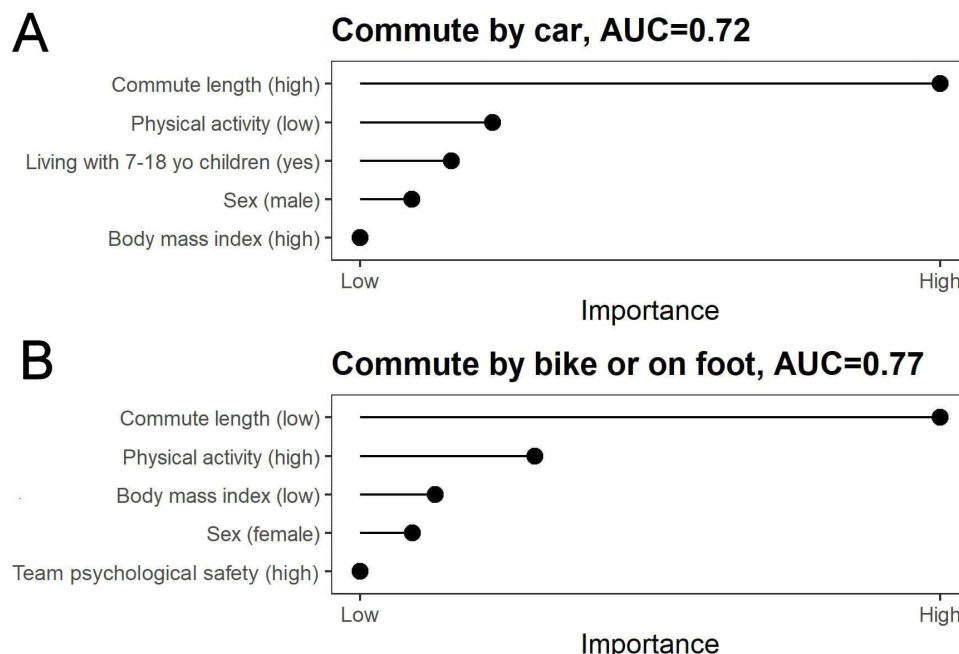## Commute by bike or on foot

Figure 2 Bivariate associations between predictor items and commute mode choice for active commuting in winter weather. The individual items are grouped with colours to indicate themes. The horizontal red line indicates Bonferroni corrected statistical significance (p=0.05/87=0.0006). Predictors with the highest −log10 (p values) are labelled.

commute length, high physical activity, low body mass index, not smoking and marriage or cohabitation (online supplemental table 2 and figure 4). The final sensitivity analysis for winter weather active commuting a minimum of a few days per week produced a model with five predictors: shorter commute length, higher physical activity, lower body mass index, being a non-smoker and having better worktime control. The model performed weaker than the main model (AUC 0.81) (online supplemental table 4 and figure 7).

DISCUSSION

In this study, we developed prediction models for commuting modes on short journeys using a sizeable public sector workplace survey of 10 983 employees. We built four models: two for active commuting modes of walking and cycling during summer and winter weather and another two for passive commuting by car during summer and winter weather. The models were good at identifying characteristics of individuals who are likely to choose either an active or a passive mode for commuting short distances.

Figure 3 The remaining predictor variables in the main Lasso models. (A) Commuting by car and (B) commuting by bike or on foot in summertime.

All models resulted in four or five predictors. Passive commuting by car in summer weather was associated with higher journey length, lower physical activity, living with 7–18 years old children, male sex and higher body mass index. Active commuting in summertime was associated with shorter commute length, higher physical activity, lower body mass index, female sex and better team psychological safety. Passive winter weather commuting was associated with longer commute length, living with 7–18 years old children, lower physical activity, higher body mass index and marriage or cohabitation. Active commuting in wintertime was associated with shorter commute length, higher physical activity, lower body mass index and higher age. The results were largely replicated in multiple sensitivity analyses.
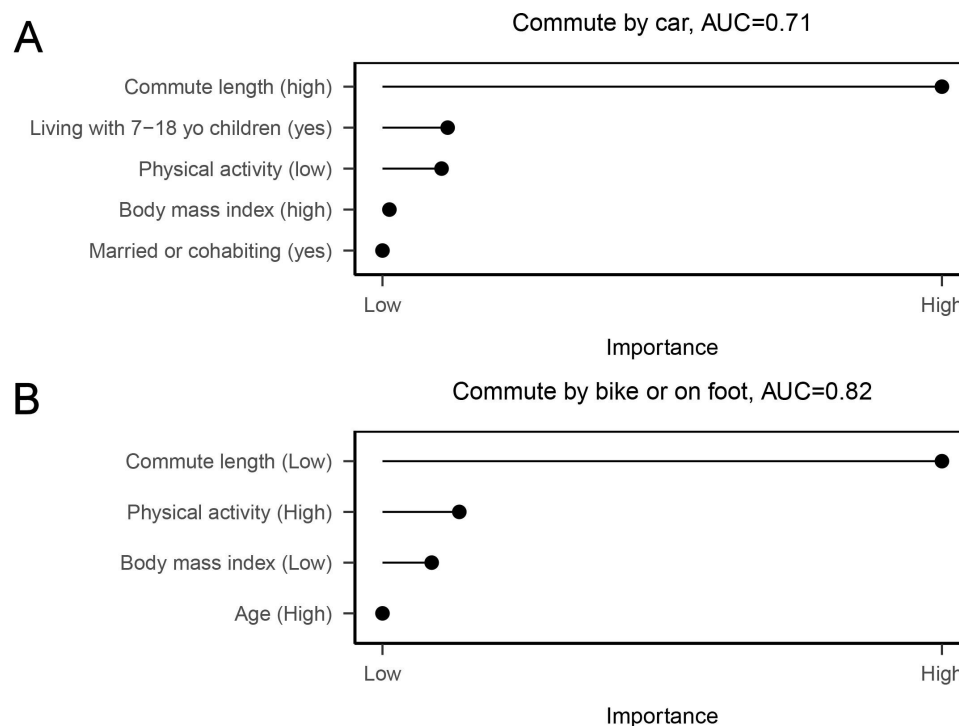
Figure 4 The remaining predictor variables in main Lasso model. (A) Commuting by car and (B) commuting by bike or on foot in the winter weather.

Table 2 The predictors of summer weather commute mode choice

Predictor	Commute by car		Predictor	Commute by bike or on foot	
	OR	95% CI		OR	95% CI
Commute length (km) (high)	3.86	3.41, 4.36	Commute length (km) (low)	0.13	0.12, 0.15
Physical activity (MET) (high)	0.54	0.47, 0.62	Physical activity (MET) (high)	2.20	1.97, 2.46
Living with 7–19 years old children* (yes)	1.58	1.40, 1.77	Body mass index (kg/m^2) (low)	0.64	0.58, 0.71
Sex† (male)	1.63	1.41, 1.89	Sex† (female)	0.60	0.53, 0.69
Body mass index (kg/m^2) (high)	1.34	1.20, 1.50	Team's psychological safety (high)	1.25	1.13, 1.56

The ORs with their 95% CI for continuous variables represent the effect of each 2 SD increase.
*Reference category=no.
†Reference category 1=female.

Previous research on the identified predictors of commute mode

Although the commute length was limited to short commutes of 5 km or less, the strongest predictor of commuting mode choice was length in all four models. This finding is supported by previous studies.[7–10] Shorter commute was associated with walking or cycling, and longer commute with driving a car. While more than half of the study population (56%) cycled or walked every day to work in summer weather, still 20% of the participants drove to work daily despite the short journey, which takes a maximum of 30 min by cycle.

Physical activity and body mass index were linked to both active and passive modes of commuting in all models. Previous studies also demonstrate a robust association between these factors and commute mode.[37] The negative association between body mass index and active commuting is consistent with previous cross-sectional studies.[38–40] A longitudinal study[41] showed that switching from passive commuting to active commuting and maintaining active commuting can help control body weight among working-aged adults of both sexes. As the prevalence of overweight and obesity increases,[40 42 43] the role of active commuting may prove an effective weight control strategy in working-age adults. E-bikes could be a promising alternative, as they lower the barrier to start cycling and help people maintain cycling, contributing to meeting physical activity guidelines. Although e-bikes require less effort than regular bicycles, several studies have found that they still provide moderate-intensity physical activity.[44]

One consistent finding was that family situation was associated with passive commuting. Living with children aged from 7 to 18 was associated with a higher likelihood to commute by car during both summer and winter weather. Living with a partner was associated with a higher likelihood of driving during winter weather. Earlier findings regarding family situations and commuting modes have been varying. Although there are also opposite results from high cycling countries, Germany and Netherlands,[5 45 46] several studies indicate that parenthood, of school-aged children in particular, is associated with increased car use and decreased active travel.[11 17 40 47] This effect is more pronounced in men than women.[48] Possible explanations exist for the association between living with school-aged children and driving. Time constraints and greater necessity for trip chaining, such as dropping off or picking up children from schools or daycare centres, are direct effects of parenthood affecting commute mode choice.[48 49] There are also indirect effects, such as increased car ownership, that may significantly impact commute mode choice.[20 47] Living in a common law or marital relationship can also increase the likelihood of purchasing a car, as the financial situation is often significantly better.

Table 3 The predictors of winter weather commute mode choice

Predictor	Commute by car		Predictor	Commute by bike or on foot	
	OR	95% CI		OR	95% CI
Commute length (km) (high)	4.03	3.59, 4.53	Commute length (km) (low)	0.06	0.53, 0.59
Living with 7–18 years old children* (yes)	1.55	1.39, 1.73	Physical activity (MET) (high)	1.82	1.63, 2.04
Physical activity (MET) (low)	0.64	0.56, 0.72	Body mass index (kg/m^2) (low)	0.60	0.54, 0.68
Body mass index (kg/m^2) (high)	1.35	1.22, 1.51	Age (high)	1.35	1.21, 1.50
Married or cohabiting† (yes)	1.39	1.23, 1.57			

The ORs with their 95% CIs for continuous variables represent the effect of each 2 SD increase.
*Reference category=no.
†Reference category=no.

Owning a car, in turn, can partly explain the connection to choosing passive commute mode.[6 19 20]

Although sex did not emerge in the winter weather prediction models, it explained cycling and driving in the summer weather, with men being more likely to drive and women to cycle. The result agrees with previous research showing that women are less dependent on private cars than men are.[6 9 11 12] In Finland, men drive more and longer journeys than women in all age groups.[6] One explanation may be that men, more often than women, have the identity of a driver. Thus, men use the car in the household. Previous research showed that people with car-related self-image who value travel time, comfort and freedom,[50] or status power,[51 52] are more likely to commute by car.[16]

Our study finds that older employees are more likely to cycle to work during winter. Countries with higher cycling rates tend to have more older cyclists than low-cycling countries.[5 53] Our research cannot address the factors that explain the observed relationships between the variables. However, as parents of young children are living through the busiest years and often need to run errands during their commute, they may, therefore, rely on cars to save time. In contrast, older working-age individuals who have passed the busiest years have less urgency and more time for commuting by bicycle.

We found no previous research on the role of psychosocial or other work characteristics in commute mode choices. Our data included several such measures. One of these measures, namely, team psychological safety, remained in the final prediction model. Higher team psychological safety was associated with active commuting in summer weather. When the work team is perceived as supportive and helpful, people may have lower anxiety and stress levels,[54] indicating a positive mental state, and thus they may be more willing and have more energy to cycle or walk to work. A psychologically safe work environment may foster a healthy work–life balance, motivating employees to take care of their overall well-being, including physical activity. Moreover, a strong sense of community and teamwork often prevails in workplaces with high psychological safety.[54] Employees may feel supported and encouraged by their peers to engage in active commuting, whether through group initiatives or informal encouragement. A previous study suggests that for women, social environment may play a significant role in choosing the commute mode, whereas men perceive bike-friendly facilities (eg, changing rooms, showers, bike storage) as more important determinants.[55] Active commute might also improve perceptions of team psychological safety. In an intervention study, the active travel group indicated more positive organisational behaviour and positive feelings than the passive travel group.[56]

Strengths and limitations

The strengths of this study include a large sample from a cohort representative of FPS employees across various occupations. The FPS is the most extensive survey examining the work and health of public sector staff in Finland. The response rate was high, 72%. Our analytical sample (table 1) represented the survey population well, in which the proportion of women was 78.9 %, and the mean age was 46 years, with a SD of 11. Moreover, considering both summer and winter weather, commute distance and frequency of the outcome measure, we gained more detailed insights into commute mode choices and enhanced the robustness of our main findings. Additionally, to our knowledge, the study included factors associated with commute mode, such as the psychosocial work environment, which have not been previously explored.

However, the study also has limitations. All survey items were self-reported and therefore subject to recall and/or social desirability bias. Furthermore, while we did have more than 1700 men in our sample, most participants were women, which is typical in the FPS. Thus, the generalisability of our findings to male-dominated private sector employees, other environments, societies and cultures remains uncertain. Despite including many predictors, we lacked information on employer-provided parking facilities, commute infrastructure and the environment, which could influence commuting choices. Previous evidence indicates that eliminating free workplace parking effectively reduces car-based commutes, as incentives, convenience and habit encourage car use.[57–59] In the future, it might also be useful to separate walkers and cyclists in the analyses, as the correlates may differ. Lastly, the data collection occurred in September–October 2020, after the outbreak of the COVID-19 pandemic. The pandemic may have led to more people commuting by car or active modes than normally. Nevertheless, our results mainly align with previous findings of predictors of commute mode choice from both pre-pandemic[7 37 41] and post-pandemic[23] pandemic periods.

Conclusions

We created prediction models to identify the characteristics of individuals who are likely to choose a bike and those of a car commuter. Our prediction models, with eight predictors, were accurate at identifying the characteristics of these individuals who were active or passive commuters in summer and winter weather.

Alongside the better-known correlates, such as commute length and physical activity, we discovered a new, previously unidentified psychosocial factor associated with active commuting: psychological safety. Better team psychological safety was linked to daily active commuting during the summer weather. Notably, this factor can be modified with interventions. Workplaces with high psychological safety are likely also committed to the overall well-being of their employees, including healthy work–life balance and promotion of physical activity. Psychological safety may be accompanied by empowerment and autonomy.[54] Employees in control of their work and personal lives may be more inclined to take proactive measures to enhance their health and well-being, including opting for active commuting.

Our models can help inform intervention studies. If our findings can be replicated in other cohorts and are supported by intervention studies, policy-makers and urban planners can justify investments in biking routes and facilities, and employers in workplace environments that support active commuting, ultimately fostering healthier and more sustainable communities.

X Jaana I Halonen @jaana_halonen and Jenni Ervasti @JenniErvasti1

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the Ethical Committee of the Helsinki and Uusimaa hospital district (HUS/1210/2016). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

ORCID iDs
Anna Makkonen http://orcid.org/0000-0001-5615-6547
Jaakko Airaksinen http://orcid.org/0000-0001-7773-7615
Jaana I Halonen http://orcid.org/0000-0003-1142-0388
Jenni Ervasti http://orcid.org/0000-0001-9113-2428

REFERENCES

1 World Health Organization. Promoting walking and cycling, Available: https://www.who.int/activities/promoting-walking-and-cycling
2 Petrunoff N, Rissel C, Wen LM. The effect of active travel interventions conducted in work settings on driving to work: A systematic review. *Journal of Transport & Health* 2016;3:61–76.
3 Jääskeläinen S. Kävelyn ja pyöräilyn edistämisohjelma, 2018. Available: https://julkaisut.valtioneuvosto.fi/handle/10024/160720
4 Publikationsdatabase. CPH 2025 Climate Plan -Roadmap 2021-2025, 2020. Available: https://kk.sites.itera.dk/apps/kk_pub2/index.asp?mode=detalje&id=2062
5 Schepers P, Helbich M, Hagenzieker M, et al. n.d. The development of cycling in European countries since 1990. *Eur J Transp Infrastruct Res*.
6 Traficom. Henkilöliikennetutkimus 2021: Suomalaisten liikkuminen. 2023. Available: https://www.traficom.fi/fi/julkaisut/henkiloliikennetutkimus-2021-suomalaisten-liikkuminen
7 Carse A, Goodman A, Mackett RL, et al. The factors influencing car use in a cycle-friendly city: the case of Cambridge. *J Transp Geogr* 2013;28:67–74.
8 Hansen KB, Nielsen TAS. Exploring characteristics and motives of long distance commuter cyclists. *Transp Policy (Oxf)* 2014;35:57–63.
9 Ko J, Lee S, Byun M. Exploring factors associated with commute mode choice: An application of city-level general social survey data. *Transp Policy (Oxf)* 2019;75:36–46.
10 Fitch DT, Gao Z, Noble L, et al. Effectiveness of Free Bikes and E-Bikes for Commute Mode Shift: The Case of Google's Lending Program. *Front Future Transp* 2022;3:886760.
11 Limtanakool N, Dijst MJ, Schwanen T. The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium- and longer-distance trips. *J Transp Geogr* 2006;14:327–41.
12 Mattisson K, Idris AO, Cromley E, et al. Modelling the association between health indicators and commute mode choice: a cross-sectional study in southern Sweden. *Journal of Transport & Health* 2018;11:110–21.
13 Pucher J, Buehler R. Making Cycling Irresistible: Lessons from The Netherlands, Denmark and Germany. *Transp Rev* 2008;28:495–528.
14 Pucher J, Buehler R. Cycling towards a more sustainable transport future. *Transp Rev* 2017;37:689–94.
15 Clark B, Chatterjee K, Melia S. Changes to commute mode: The role of life events, spatial context and environmental attitude. *Transp Res Part A: Policy Pract* 2016;89:89–105.
16 Soza-Parra J, Cats O. The role of personal motives in determining car ownership and use: a literature review. *Transp Rev* 2024;44:591–611.
17 Rasca S, Saeed N. Exploring the factors influencing the use of public transport by commuters living in networks of small cities and towns. *Trav Behav Soc* 2022;28:249–63.
18 Rachele JN, Kavanagh AM, Badland H, et al. Associations between individual socioeconomic position, neighbourhood disadvantage and transport mode: baseline results from the HABITAT multilevel study. *J Epidemiol Community Health* 2015;69:1217–23.
19 Chng S, White M, Abraham C, et al. Commuting and wellbeing in London: The roles of commute mode and local public transport connectivity. *Prev Med* 2016;88:182–8.
20 Sierra Muñoz J, Duboz L, Pucci P, et al. Why do we rely on cars? Car dependence assessment and dimensions from a systematic literature review. *Eur Transp Res Rev* 2024;16:17–16.
21 Ervasti J, Pentti J, Seppälä P, et al. Prediction of bullying at work: A data-driven analysis of the Finnish public sector cohort study. *Soc Sci Med* 2023;317:S0277-9536(22)00896-6.
22 Kalliolahti E, Aalto V, Salo P, et al. Associations between commute mode use and self-rated health and work ability among Finnish public sector employees. *Scand J Public Health* 2024;52:468–75.
23 Haukka E, Gluschkoff K, Kalliolahti E, et al. Changes in active commuting and changes in health: Within- and between-individual analyses among 16 881 Finnish public sector employees. *Prev Med* 2023;177:107744.
24 Karasek RAJD. Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Adm Sci Q* 1979;24:285.
25 Karasek RA, Theorell T. n.d. Healthy Work: Stress, Productivity and the Reconstruction of Working Life.
26 Kivimäki M, Vahtera J, Elovainio M, et al. Effort-reward imbalance, procedural injustice and relational injustice as psychosocial predictors of health: complementary or redundant models? *Occup Environ Med* 2007;64:659–65.
27 Ala-Mursula L, Vahtera J, Pentti J, et al. Effect of employee worktime control on health: a prospective cohort study. *Occup Environ Med* 2004;61:254–61.
28 Vahtera J, Laine S, Virtanen M, et al. Employee control over working times and risk of cause-specific disability pension: the Finnish Public Sector Study. *Occup Environ Med* 2010;67:479–85.
29 Moorman RH. Relationship between organizational justice and organizational citizenship behaviors: Do fairness perceptions influence employee citizenship? *J Appl Psychol* 1991;76:845–55.
30 Ervasti J, Airaksinen J, Pentti J, et al. Does increasing physical activity reduce the excess risk of work disability among overweight individuals? *Scand J Work Environ Health* 2019;45:376–85.
31 Halonen JI, Kivimäki M, Pentti J, et al. Green and blue areas as predictors of overweight and obesity in an 8-year follow-up study. *Obesity (Silver Spring)* 2014;22:1910–7.
32 Tibshirani R. Regression Shrinkage and Selection via The Lasso: A Retrospective. *J R Stat Soc Ser B* 2011;73:273–82.

33 Bach F. Bolasso: model consistent Lasso estimation through the bootstrap, 2008. Available: http://arxiv.org/abs/0804.1302

34 Chen Y, Yang Y. The One Standard Error Rule for Model Selection: Does It Work? *STATS* 2021;4:868–92.

35 Hastie T. *The elements of statistical learning: data mining, inference, and prediction*. New York: Springer, 2009.

36 Harrell FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis. 2015.

37 Henriques-Neto D, Peralta M, Garradas S, *et al*. Active Commuting and Physical Fitness: A Systematic Review. *Int J Environ Res Public Health* 2020;17:2721.

38 Flint E, Cummins S. Active commuting and obesity in mid-life: cross-sectional, observational evidence from UK Biobank. *Lancet Diabetes Endocrinol* 2016;4:420–35.

39 Flint E, Cummins S, Sacker A. Associations between active commuting, body fat, and body mass index: population based, cross sectional study in the United Kingdom. *BMJ* 2014;349:g4887.

40 Wojan TR, Hamrick KS. Can Walking or Biking to Work Really Make a Difference? Compact Development, Observed Commuter Choice and Body Mass Index. *PLoS ONE* 2015;10:e0130903.

41 Kuwahara K, Noma H, Nakagawa T, *et al*. Association of changes in commute mode with body mass index and visceral adiposity: a longitudinal study. *Int J Behav Nutr Phys Act* 2019;16:101.

42 de Haas M, Kroesen M, Chorus C, *et al*. Causal relations between body-mass index, self-rated health and active travel: An empirical study based on longitudinal data. *J Transp Health* 2021;22:101113.

43 Pineda E, Sanchez-Romero LM, Brown M, *et al*. Forecasting Future Trends in Obesity across Europe: The Value of Improving Surveillance. *Obes Facts* 2018;11:360–71.

44 Bourne JE, Sauchelli S, Perry R, *et al*. Health benefits of electrically-assisted cycling: a systematic review. *Int J Behav Nutr Phys Act* 2018;15:116.

45 Ingvardson JB, Thorhauge M, Kaplan S, *et al*. Incorporating psychological needs in commute mode choice modelling: a hybrid choice framework. *Transportation (Amst)* 2022;49:1861–89.

46 Heinen E. Identity and travel behaviour: A cross-sectional study on commute mode choice and intention to change. *Transp Res Part F: Traffic Psychol Behav* 2016;43:238–53.

47 Oakil ATM, Manting D, Nijland H. Dynamics in car ownership: the role of entry into parenthood. *Eur J Transp Infrastruct Res* 2016;16.

48 Chakrabarti S, Joh K. The effect of parenthood on travel behavior: Evidence from the California Household Travel Survey. *Transp Res Part A: Policy Pract* 2019;120:101–15.

49 Motte-Baumvol B, Bonin O, Belton-Chevallier L. Who escort children: mum or dad? Exploring gender differences in escorting mobility among parisian dual-earner couples. *Transportation (Amst)* 2017;44:139–57.

50 Steg L. Car use: lust and must. Instrumental, symbolic and affective motives for car use. *Transp Res Part A: Policy Pract* 2005;39:147–62.

51 Paulssen M, Temme D, Vij A, *et al*. Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation (Amst)* 2014;41:873–88.

52 Steg L, Vlek C, Slotegraaf G. Instrumental-reasoned and symbolic-affective motives for using a motor car. *Transp Res Part F: Traffic Psychol Behav* 2001;4:151–69.

53 Goel R, Goodman A, Aldred R, *et al*. Cycling behaviour in 17 countries across 6 continents: levels of cycling, who cycles, for what purpose, and how far? *Transp Rev* 2022;42:58–81.

54 Newman A, Donohue R, Eva N. Psychological safety: A systematic review of the literature. *Hum Resour Manag Rev* 2017;27:521–35.

55 Patterson R, Ogilvie D, Panter J. The social and physical workplace environment and commute mode: A natural experimental study. *Prev Med Rep* 2020;20:101260.

56 Page NC, Nilsson VO. Active Commuting: Workplace Health Promotion for Improved Employee Well-Being and Organizational Behavior. *Front Psychol* 2016;7:1994.

57 Julsrud TE, Hjorthol R. Commuting in knowledge intensive organizations: An outline of six different practices. *Int J Sustainable Transp* 2021;15:943–57.

58 Khordagui N. Parking prices and the decision to drive to work: Evidence from California. *Transp Res Part A: Policy Pract* 2019;130:479–95.

59 Christiansen P, Engebretsen Ø, Fearnley N, *et al*. Parking facilities and the built environment: Impacts on travel behaviour. *Transp Res Part A: Policy Pract* 2017;95:198–206.

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

PREDICTION OF COMMUTE MODE CHOICE

Web appendix 1.

Development of a Multifactorial Prediction Model for Commute Mode Choice in 10,983 Finnish Public Sector Employees: A Cross-sectional Study

Detailed description of variables included in the prediction models

Sociodemographic variables

We had information on the respondents' *Age* (as a continuous variable), *sex* (man, woman), and occupation collected from the employers' registers. We also had information on the type of job contract (temporary/permanent) and occupational title (ISCO-coded), Job tenure (years with employer and years in the current position), and working time (full or part-time; day or shift work; years in shift work).

Marital status (having a partner; yes/no), *children aged 0-6 years* (yes/no), *children aged 7-18 years* (yes/no), and *individual commute length* (as a continuous variable; one-way commute distance in kilometers) were self-reported.

Work unit level variables were created as: average age; percentage of women; percentage of temporary employees; percentage of managers, senior officials, and professionals (ISCO codes 1-2); percentage of manual workers (ISCO codes 5-9); turnover percentage (share of new employees as compared to previous survey); size of the work unit; average job tenure; average years in current job; percentage of part-time workers, the percentage in day job; average years in shift work.

Work characteristics

1

PREDICTION OF COMMUTE MODE CHOICE

The survey included measures of *job demands and job control*[1,2]. Job demands were measured with 5 items: time pressures and deadlines, lack of time to do what was expected, and work overload. The job control scale combines two concepts: skill discretion (the opportunities of an individual to develop his or her special abilities within the job, 6 items) and decision authority (individual's abilities to be part of the decision-making process within the organization, 3 items).

For the *Team's psychological safety,* the survey included eight items "My work is valued at my workplace", "Our workplace is supportive", "people at our workplace can really be trusted", "there is an open camaraderie at our workplace characterized by mutual helpfulness", and with reverse coded items: "Bullying occurs in our workplace", "gossip and jealousy occur at our workplace", "I face discrimination at our workplace", " people on sick leave are easily labeled as truants".[3]

Efforts and rewards at work (4 items): Effort was measured with one question ("How much do you feel you invest in your job in terms of skill and energy?") and reward was assessed with three questions about feelings of getting a return from work in terms of (1) income and job benefits, (2) recognition and prestige, and (3) personal satisfaction. The response format was a five-point scale from 1 =very much to 5=very little.[4] The ratio of the self-assessed effort score to the mean of the self-assessed reward scores formed the *effort-reward imbalance* (ERI). Scores above the median was defined as high ERI.

Worktime control (7 items) was measured using a questionnaire in which the participants were asked to evaluate on a scale from 1 (very much) to 5 (very little) how much they could influence the following aspects of their working time: length, starting and ending

2

PREDICTION OF COMMUTE MODE CHOICE

times, breaks, and handling of private matters during the workday, scheduling of work shifts, vacations and paid days off, and the taking of unpaid leave.[5,6]

Job insecurities were measured with 5 items: 'Does your job involve the following insecurities': the threat that some work tasks will be terminated; involuntary transfer into another work tasks; the threat of temporary lay-offs; the threat of permanent lay-offs; the threat of excessive workloads. The scale was from 1=very much to 5=very little.

Changes at work were measured with 2 items: 'When you think about all the changes that have happened in your work during the last 12 months, how would you describe those from your own point of view?' with a scale from 1=mostly positive to 7=mostly negative. We also enquired whether the respondent felt she/he could take part when changes were planned with a response scale 1=I have very much influence over the changes; 2=I have some influence over the changes; 3=Most often the changes occur unexpectedly, I don't have influence over them.

Leadership and management

Procedural justice (7 items): The scale considers whether the decision-making procedures at the workplace are accurate, correctable, consistently applied, and whether the procedures include opinions from the people involved.[7]

Relational justice (6 items): The scale includes items evaluating whether the supervisors use kindness and consideration, are truthful, and can suppress personal biases.[7]

Supervisor support (4 items): The scale includes items evaluating the extent to which the supervisor supports and encourages, rewards from good performance, trusts, and encourages employees to educate and develop themselves in their work.

PREDICTION OF COMMUTE MODE CHOICE

In *support from the work unit to supervisor* (4 items), participants were asked to evaluate the extent to which employees perceive employees to have a role in successful leadership, inform their supervisor on work-related matters, value the competence of their supervisor, and support their supervisor.

Performance appraisals/career development discussions were measured with two items: having had such a discussion within the last 12 months (1=No, 2=Yes), and whether the discussion was perceived useful (1=Useful, 2=Not useful, but not totally useless, 3=Useless).

Work unit/team climate

Team climate (14 items): The work unit cooperation and interaction was measured using the short version[8] of the Team Climate Inventory (TCI)[9]. TCI conceptualizes team climate into four dimensions: participations safety (4 items), support for innovation (3 items), vision (4 items), and task orientation (3 items).

Discrimination at work was measured with a single item: Is there discrimination due to age, gender, education, opinion, status, origins, language, religion, believes/convictions, political activity, trade union activity, health, disability, sexual orientation, or gender identity/gender expression? (1=No, 2=Yes).

Satisfaction with job, satisfaction with employer, and intentions to leave

Job satisfaction was measured with 5 items: Employees were asked to evaluate the extent to which they were satisfied with their personal growth and development; the feeling of accomplishing something significant; possibilities to think and act independently at work; challenges provided by the job. The scale was from 1=very unsatisfied to 5=very satisfied. In

4

PREDICTION OF COMMUTE MODE CHOICE

addition, we asked whether the respondent would recommend the current employer to a friend, with a scale dichotomized into 1=Yes; 2=No.

Retirement intentions were measured with a single item: "Do you see yourself working until your personal retirement age?" with response scale 1=Yes, I see myself working until my retirement age; 2=Yes, I see myself working even after my retirement age; 3=No, I don't see myself working until retirement age.

Turnover intentions were measured with a single item: "Have you considered changing employer?" with response scale 1=No, I want to keep working for my current employer; 2=Yes, I have considered changing employer; 3=I have already recruited to another employer.

Health and lifestyle

Psychological distress (symptoms of depression and anxiety) was measured with the 12-item General Health Questionnaire, where respondents rate the extent to which they are affected by 12 symptoms of distress. Four or more symptoms were coded as cases.[10]

Sleep problems were measured with the Finnish version of the Jenkins Sleep Scale, which comprises four items: 1) the difficulty to fall asleep, 2) wake up at night, 3) difficulty to stay asleep, and 4) non-restorative sleep. The response format was a five-point scale, ranging from 0 (never) to 5 (almost every night), with higher scores indicating more sleep problems.[11]

Perceived health was a single-item measure "How do you rate your health?" with response options from 1=poor to 5=good. The question is widely used in surveys and recommended for standard indicator of health.[12]

PREDICTION OF COMMUTE MODE CHOICE

Self-rated work ability was a single-item measure from the Work Ability Index[13,14]: "Let's assume that your work ability at its all-time best would be given 10 points, and 0 points would indicate that you are completely unable to work. How would you score your current work ability?" Response options ranged from 0 to 10. This single item and the entire Work Ability Index are very strongly associated and showed similar associations with sickness absence health, and symptoms.[15]

Self-reported Body mass index (weight in kg divided by height in m^2) was dichotomized as less than 25 (non-overweight) and 25 or more (overweight).[16] Body mass index is a measure of body fat, commonly used as an indicator of overall health and risk for various health conditions.

Alcohol use was defined through questions on weekly consumption. One drink was approximately equivalent to one unit or one glass of alcoholic drink or 12 g of alcohol. Alcohol use was dichotomized into no use or moderate use (max of 140 g or 11 units for women and 280 g or 23 units for men) versus alcohol use greater than this.[17]

Smoking was dichotomized into current smoker and non-smoker (including never and ex-smokers).[18]

Total Physical activity was measured with questions on average weekly hours of physical activity or exercise during leisure time or commuting within the previous 12 months, with varying intensity corresponding to walking, brisk walking, jogging, and running. The response categories were: < 30 minutes, 1 hour, 2–3 hours, and > 4 hours. We used the following scales for calculations: <30 minutes = 15 minutes, 1 hour = 45 minutes, 2–3 hours = 2.5 hours, and >4 hours = 5 hours. The time spent on activity at each intensity level in hours per week was multiplied by the average energy expenditure of each activity and expressed in MET. Physical

PREDICTION OF COMMUTE MODE CHOICE

activities evaluated to correspond to walking, brisk walking, jogging, and running, were given

MET values 3.5, 5, 8, and 11 MET, respectively.[19]

PREDICTION OF COMMUTE MODE CHOICE
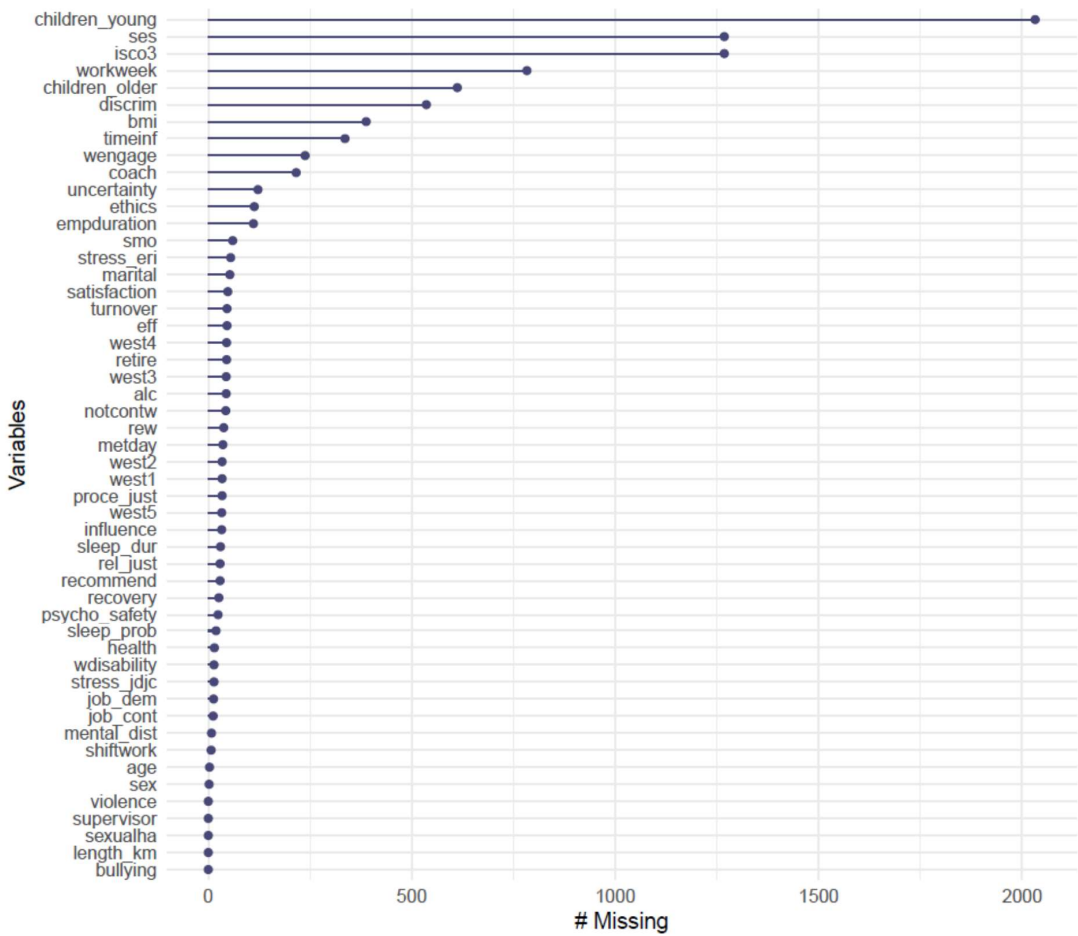
Web figure 1 The level of missing data before imputation

PREDICTION OF COMMUTE MODE CHOICE

Web Table 1 Predictors of commuting by car and by bike or on foot in summer weather, excluding physical activity. Lasso regression. The odds ratios with their 95% confidence intervals (CI) for continuous variables represent the effect of each two standard deviation increase.

	COMMUTE BY CAR			COMMUTE BY BIKE OR ON FOOT	
Predictor	OR	95% CI	Predictor	OR	95% CI
Commute length (high) km	3.79	3.36,4.28	Commute length (low)	0.14	0.13,0.16
Body mass index (kg/m^2) (high)	1.54	1.38,1.71	Body mass index (kg/m^2) (low)	0.55	0.49,0.60
Living with 7-18-year-old children[1] (yes)	1.52	1.36,1.71	Smoking[2] (no)	0.65	0.56,0.76
			Sex[3] (female)	0.73	0.64,0.84
			Efforts (high)	1.20	1.09,1.33
			Team psychological safety (high)	1.20	1.09,1.32

[1]Reference category=no, [2]reference category=no, [3]Reference category=female

PREDICTION OF COMMUTE MODE CHOICE

Web Table 2 Predictors of commuting by bike or on foot for a minimum of a few days per week in summer weather. Lasso regression. The odds ratios for continuous variables with their 95% confidence intervals (CI) represent the effect of each two standard deviation increase.

	COMMUTE BY BIKE OR ON FOOT	
Predictor	OR	95% CI
Commute length (km) (low)	0.12	0.11,0.13
Physical activity (MET) (high)	2.29	2.02,2.60
Body mass index (kg/m^2) (low)	0.66	0.60,0.74
Smoking[1] (no)	0.67	0.57,0.78
Married or cohabiting[2] (yes)	1.25	1.12,1.39

[1]Reference category=no, [2]reference category=yes

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

BMJ Open

PREDICTION OF COMMUTE MODE CHOICE

Web figure 2. Model performance of commuting by car without physical activity in summer weather. Panel A: ROC-curve for prediction of commuting by car (95% threshold for predictor selection). Panel B: remaining predictors in the model.

11

PREDICTION OF COMMUTE MODE CHOICE

Web figure 3. Model performance without physical activity in summer weather. Panel A: ROC-curve for prediction of commuting by walk or cycle (95% threshold for predictor selection). Panel B: remaining predictors in the model.
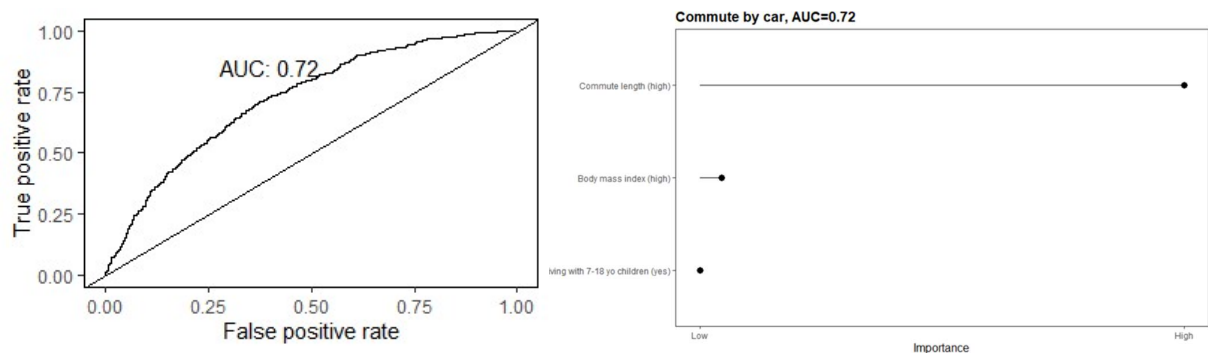
PREDICTION OF COMMUTE MODE CHOICE
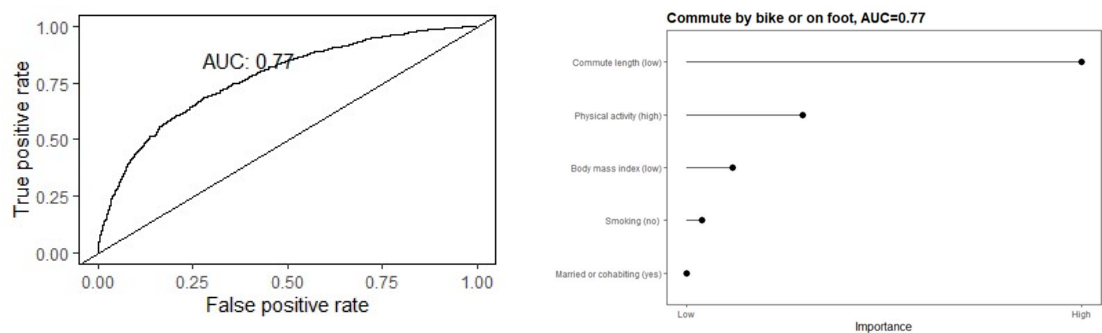
Web figure 4. Model performance for commuting by bike or on foot for a minimum of a few days per week in summer weather. Panel A: ROC-curve for prediction of commuting by bike or on foot (95% threshold for predictor selection). Panel B: remaining predictors in the model.

13

PREDICTION OF COMMUTE MODE CHOICE

Web Table 3. Predictors of commuting by car and by bike or on foot in winter weather, excluding physical activity. Lasso regression. The odds ratios with their 95% confidence intervals (CI) for continuous variables represent the effect of each two standard deviation increase.

COMMUTE BY CAR			COMMUTE BY BIKE OR ON FOOT		
Predictor	OR	95% CI	Predictor	OR	95% CI
Commute length (km) (high)	´3.95	´3.52, 4.43	Commute length (km) (low)	0.06	0.05, 0.07
Body mass index (kg/m2) (high)	´1.49	´1.34, 1.65	Body mass index (kg/m2) (low)	´0.54	0.49, 0.61
Living with 7−18 yo children[1] (yes)	´1.51	´1.35, 1.68			
Married or cohabiting[2] (yes)	´1.38	´1.23, 1.56			

[1]Reference category=no, [2]Reference category=no

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

PREDICTION OF COMMUTE MODE CHOICE

Web table 4. Predictors of active commuting by bike or on foot in winter weather for a minimum of a few days per week. Lasso regression. The odds ratios with their 95% confidence intervals (CI) for continuous variables represent the effect of each standard deviation increase.

COMMUTE BY BIKE OR ON FOOT

Predictor	OR	95% CI
Lenght (km) (low)	0.07	0.06, 0.08
Physical activity (MET) (high)	´1.96	1.76, 2.20
Body mass index(kg/m2) (low)	´0.63	0.57, 0.71
Smoking[1] (no)	´0.61	0.52, 0.72
Worktime control (high)	´1.27	1.17, 1.41

[1]Reference category=no

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

PREDICTION OF COMMUTE MODE CHOICE
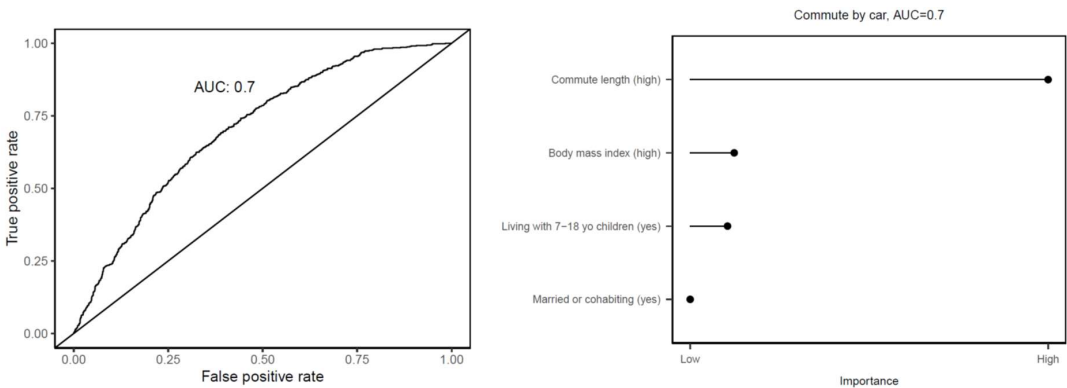
Web figure 5. Model performance of commuting by car without physical activity in winter weather. Panel A: ROC-curve for prediction of commuting by car (95% threshold for predictor selection). Panel B: remaining predictors in the model.

16

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

PREDICTION OF COMMUTE MODE CHOICE
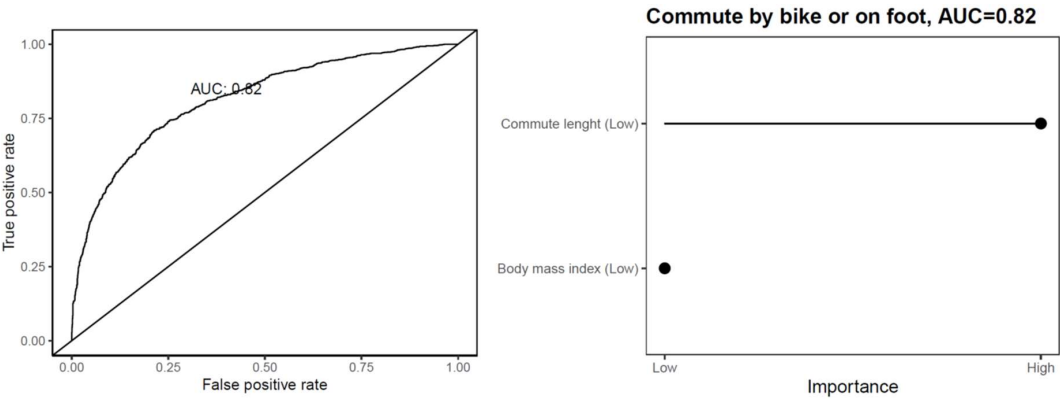
Web figure 6. Model performance for commuting by bike or on foot in the winter weather without physical activity. Panel A: ROC-curve for predicting commuting by bike or on foot in winter weather (95% threshold for predictor selection). Panel B: remaining predictors in the model.

17

PREDICTION OF COMMUTE MODE CHOICE
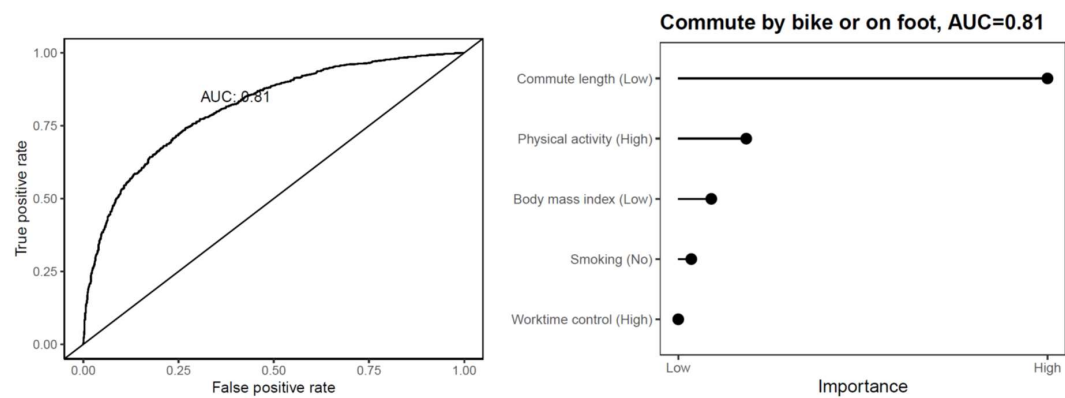
Web figure 7. Model performance for commuting by bike or on foot in the winter weather for a minimum of a few days per week. Panel A: ROC-curve for predicting commuting by bike or on foot (95% threshold for predictor selection). Panel B: remaining predictors in the model.

18

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

PREDICTION OF COMMUTE MODE CHOICE
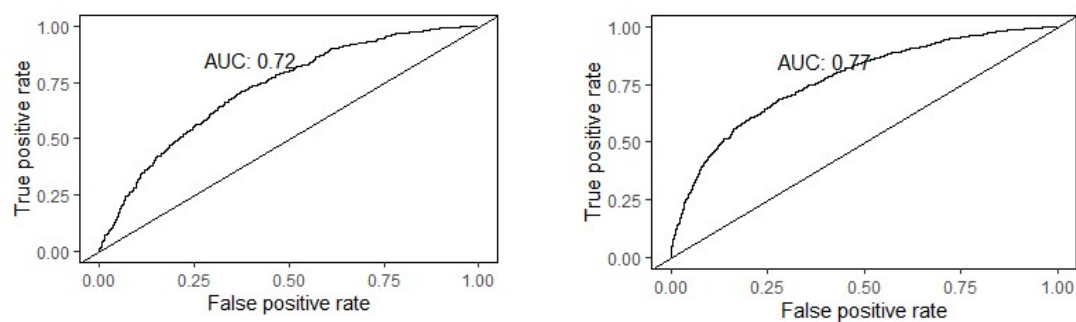
Web figure 8. The model performance for the main models in summer weather. Panel A ROC-curve for the model predicting commuting by car and panel B ROC-curve for commuting by bike or on foot.
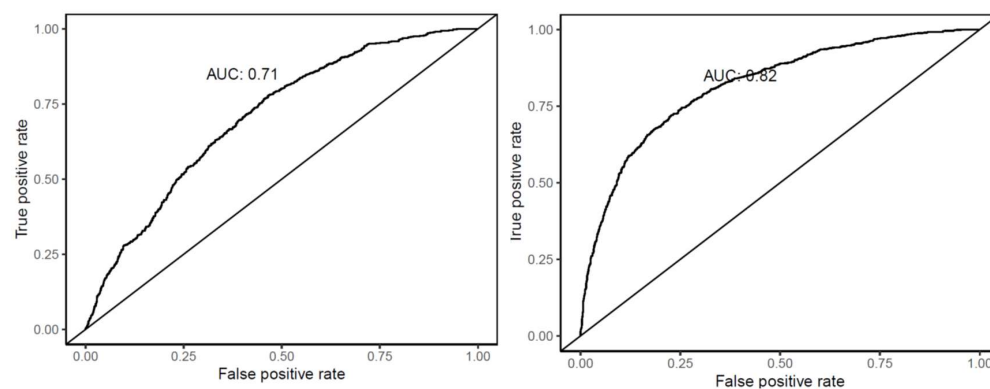
19

PREDICTION OF COMMUTE MODE CHOICE

Web figure 9. The model performance for the main models in winter weather. Panel A ROC-curve for the model predicting commuting by car and panel B ROC-curve for commuting by bike or on foot.

20

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

PREDICTION OF COMMUTE MODE CHOICE

Example analytic code used in the study

R-code for precting active commuting:

Code for walking and cycling daily during summer time (walkbikekes=walking or cycling during summer)

```
#source(here("scripts", "0 wrangling.R"))

dyhtkeskp <- dyhtimp %>%

select(-carkes,-cartal, -nocarkes,-walkbiketal,

    -walkbikekes,) %>% # optional outcomes

 #journey distance:

 #filter(length_km<=5) %>%

#standardization of the continuous variables (2 sd)

mutate_if(is.numeric, scale_this)

set.seed(2022)

splits <- initial_split(dyhtkeskp, strata = walkbikekes_c)

train <- training(splits)

test <- testing(splits)

y <- train %>% select(walkbikekes) %>% data.matrix()

x <- train %>% select(-walkbikekes) %>% data.matrix()
```

21

PREDICTION OF COMMUTE MODE CHOICE

```
bofit <- bolasso(x=x,y=y, BM=100, kfold=10, family="binomial", implement="glmnet",

        standardize=F)

vars <- selected_vars(bofit, threshold = .95,  select = "lambda.1se") #lambda.1se

model <- as.formula(paste0("walkbikekes_c~",paste(vars$variable[-1], collapse="+")))

newfit <- glm(model,data=train, family="binomial")

wb_or<-tidy(newfit, exp=T, conf.int = T)

vipvalues<-newfit %>% vi(method = "model")

vipvalues

walk5plot<-   ggplot(vipvalues,aes(y=reorder(Variable,   Importance),x=Importance))+   #
color=Sign))  +

geom_point(size=3) +

geom_segment(aes(x=min(Importance), xend=Importance, y=reorder(Variable, Importance),
yend=reorder(Variable, Importance))) +

theme_apa() +

theme(legend.position="none") +

#scale_color_manual(values=c("POS"="red", "NEG"="blue")) +

scale_x_continuous(breaks = range(vipvalues$Importance),

        labels=c("Low", "High")) +
```

22

PREDICTION OF COMMUTE MODE CHOICE

```
labs(y="")

test$pred <- predict(newfit, newdata=test)

pred <- prediction(test$pred,test$walkbikekes_c)

perf <- performance(pred,"tpr","fpr")

auc <- performance(pred, measure = "auc")

auc <- auc@y.values[[1]]

auc <- round(auc,2)

auc_aggre_data <- data.frame(x=perf@x.values[[1]], y=perf@y.values[[1]])

#AUC curve!

auc_aggre_data <- data.frame(x=perf@x.values[[1]], y=perf@y.values[[1]])

ggplot(auc_aggre_data, aes(x,y))+

geom_line()+

geom_abline(intercept = 0, slope = 1)+

labs(x="False positive rate", y="True positive rate")+

theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),

panel.background = element_blank(), axis.line = element_line(colour = "black"),

panel.border = element_rect(colour = "black", fill=NA))+
```

23

PREDICTION OF COMMUTE MODE CHOICE

```
annotate("text",.4,.85,label=paste0("AUC: ",auc))

walk5plot <-walk5plot + ggtitle(paste0("Commute by bike or on foot, AUC=",auc)) +

 #scale_y_discrete(labels=c("Body mass index (low)","Physical activity (high)","Commute
length (low)"))

 #scale_y_discrete(labels=c("Sex (female)","Smoking (no)","Body mass index (low)",
"Commute length (low)"))

 #scale_y_discrete(labels=c("Work climate (good)", "Sex (female)", "Body mass index
(low)","Physical activity (high)","Commute length (low)"))

#scale_y_discrete(labels=c("Work climate (good)","Efforts (high)", "Sex (female)","Smoking
(no)","Body mass index (low)","Commute length (low)"))

walk5plot
```

PREDICTION OF COMMUTE MODE CHOICE

References

1. Karasek, R. A. Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Adm. Sci. Q.* **24**, 285 (1979).

2. Karasek, R. A. & Theorell, T. *Healthy Work: Stress, Productivity and the Reconstruction of Working Life.*

3. Ervasti, J. *et al.* Prediction of bullying at work: A data-driven analysis of the Finnish public sector cohort study. *Soc. Sci. Med.* **317**, 115590 (2023).

4. Kivimaki, M., Vahtera, J., Elovainio, M., Virtanen, M. & Siegrist, J. Effort-reward imbalance, procedural injustice and relational injustice as psychosocial predictors of health: complementary or redundant models? *Occup. Environ. Med.* **64**, 659–665 (2007).

5. Ala-Mursula, L. Effect of employee worktime control on health: a prospective cohort study. *Occup. Environ. Med.* **61**, 254–261 (2004).

6. Vahtera, J. *et al.* Employee control over working times and risk of cause-specific disability pension: the Finnish Public Sector Study. *Occup. Environ. Med.* **67**, 479–485 (2010).

7. Moorman, R. H. Relationship between organizational justice and organizational citizenship behaviors: Do fairness perceptions influence employee citizenship? *J. Appl. Psychol.* **76**, 845–855 (1991).

8. Kivimaki, M. & Elovainio, M. A short version of the Team Climate Inventory: Development and psychometric properties. *J. Occup. Organ. Psychol.* **72**, 241–246 (1999).

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

PREDICTION OF COMMUTE MODE CHOICE

9. Anderson, N. & West, M. A. The team climate inventory: Development of the tci and its applications in teambuilding for innovativeness. *Eur. J. Work Organ. Psychol.* **5**, 53–66 (1996).

10. GOLDBERG, D. P. *et al.* The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychol. Med.* **27**, 191–197 (1997).

11. Juhola, J. *et al.* Internal consistency and factor structure of Jenkins Sleep Scale: cross-sectional cohort study among 80 000 adults. *BMJ Open* **11**, e043276–e043276 (2021).

12. Robine, J. & Jagger, C. CREATING A COHERENT SET OF INDICATORS TO MONITOR HEALTH ACROSS EUROPE: THE EUROREVES 2 PROJECT. *The Gerontologist* **42**, 35–35 (2002).

13. Ilmarinen, J., Tuomi, K. & Klockars, M. Changes in the work ability of active employees over an 11-year period. *Scand. J. Work. Environ. Health* **23**, 49–57 (1997).

14. Tuomi, K., Ilmarinen, J., Martikainen, R., Aalto, L. & Klockars, M. Aging, work, life-style and work ability among Finnish municipal workers in 1981—1992. *Scand. J. Work. Environ. Health* **23**, 58–65 (1997).

15. Ahlstrom, L., Grimby-Ekman, A., Hagberg, M. & Dellve, L. The Work Ability Index and single-item Question: Associations with Sick Leave, Symptoms, and health-a Prospective Study of Women on long-term Sick Leave. *Scand. J. Work Environ. & Health* **36**, 404–412 (2010).

16. Halonen, J. I. *et al.* Green and blue areas as predictors of overweight and obesity in an 8-year follow-up study. *Obes. Silver Spring Md* **22**, 1910–1917 (2014).

PREDICTION OF COMMUTE MODE CHOICE

17. Ervasti, J. *et al.* Sickness absence diagnoses among abstainers, low-risk drinkers and at-risk drinkers: consideration of the U-shaped association between alcohol use and sickness absence in four cohort studies. *Addict. Abingdon Engl.* **113**, 1633–1642 (2018).

18. Heikkilä, K. *et al.* Job strain and tobacco smoking: an individual-participant data meta-analysis of 166,130 adults in 15 European studies. *PloS One* **7**, e35463–e35463 (2012).

19. Ervasti, J. *et al.* Does increasing physical activity reduce the excess risk of work disability among overweight individuals? *Scand. J. Work. Environ. Health* **45**, 376–385 (2019).

TRIPOD Checklist: Prediction Model Development

Section/Topic	Item	Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2-3
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	2,10, 11,12, 22, 23
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	2, 3, 10, 11
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	7
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	7
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	7
	5b	Describe eligibility criteria for participants.	8
	5c	Give details of treatments received, if relevant.	-
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	8-9
	6b	Report any actions to blind assessment of the outcome to be predicted.	-
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	8,9,10 + appendix 1
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	-
Sample size	8	Explain how the study size was arrived at.	7
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	8
Statistical analysis methods	10a	Describe how predictors were handled in the analyses.	10-12
	10b	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	10-12
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	10-12
Risk groups	11	Provide details on how risk groups were created, if done.	-
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	7
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	7, 13+ appendix 1 p.8
Model development	14a	Specify the number of participants and outcome events in each analysis.	10-12
	14b	If done, report the unadjusted association between each candidate predictor and outcome.	-
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	14-17
	15b	Explain how to the use the prediction model.	10-12, 23
Model performance	16	Report performance measures (with CIs) for the prediction model.	15-16
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	3, 22
Interpretation	19b	Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence.	18-22
Implications	20	Discuss the potential clinical use of the model and implications for future research.	23
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	24, web appendix
Funding	22	Give the source of funding and the role of the funders for the present study.	1

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

TRIPOD Checklist: Prediction Model Development

We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.