Statistical analysis plan for ADDITION PRO study

1 Title: Long-term weekly heart rate variability association with cardiovascular disease in prediabetes – a prospective cohort-study from the ADDITION-PRO

Jonas R. Schaarup[1], Lasse Bjerg[2], Christian S. Hansen[3], Annelli Sandbæk[1,2], Dorte Vistisen[3,4], Daniel R. Witte[1,2]

[1]Department of Public Health, Aarhus University, Denmark; [2] Steno Diabetes Center Aarhus, Aarhus University Hospital, Denmark; [3] Steno Diabetes Center Copenhagen, Copenhagen, Denmark; [4]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

ORCID identifier(s)

…….

Corresponding author

Jonas R. Schaarup

Work address: Steno Diabetes Center Aarhus, Palle Juul-Jensens Boulevard 11, Aarhus N, 8200, DK

E-mail: jfrscha@ph.au.dk

Phone number: +45 29936899

1.1 Description

Inclusion criteria for the participants was having a least 48 hour measured heart rate and no prior CVD events before inclusion.

Frequencies (percentages) and mean (lSD) will be calculated to describe categorical and continuous variables, respectively. These descriptive statistics will be presented in the table. Table 1 gives an overview of the study population general characteristics (demographic, lifestyle, bio markers). Prevalence and incidence rate of cardiovascular events (myocardial infarction, stroke, and heart failure), will be shown in table 2. Flow of participation will be shown in figure 1.

1.1.1 Calculation of HRV time-domain indices

Based on actiheart, we had measures of heart rate traces for up to 7 days. Mean heart rate with prediction interval were obtained every 30-second epoch. Minimum and maximum as well as 2nd lowest minimum and 2nd highest maximum inter-beat-interval where measured from the latest 16 heartbeat in each 60-second epoch.

We did not have access to time-series of successive IBI, in the period of measurement. Therefore, we generated IBIs in time-spand of every 30-second interval based on 30-second epoch of mean heart rate and prediction intervals. As earlier data from studies have shown that IBI are normally distributed per 30-second epoch, we generated IBI 30-second distribution by using mean heart rate and its standard deviation. In order to calculate SD from prediction intervals, we assured that the prediction intervals symmetric differed from the mean by calculating the difference between the upper and lower prediction interval from the mean heart rate and visually observing their symmetry over time. Using RHRV package in R, we calculated HRV indices. As we did not have successive time-series measurement, we can only use HRV indices that is based on distribution of RR intervals. Therefore, frequency domain measures were not included in the study. We included HRV indices standard deviation between normal to normal heart beat intervals (SDNN), standard Deviation of the 5 minute Average NN intervals (SDANN), SDNN index (SDNNi), HRV triangular index, triangular interpolation of NN interval histogram (TINN). Additionally, by taking the mean of the difference between 2nd highest and 2nd lowest inter-beat-interval for each 60-second epoch, we estimated the HRV index, IBI difference, a measurement (use Søren Brage paper description). All HRV indices where calculate by week, day, hour per day, and circadian block per day (00:00-06:00, 06:00-12:00, 12:00-18:00, 18:00-00:00).

Furthermore, mean heart rate and resting heart rate was included. Resting heart rate was determined by ?the lowest heart for a time period, in the resting stage in supine position. this was cross-checked by accelerometry data?

1.1.2 Statistical analysis

In the statistical analysis, we would like to use two approaches. One traditional statistical investigating the association between HRV. In both approaches, we want to examine the temporal changes in HR and HRV from the circadian rhythm.

Traditional statistical approach

We will investigate two aims in our analysis: 1) Determine the association between week-long HRV and CVD; 2) Determine the diurnal variation in hourly HRV association to future CVD.

For the first objective, we will use poisson regression models to investigate the association between week long heart rate variability indices and hard cardiovascular diseases outcomes (including myocardial infarction, stroke, and heart failure) as well as all-cause mortality. We will fit three models. Model 1 will include adjustments of age and sex. In addition to these, Model 2 will include alcohol consumption, smoking behavior, diet, physical activity, education, systolic blood pressure, body mass index, total cholesterol, and Hba1c, and Model 3 with further adjustments of anti-hypertensive, and glucose-lowering medication. The incidence rate by increase in HRV from the poisson models will be presented in table 3. Because of biological differences in sex with regard to the development CVD and women genuinely having a lower heart rate variability than men, we want to investigate the stratified association by sex.

To investigate for non-linearity (splines), we included defined knots based on percentiles in HRV distribution. The results from the spline models will by visualized in figure 2.

We will perform similar analysis for each hour of HRV in a day. Point incidence estimates with 95% confidence interval will be plotted by each hour (see figure 3). For this analysis, we will not include knot for HRV analysis.

Both with multiple imputation and complete case analyses will be conducted in the R statistical computing environment (version X).

Exploratory ideas

Additional suggestion to account for: actiheart was performed in different season, hence it might by interesting to account or examine seasonal variations. Multi-stage from prediabetes and HRV measurement to T2D to CVD risk, compared to prediabetes and HRV to CVD risk. Multi-stage to different types of anti-hypertensive medication (ace-inhibitors, beta blockers, diuretics), and to CVD.

Machine learning approach

Use all available data as detailed as possible to look into predicting CVD events. These models can include Random Forrest, Bayesian Additive Regression Trees (BART) for simple data, and neural network for more complicated models using time-series data. Using neural network, we

will train time-series data from actiheart unsupervised and then upload the model to DST to test is performance in detecting high risk CVD individuals.

1.2 Tables and Figures

List of variables

Variables from ADDITION-PRO dataset	Available at DST	variable name
Age (years)	x	age_fup
Sex (male)	x	p_gv_sex
Socioeconomic status	x	
Smoking status	x	smoke_q_4_a smoke_cigarettes_day_q4_b p_gq_smoke
Physical activity (PAEE kj_kg_day)		PAEE_kj_kg_day
Step test results 1. VO2 max (Cardiorespiratory function)		Pred.VO2max
Alcohol comsuption (units per week)	x	alkohol_week_unit_total
BMI (kg/m^2)	x	bmi
Height (cm)		height
Weight (kg)		?????
Waist circumference (cm)	x	waist_av
Hip circumference (cm)		hip_av
Fat percentage (%)	x	fat_pc
Visceral fat		vat
Subcutaneous fat		sat
Systolic blood pressure (mm hg)	x	sbp_av
Diastolic blood pressure (mm hg)	x	dbp_av
Hours fasting		hours_fast~g
HbA1c (%)		p_lab_hba1c
LDL cholesterol (mmol/L)		p_lab_ldl
HDL cholesterol (mmol/L)		p_lab_hdlc
Total cholesterol (mmol/L)		p_lab_chol
Triglycerides (mmol/L)		p_lab_trig
Insulin resistance		homa_ir
Beta cell function		homa_b
Insulin sensitivity		homa_s
Insulin sensitivity (Gutt's index)	Doubblecheck with DW	isi
2 hour glucose tolerance test (0 min plasma glucose)		p_lab_pglu_0
2 hour glucose tolerance test (30 min plasma glucose)		p_lab_pglu_30

Variables from ADDITION-PRO dataset	Available at DST	variable name
2 hour glucose tolerance test (120 min plasma glucose)		p_lab_pglu_120
2 hour glucose tolerance test (0 min GIP)		p_lab_gip_0
2 hour glucose tolerance test (30 min GIP)		p_lab_gip_30
2 hour glucose tolerance test (120 min GIP)		p_lab_gi~120
2 hour glucose tolerance test (0 min GLP-1)		p_lab_glp1_0
2 hour glucose tolerance test (30 min GLP-1)		p_lab_glp~30
2 hour glucose tolerance test (120 min GLP-1)		p_lab_gpl~120
2 hour glucose tolerance test (0 min insulin response)		p_lab_insu_0
2 hour glucose tolerance test (30 min insulin response)		p_lab_insu_30
2 hour glucose tolerance test (120 min insulin response)		p_lab_in~120
Kidney function markers	ask marianne	
Mean heart rate (bpm)		
Median heart rate (bpm)		
Heart rate variability indices		
• Standard deviation of Normal to Normal intervals (ms)		
• SDANN		
• SDNNi		
• HRVi		
• TINN		
Medication use?		
• Glucose lowering medication		
• Cardioprotective medication		
– antihypentensive		
∗ types of antihypertensive		
• lipid lowering medication		
Variables from DST - follow up on		
cardiovascular disease events		
Myocardial infarction	from 2022	
Stroke	from 2022	

Variables from ADDITION-PRO dataset	Available at DST	variable name
Heart failure	from 2022	
All-cause mortality	from 2022	

Table 1 Population characteristics by diabetes status

	Study population in ADDITION-PRO
Age (years)	
Sex (male)	
Socioeconomic status	
Smoking status	
Physical activity (PAEE kj_kg_day)	
Alcohol comsuption (units per week)	
BMI (kg/m^2)	
Waist circumference (cm)	
Fat percentage (%)	
Systolic blood pressure (mm hg)	
Diastolic blood pressure (mm hg)	
HbA1c (%)	
LDL cholestorol (mmol/L)	
HDL cholesterol (mmol/L)	
Triglycerides (mmol/L)	
Mean heart rate (bpm)	
Median heart rate (bpm)	
Standard deviation of Normal to Normal intervals (ms)	
Root mean squared of successive RR intervals (ms)	*Not possible*
pNN50 (ms)	*Not possible*
SDSD (ms)	
High frequency (ms^2)	*Not possible*
Low frequency (ms^2)	*Not possible*
Low frequency / high frequency (ratio)	*Not possible*
Ultra low frequency (ms^2)	*Not possible*
Ultra high frequency (ms^2)	*Not possible*

Table 2 CVD and mortality incident in study population

	Incidents CVD (cases per 1000 year)	Incidents mortality (cases per 1000 year)
All		
1st tertile HRV		
2nd tertile HRV		
3rd tertile HRV		

Table 3 Heart rate variability indices associated with CVD

	Model 1: IRR (95% CI)		**Model 2: IRR (95% CI)**	
	CVD events	All cause mortality	CVD events	All cause mortality
	n=x		n=x	
SDNN (per unit)	0.xx (0.xx - 0.xx)	0.xx (0.xx - 0.xx)	0.xx (0.xx - 1.xx)	0.xx (0.xx - 1.xx)
mean HR (per unit)	0.xx (0.xx - 0.xx)	0.xx (0.xx - 0.xx)	0.xx (0.xx - 0.xx)	0.xx (0.xx - 0.xx)

Figure 1

Flow chart

Figure 2 (example)

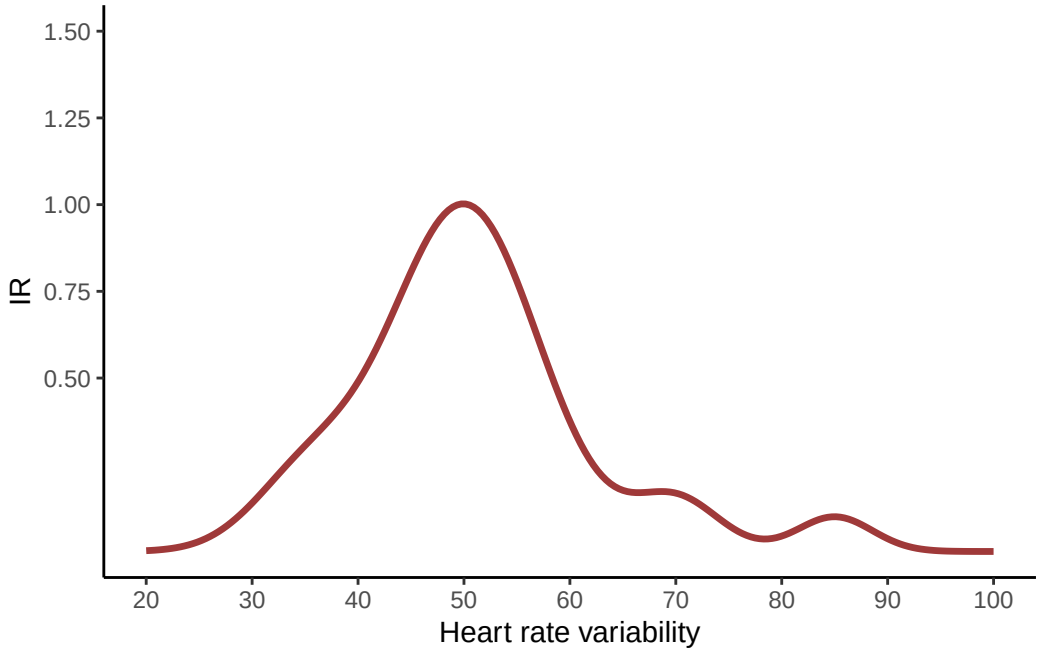
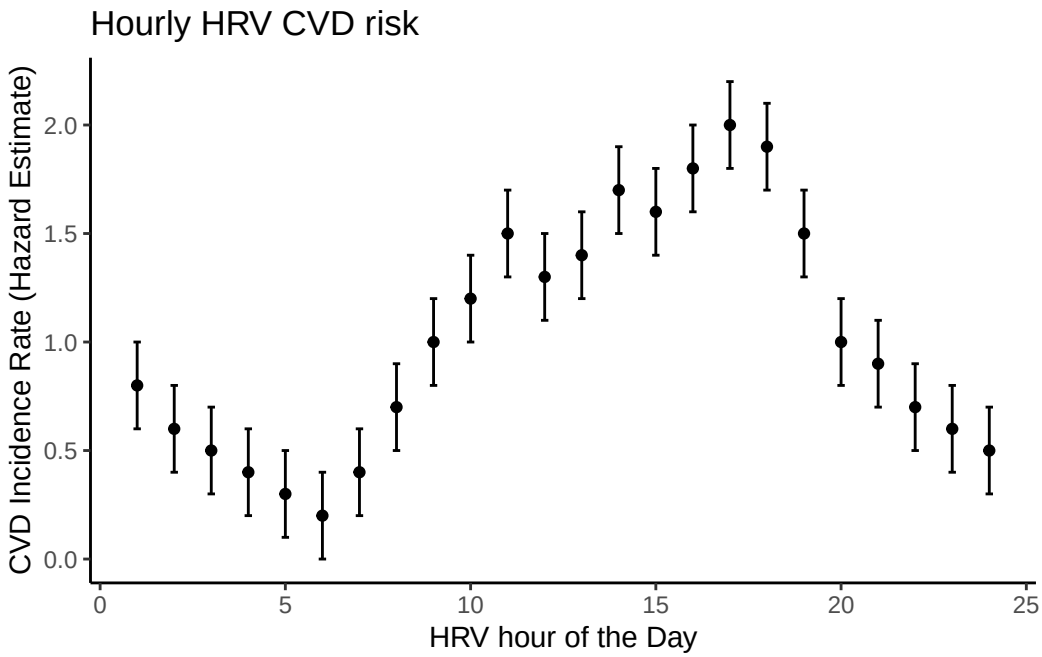


Figure 3 **(example)**


Hourly HRV CVD risk
CVD Incidence Rate (Hazard Estimate)
0.0
0.5
1.0
1.5
2.0
0
5
10
15
20
25
HRV hour of the Day