

A linear regression-based machine learning pipeline for the discovery of clinically relevant correlates of gait speed reserve from multiple physiological systems

James Davis
The Irish Longitudinal Study on
Ageing (TILDA),
Trinity College Dublin,
Dublin, Ireland
ORCID: 0000-0001-6388-2184

Orna A. Donoghue
The Irish Longitudinal Study on
Ageing (TILDA),
Trinity College Dublin,
Dublin, Ireland
ORCID: 0000-0002-2091-3478

Silvin P. Knight
The Irish Longitudinal Study on
Ageing (TILDA),
Trinity College Dublin,
Dublin, Ireland
ORCID: 0000-0003-1245-4870

Rose Anne Kenny
The Irish Longitudinal Study on
Ageing (TILDA),
Trinity College Dublin,
Dublin, Ireland
ORCID: 0000-0002-9336-8124

Rossella Rizzo
The Irish Longitudinal Study on
Ageing (TILDA),
Trinity College Dublin,
Dublin, Ireland
ORCID: 0000-0003-4486-4794

Roman Romero-Ortuno
The Irish Longitudinal Study on
Ageing (TILDA),
Trinity College Dublin,
Dublin, Ireland
ORCID: 0000-0002-3882-7447

Abstract— Frailty in older adults is characterized by reduced physiological reserve. Gait speed reserve (GSR: maximum minus usual gait speed) could help identify frailty and act as a proxy for physiological reserve. Utilizing data from 2397 participants aged 50+ from wave 3 of The Irish Longitudinal Study on Ageing, we developed a stepwise linear regression-based machine learning pipeline to select the most important GSR predictors from 34 manually selected features across multiple domains. Variables were selected one at a time such that they maximized the mean adjusted r-squared score from a 5-fold cross-validation. A peak score of (0.16 ± 0.03) was achieved with 14 variables (giving adjusted-r-squared of 0.18 and 0.16 on 80% training and 20% test data, respectively). The first 7 variables explained 95% of the peak score: grip strength, MOCA cognitive score, third level education, chair stands time, sex, age, and body mass index (BMI). Of the 14 selected features, 11 had statistically significant ($p < 0.05$) effects in the model: sex, MOCA, third level education, chair stands time, age, BMI, grip strength, cardiac output, number of medications, fear of falling, and mean choice reaction time. Associations between GSR and multi-domain features suggest that a network physiology approach is necessary for assessing physiological reserve.

Keywords— *Gait Speed Reserve; Linear Regression; Machine Learning; Network Physiology; Aging; Frailty.*

I. INTRODUCTION

Gait speed is a measure of general fitness and is associated with the ability to meet occupational demands in younger adults [1] and functional decline and morbidity in older adults [2, 3]. Even though comfortable walking speed and maximum walking speed are significantly intercorrelated [4], changing from comfortable (or usual) speed to maximum speed requires a general effort across many body systems. We propose that the difference between these two gait speeds, which we refer to here as gait speed reserve (GSR: maximum minus usual gait speed), may be a useful proxy measure of physiological reserve. There is increasing medical interest to understand what the main determinants of physiological reserve are and ways in which to quantify it. In addition, it

has been postulated that in community-dwelling older adults, the simultaneous consideration of both usual and maximum walking speed, as with GSR, could increase the accuracy of the identification of frailty [5]. Indeed, in older adults, frailty is defined as a state of reduced physiological reserve increasing the risk of decompensation and adverse clinical outcomes triggered by relatively minor stressors [6]. In women aged over 70 years, a stronger association was seen between gait reserve and frailty than for usual gait speed alone [7]. However, there have been no studies attempting to model predictors of GSR in a large sample of community-dwelling older adults where many demographic, anthropometric and clinical features were measured across multiple physiological systems. In the present study, we developed a regression-based machine learning pipeline for the discovery of clinically relevant predictors of GSR across multiple physiological systems using data from The Irish Longitudinal Study on Ageing (TILDA).

II. METHODS

A. Design and setting

We analyzed data from adults aged 50+ from Wave 3 of TILDA, a population-based longitudinal study of ageing. TILDA study design, as well as the full cohort profile, have been previously described in detail [8, 9]. Wave 3 data collection took place in 2014 and 2015 and included a computer-assisted personal interview conducted by social interviewers in the participants' home, a self-completion questionnaire completed in the participants' own time and a detailed suite of technology-aided health assessments conducted by trained research nurses at a dedicated health centre. Ethical approval was obtained from the Faculty of Health Sciences Research Ethics Committee at Trinity College Dublin, Ireland. All participants provided written informed consent.

This study was funded by a grant from Science Foundation Ireland (SFI) [18/FRL/6188]. TILDA is funded by Atlantic Philanthropies, the Irish Department of Health and Irish Life. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

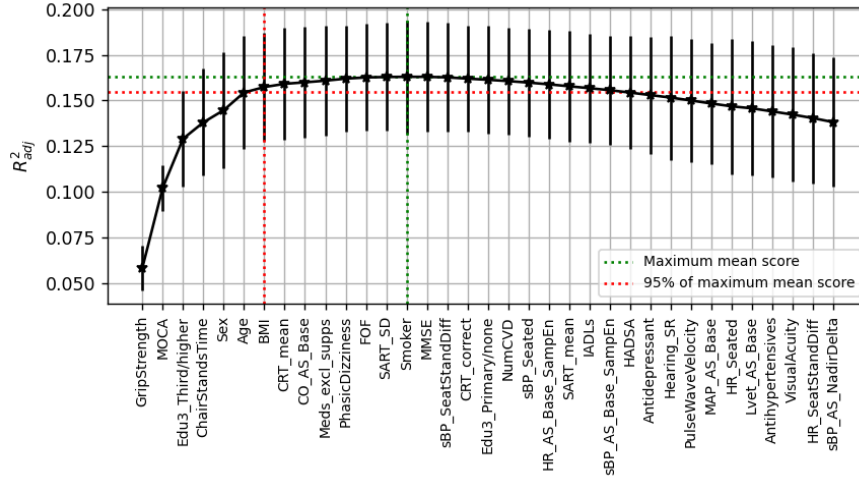


Fig. 1: automated feature selection curve. The x-axis shows the names of the features in the order they are selected for the model. All features to the left of a given feature are included in the model with that feature. The y-axis represents R^2_{adj} . Mean and standard deviation (SD) of R^2_{adj} values from the 5-fold CV are shown. The green lines indicate the peak mean score and the feature at which it occurred. The red lines indicate the 95% on the peak mean score and the feature at which it is achieved.

B. Measures

At Wave 3 of TILDA, gait speed was measured as part of a health centre assessment. Measurements were made using a 4.88 m computerized walkway (GAITRite, CIR Systems, NY, USA). A two-meter space before and after the walkway was used for acceleration and deceleration. Participants were first asked to walk at their normal (usual) pace and then as fast as they safely could (maximum). Two walking trials were obtained in each condition and the mean value for each was used in this analysis. GSR was defined as maximum gait speed – usual gait speed.

Other measures collected included demographics (age, sex), sociodemographic information (level of education), anthropometrics (body mass index (BMI)), self-reported medical history (medical conditions, medications), behavioral health (smoking status), difficulty with instrumental activities of daily living, psychological variables (anxiety, fear of falling), and two paper-based cognitive assessments: the Montreal Cognitive Assessment (MOCA) and the Mini Mental State Examination (MMSE). Participants also underwent performance-based physical assessments including a five chair stands test (for the assessment of lower limb muscle strength), dynamometer handgrip strength (as a measure of upper limb strength), a visual acuity test, and blood pressure and heart rate measurements (sitting and standing). In addition, participants underwent non-invasive continuous haemodynamic monitoring using a Finometer MIDI device (Finapres Medical Systems BV, Amsterdam, the Netherlands) during an active stand test, which provided additional cardiovascular function parameters such as mean arterial pressure, left ventricular ejection time, and cardiac output. The active stand test has the participant laying supine for 10 minutes before standing up as quickly as is comfortably possible and remaining standing still for a further 3 minutes. Participants are asked if they felt dizzy on standing. Baseline cardiovascular values were obtained from the resting state time window. Cardiovascular signal complexity (sample

entropy) for blood pressure and heart rate were also derived [10]. Subjects also underwent non-invasive measurement of pulse wave velocity (a measure of arterial stiffness). Computerised cognitive tests of sustained attention and choice reaction time were also carried out. A summary of all the measures considered in this analysis is shown in Appendix I. Further details on all these TILDA tests have been described elsewhere [8].

C. Feature selection pipeline

We began with a total of 34 features manually selected from Wave 3 of TILDA (Appendix I). The features were chosen from multiple domains including demographics, cardiovascular system, cognition, psychology, hearing and vision, physical strength, and medications. In this analysis, participants with missing values for any of the above-mentioned 34 variables were removed. The data were divided via an 80/20 train/test split. From there, using the training data, an automatic feature selection was employed that utilizes a stepwise approach in which features are added to a linear regression model one at a time such that they maximize the mean adjusted r-squared, $\overline{R^2_{adj}}$, which was calculated as the mean R^2_{adj} value obtained from a 5-fold cross-validation (CV). The 5-fold CV was introduced to reduce overfitting to the entire dataset and help identify the features that performed best across multiple subsets of the data. On each iteration of the CV, a pipeline consisting of a standard scaler and a linear regression was employed. The features selected for the final model were those corresponding to the peak of a $\overline{R^2_{adj}}$ vs. added features plot. The selected features were used to train a model on the training data. The model was tested on the remaining 20% of the data. This process was implemented using the Scikit-learn package (v0.19.1) for Python (v3.8.3).

D. Feature importance

The importance of the selected features was assessed by fitting an ordinary least squares (OLS) model to the training data. The OLS model used was from the Statsmodels package (v0.13.0) for Python. The statistically significant ($p < 0.05$)

features from the model, and the features that built up to 95% of the maximum \bar{R}_{adj}^2 were retained for discussion.

III. RESULTS

Of the 4309 participants that took part in the TILDA Wave 3 health assessment centre, 3925 aged 50+ completed both the usual and maximum walking tests required to generate the GSR data. Of those, 2397 (61%) were included in the linear regression analysis as they had no missing values for any of the features. Female sex made up 52.9% of the 2397 participants.

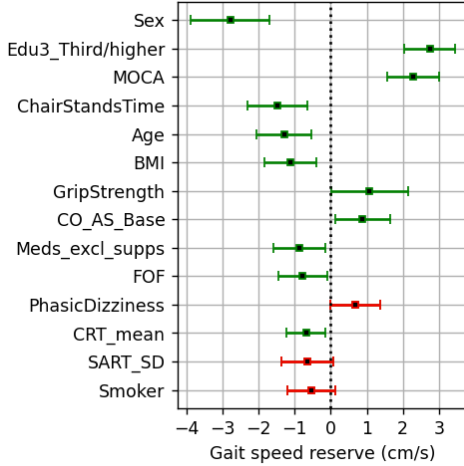


Fig. 2: Visual summary of regression coefficients for the standardized input features. The y-axis presents the final model features in order of descending coefficient magnitude from top to bottom. The x-axis shows the coefficients effect size with 95% confidence interval in terms of absolute GSR with units of cm/s. Green markers represent statistically significant ($p < 0.05$) effects. Sex is coded as male = 0 and female = 1.

Fig. 1 shows the $\bar{R}_{adj}^2 \pm SD$ as features are added to the model. The features appear on the x-axis in a cumulative manner, i.e. the \bar{R}_{adj}^2 value at a given x-coordinate corresponds to that of a model containing the feature at that coordinate plus all the other features to its left. The maximum \bar{R}_{adj}^2 of 0.16 ± 0.03 was achieved in the feature selection with 14 features. A linear regression model containing these 14 features returned an \bar{R}_{adj}^2 of 0.18 and 0.16 on the training and test data, respectively. However, just the first seven predictors built 95% of the peak \bar{R}_{adj}^2 score. In order of addition to the model, these seven features were: grip strength, MOCA score, third level education, chair stands time, sex, age, and BMI.

The effect sizes with 95% confidence intervals of the regression coefficients from a model trained using all 14 selected features are shown in Fig. 2. The input features are standardized, and the effect size (shown along the x-axis) is in terms of GSR (units of cm/s). The y-axis orders the model features in order of decreasing coefficient magnitude from top to bottom. The 11 statistically significant coefficients (shown in green in Fig. 2) are: sex, third level education, MOCA, chair stands time, age, BMI, grip strength, cardiac output at resting state, number of medications, fear of falling, and mean choice reaction time. Female sex, chair stands time, age, BMI number of medications, fear of falling, and mean

choice reaction time were associated with a decrease in GSR, whilst third level education, MOCA score, grip strength, and baseline cardiac output were associated with increased GSR.

IV. DISCUSSION

We developed a linear regression-based machine learning pipeline for the discovery of clinically relevant predictors of GSR across multiple physiological systems in TILDA. The first 7 of the 14 selected predictors (grip strength, MOCA score, third level education, chair stands time, sex, age, and BMI) explained 95% of the maximum \bar{R}_{adj}^2 achieved (0.16). When examining the regression coefficients, we found that 11 variables were statistically significant: sex, third level education, MOCA, chair stands time, age, BMI, grip strength, cardiac output at resting state, number of medications, fear of falling, and mean choice reaction time.

The results show that there were significant associations between GSR and features from multiple physiological systems (e.g. cognitive, psychological, musculoskeletal, cardiovascular), which supports that GSR could be useful as an indicator of overall physiological reserve.

Our results are consistent with previous research showing significant differences in maximum walking speed for different ages and between men and women [11, 12]. Our results are also consistent with previous findings that higher BMI and physical workload among those with lower education contributed most to the educational disparities in age-related decline in maximum gait speed [13]. Cognitive performance has also been cited as a significant predictor of maximum gait speed [14] and from our study, we can conclude that the MOCA may be more predictive than the MMSE test in this regard. Recent work has suggested that longer choice reaction time may be associated with longitudinal mobility decline [15]. Furthermore, our results underscore that even though maximum gait speed can be expected to be reduced in individuals with weaker lower extremity muscle strength [16, 17], upper limb strength assessment also needs to be considered for GSR prediction. Indeed, previous studies have shown that the movement velocity of the upper limbs is a significant determinant of maximum gait speed, suggesting that the ability to move any region rapidly might be a critical factor in maximum gait speed [18]. The importance of both upper and lower limb muscles as predictors of GSR offers clinical opportunities for strengthening exercises as a way to improve physiological reserve [19]. Fear of falling being predictive of decreased GSR is also clinically plausible, which offers opportunities for psychological interventions. Our results also suggest that there is clinical scope for cardiovascular health optimization in the context of GSR improvement.

A limitation of our study is that due to the inclusion of participants with complete data for the 34 initial features, the analytical sample size was reduced to 61% of the original sample. In future work, this limitation could be addressed by exclusion of features (where clinically acceptable) with excessive proportion of missing data, or by performing multiple imputation of missing data. In addition, even though GSR had significant predictors in our study, the total variance

explained (judged by R^2_{adj}) was low to moderate [20], suggesting that there is further scope for consideration of additional variables. Future feature discovery should attempt to increase the amount of variance explained with additional predictors and/or implementation of the study in clinical cohorts. However, our level of explained variance is in keeping with previous observations that maximum gait speed as a single-item tool is limited to fully predict future falls in community-dwelling older persons [21]. Another limitation to consider is that measurement of GSR restricts analysis to TILDA participants who attended the health assessment centre. Those who could not attend are likely to be frailer than those who did.

Given the observed association between GSR and sex, a stratification by sex may have revealed more nuanced differences between the two groups. Furthermore, an analysis by age groups could be considered to explore whether associations with GSR might vary with age. In addition, in future work we could compare current results with those from non-linear and non-parametric machine learning models.

According to previous research, an older persons' probability of being frail with an insufficient GSR could be around 40% [5], which further bolsters the clinical relevance of the present results, since many of the identified associations are clinically modifiable. Indeed, improved education, obesity prevention, cardiovascular risk reduction, cognitive training, appropriate prescribing and monitoring of medication, neuro-psychological interventions against fear of falling, and muscle strengthening could all potentially improve GSR in older populations. GSR could also have safety implications during daily activities that require a sudden increase in pace such as crossing the road, running for the bus, reacting to hazards, etc., and hence be important to maintain older people's functional independence. Our results demonstrate the importance of a network physiology approach for the understanding of frailty and resilience in ageing, where systems work together towards the generation of physiological reserve [22].

ACKNOWLEDGMENT

The authors would like to acknowledge the continued commitment and cooperation of the TILDA participants and research team.

REFERENCES

- [1] C. Aldridge *et al.*, "Walking at work: Maximum gait speed is related to work ability in hospital nursing staff," *J Occup Health*, vol. 62, no. 1, p. e12171, Jan 2020, doi: 10.1002/1348-9585.12171.
- [2] R. W. Bohannon, "Walking after stroke: comfortable versus maximum safe speed," *Int J Rehabil Res*, vol. 15, no. 3, pp. 246-8, 1992. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/1428391>.
- [3] H. Kawajiri *et al.*, "Maximum Walking Speed at Discharge Could Be a Prognostic Factor for Vascular Events in Patients With Mild Stroke: A Cohort Study," *Arch Phys Med Rehabil*, vol. 100, no. 2, pp. 230-238, Feb 2019, doi: 10.1016/j.apmr.2018.05.025.
- [4] B. Kollen, G. Kwakkel, and E. Lindeman, "Hemiplegic gait after stroke: is measurement of maximum speed required?," *Arch Phys Med Rehabil*, vol. 87, no. 3, pp. 358-63, Mar 2006, doi: 10.1016/j.apmr.2005.11.007.
- [5] M. do Carmo Correia de Lima, T. Loffredo Bilton, W. Jefferson de Sousa Soares, L. Paccini Lustosa, E. Ferrioli, and M. Rodrigues Perracini, "Maximum Walking Speed Can Improve the Diagnostic Value of Frailty among Community-Dwelling Older Adults: A Cross-Sectional Study," *J Frailty Aging*, vol. 8, no. 1, pp. 39-41, 2019, doi: 10.14283/jfa.2018.44.
- [6] A. Clegg, J. Young, S. Iliffe, M. O. Rikkert, and K. Rockwood, "Frailty in elderly people," *Lancet*, vol. 381, no. 9868, pp. 752-62, Mar 2 2013, doi: 10.1016/S0140-6736(12)62167-9.
- [7] A. Noguerón García *et al.*, "Gait plasticity impairment as an early frailty biomarker," *Experimental Gerontology*, vol. 142, p. 111137, 2020, doi: 10.1016/j.exger.2020.111137.
- [8] P. M. Kearney *et al.*, "Cohort profile: the Irish Longitudinal Study on Ageing," *Int J Epidemiol*, vol. 40, no. 4, pp. 877-84, Aug 2011, doi: 10.1093/ije/dyr116.
- [9] O. A. Donoghue, C. A. McGarrigle, M. Foley, A. Fagan, J. Meaney, and R. A. Kenny, "Cohort Profile Update: The Irish Longitudinal Study on Ageing (TILDA)," *Int J Epidemiol*, vol. 47, no. 5, pp. 1398-1398I, Oct 1 2018, doi: 10.1093/ije/dyy163.
- [10] S. P. Knight, L. Newman, J. D. O'Connor, J. Davis, R. A. Kenny, and R. Romero-Ortuno, "Associations between Neurocardiovascular Signal Entropy and Physical Frailty," *Entropy (Basel)*, vol. 23, no. 1, Dec 22 2020, doi: 10.3390/e23010004.
- [11] S. Tibaek, N. Holmestad-Bechmann, T. B. Pedersen, S. M. Bramming, and A. K. Friis, "Reference values of maximum walking speed among independent community-dwelling Danish adults aged 60 to 79 years: a cross-sectional study," *Physiotherapy*, vol. 101, no. 2, pp. 135-40, Jun 2015, doi: 10.1016/j.physio.2014.08.005.
- [12] K. P. Izawa *et al.*, "Gender-related differences in maximum gait speed and daily physical activity in elderly hospitalized cardiac inpatients: a preliminary study," *Medicine (Baltimore)*, vol. 94, no. 11, p. e623, Mar 2015, doi: 10.1097/MD.0000000000000623.
- [13] S. M. Kyronlahti, S. Stenholm, J. Raitanen, S. Neupane, S. Koskinen, and K. Tiainen, "Educational Differences in Decline in Maximum Gait Speed in Older Adults over an 11-year Follow-up," *J Gerontol A Biol Sci Med Sci*, Aug 14 2020, doi: 10.1093/gerona/glaa196.
- [14] H. Umegaki *et al.*, "Maximum gait speed is associated with a wide range of cognitive functions in Japanese older adults with a Clinical Dementia Rating of 0.5," *Geriatr Gerontol Int*, vol. 18, no. 9, pp. 1323-1329, Sep 2018, doi: 10.1111/ggi.13464.
- [15] R. Chintapalli and R. Romero-Ortuno, "Choice reaction time and subsequent mobility decline: Prospective observational findings from The Irish Longitudinal Study on Ageing (TILDA)," *EClinicalMedicine*, vol. 31, p. 100676, 2021, doi: 10.1016/j.eclinm.2020.100676.
- [16] R. W. Bohannon, "Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants," *Age Ageing*, vol. 26, no. 1, pp. 15-9, Jan 1997, doi: 10.1093/ageing/26.1.15.
- [17] D. J. Clark, T. M. Manini, R. A. Fielding, and C. Patten, "Neuromuscular determinants of maximum walking speed in well-functioning older adults," *Exp Gerontol*, vol. 48, no. 3, pp. 358-63, Mar 2013, doi: 10.1016/j.exger.2013.01.010.
- [18] A. Iwata *et al.*, "Maximum movement velocity of the upper limbs reflects maximum gait speed in community-dwelling adults aged older than 60 years," *Geriatr Gerontol Int*, vol. 14, no. 4, pp. 886-91, Oct 2014, doi: 10.1111/ggi.12186.
- [19] T. Rantanen *et al.*, "Association of muscle strength with maximum walking speed in disabled older women," *Am J Phys Med Rehabil*, vol. 77, no. 4, pp. 299-305, Jul-Aug 1998, doi: 10.1097/00002060-199807000-00008.
- [20] C. J. Ferguson, "An Effect Size Primer: A Guide for Clinicians and Researchers," *Professional Psychology: Research and Practice*, vol. 40, 5, pp. 532-538, 2009.
- [21] K. T. Bongers, Y. Schoon, M. J. Graauwman, H. J. Schers, R. J. Melis, and M. G. Olde Rikkert, "The predictive value of gait speed and maximum step length for falling in community-dwelling older persons," *Age Ageing*, vol. 44, no. 2, pp. 294-9, Mar 2015, doi: 10.1093/ageing/afu151.
- [22] R. Rizzo, X. Zhang, J. Wang, F. Lombardi, and P. C. Ivanov, "Network Physiology of Cortico-Muscular Interactions," *Front Physiol*, vol. 11, p. 558070, 2020, doi: 10.3389/fphys.2020.558070.

APPENDIX 1

Demographics	Age	Age of participant assuming date of birth is 1 st of specific month
	Sex	Sex of participant (male = 0, female=1)
Behavioural health	Smoker	Current smoker: yes/no
Anthropometric	BMI	Body mass index: weight/height ² [kg/m ²]
Sociodemographic	Edu3	Highest education achieved: primary/none, secondary, or third/higher. Encoded using one-hot encoded dummy variables with the 'Secondary' dummy variable dropped to avoid multicollinearity.
Living Impairments	IADLs	Number of instrumental activities of daily living impairments
Cognitive	CRT mean	Choice reaction time: mean time [ms]
	CRT correct	Choice reaction time: number of correct choices
	MOCA	Montreal Cognitive Assessment (scale 0-30, 30 being best performance)
	SART mean	Sustained Attention to Reaction Time: mean time [ms]
	SART SD	Sustained Attention to Reaction Time: time standard deviation [ms]
	MMSE	Mini-mental state examination (30-point scale with 30 being best performance)
Psychological	HADSA	Hospital Anxiety and Depression Scale – Anxiety (scale of 0 to 21 with higher indicating more anxiety)
	FOF	Fear of Falling: yes/no.
Hearing and vision	Hearing SR	Self-rated hearing on a scale of 1-5 with higher being worse
	VisualAcuity	100-50*(LogMar Visual Acuity of best eye))
Self-reported medications and medical history	Meds excl supps	Number of medications being taken excluding supplements
	Antidepressant	Taking an antidepressant: yes/no
	Antihypertensives	Taking an antihypertensive: yes/no
	NumCVD	Number of cardiovascular diseases
Cardiovascular	PulseWaveVelocity	Pulse wave velocity [m/s]
	sBP Seated	Seated systolic blood pressure [mmHg]
	HR Seated	Seated heart rate [Hz]
	sBP SeatStandDiff	Difference between seated and standing systolic blood pressure [mmHg]
	HR SeatStandDiff	Difference between seated and standing heart rate [Hz]
Cardiovascular (active stand)	MAP AS Base	Mean Arterial Pressure at baseline of active stand [mmHg]. Measured during resting state of active stand.
	sBP AS NadirDelta	Drop in systolic blood pressure during active stand [mmHg] i.e. drop from when resting (supine) to standing.
	Lvet AS Base	Left ventricular ejection time at baseline of active stand [ms]. Measured during resting state of active stand.
	CO AS Base	Cardiac output at baseline [L/min]. Measured during resting state of active stand. Measured during resting state of active stand.
	PhasicDizziness	Did the participant experience phasic dizziness during the active stand test
	BP RS sBP SampEn	Sample entropy of baseline systolic blood pressure signal
	BP RS HR SampEn	Sample entropy of baseline heart rate signal
Physical strength	GripStrength	Maximum grip strength from 8 trails (4 with each hand) [kg]
	ChairStandsTime	Time taken to perform five chair stands [s]