Royal institute of science,
Stockholm, Sweden

Master's Thesis at the division of Mathemathical statistics

Logistic regression modelling for STHR analysis

Author:
Johan Olsén

Examinor:
Gunnar Englund

Supervisor:
Hans Hedenström

August 12, 2014

Abstract

Coronary artery heart disease (CAD) is a common condition which can impair the quality of life and lead to cardiac infarctions. Traditional criteria during exercise tests are good but far from perfect. A lot of patients with inconclusive tests are referred to radiological examinations. By finding better evaluation criteria during the exercise test we can save a lot of money and let the patients avoid unnecessary examinations.
Computers record amounts of numerical data during the exercise test. In this retrospective study 267 patients with inconclusive exercise test and performed radiological examinations were included. The purpose was to use clinical considerations as-well as mathematical statistics to be able to find new diagnostic criteria.
We created a few new parameters and evaluated them together with previously used parameters. For women we found some interesting univariable results where new parameters discriminated better than the formerly used. However, the number of females with observed CAD was small (14) which made it impossible to obtain strong significance.
For men we computed a multivariable model, using logistic regression, which discriminates way better than the traditional parameters for these patients. The area under the ROC curve was 0.90 (95 % CI: 0.83-0.97) which is excellent to outstanding discrimination in a group initially included due to their inconclusive results.
If the model can be proved to hold for another population it could contribute a lot to the diagnostics of this common medical conditions.

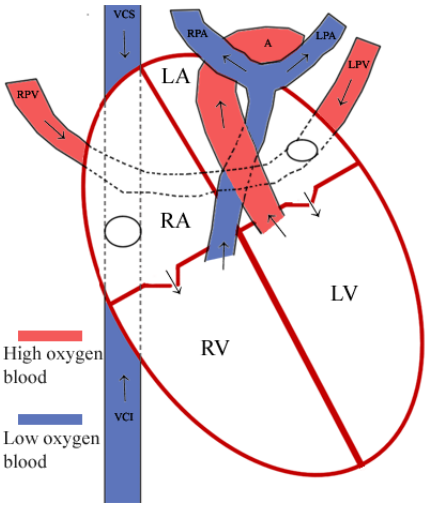
Contents

1 Introduction

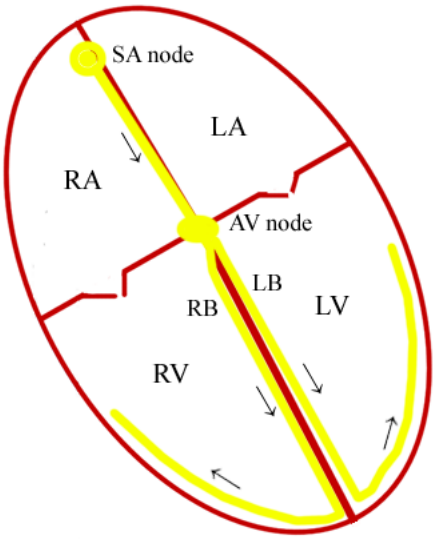
This thesis is written as a part of my master of science degree at the Royal Institute of Technology (Swedish: abbreviated KTH) in Stockholm, Sweden. But it will also be credited as a part of my medical degree at Uppsala University. I'll start with a short presentation of the medical background for mathematically oriented readers.

1.1 Medical background

The main function of the heart is being a mechanical pump. The tissues of our body consumes oxygen and produces carbon dioxide. In the capillaries of the lungs we'll be able to exchange the carbon dioxide for new oxygen. To be able to supply enough oxygen the blood has to continuously circle between lungs and other tissue. The heart is responsible for keeping the blood in motion. In addition to transport of oxygen the blood is responsible for transporting a lot of other substances.
The heart consists of 4 chambers, see figure 1a. The blood arrives to the atriums and leaves the heart through the ventricles. To be able to maintain an effective pump function the motions of the heart need to be highly coordinated. For this the heart got an electrical conduction system (see figure 1b) which determines the contractile behaviour.



(a) Heart. Schematic picture of anatomy and circulation. LA = left atrium, LV = left ventricle, RA = right atrium, RV = right ventricle, A = aorta, VCS = vena cava superior, VCI = vena cava inferior, RPV = right pulmonary vein, LPV = left pulmonary vein, RPA = right pulmonary artery, LPA = left pulmonary artery



(b) Schematic picture of the electrical conduction system. SA node = Sinoatrial node, AV node = Atrioventricular node, LB = left bundle, RB = right bundle, LA = left atrium, LV = left ventricle, RA = right atrium, RV = right ventricle

Figure 1

The heart itself needs oxygen for its hard work. It's not possible the use the blood contained by the chambers. Instead the heart is supplied by the coronary arteries who arises from the aorta. If the coronary arteries doesn't deliver enough fuel the electrical conduction and contractions will have a decreased capacity. If a coronary artery suddenly is occluded the patient will have an infarction.
But if the occlusion is slowly progressing the patient may have symptoms when the body has

increasing demands and the heart should increase its output. This will typically give chest pains during exercise and is a risk factor for future infarctions.
One way to examine the heart is electrocardiography (ECG). This will typically record the electrical activity in 12 different directions (12 leads called aVL, aVR, aVF, I, II, III, V1, V2, V3, V4, V5, V6). One part of the ECG curve, the ST-segment, is particularly interesting for evaluation of coronary artery disease (CAD).

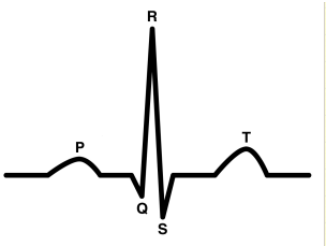


Figure 2: ECG

At the hospital a patient with suspected CAD is primarily investigated by a exercise test. The patient will cycle for a few minutes while a ECG is recorded. After the exercise phase the ECG is continuously recorded for a few minutes of the following recovery. If the analyse of the ST-segment is inconclusive the patient can be referred to a radiological examination (scintigraphy or angiography) for further investigation.

1.2 Purpose

Previously doctors evaluated ECGs manually by inspecting the curve and performing some measures with a ruler. This is still a golden standard but during exercise tests data are recorded by computers. Data for heart rate and ECG levels are recorded several times every minute. This has lead to new opportunities to find more sophisticated diagnostic criteria. The purpose of this study is to find new criteria to be able to minimize the number of inconclusive exercise test and thereby also radiological examinations. This would save the hospital a lot of money and let patients avoid extra investigations.

1.3 Study population

The subject are chosen since the result of their exercise tests were inconclusive. They where referred to radiological examination for better diagnostics.
The total number of subjects were 267 (147 men and 120 women). 46 of these where excluded due to inconclusive radiological examination. The properties for the study population are presented in table 1.

	Observed CAD	Observed no CAD	Radiologically inconclusive	total
Men	50	68	29	147
Women	14	89	17	120

Table 1: Study population

1.4 Data preparation

The raw data is delivered by the computer system as an XML file. To extract and calculate the parameters previously described (see section 2.9) I use a couple of applications and Microsoft Excel macros. These are described in detail in appendix section A.

2 Parameters

2.1 ST depression

Definition: The lowest ST level (unit μV) at the end of the work phase among all leads (aVL and V1 excluded), *(see figure 3 on page 6)*.

- If the ST level is positive at the beginning of the work phase then the ST depression is defined as the absolute deviation from zero.
- If the ST level is negative at the beginning of the work phase the ST depression is defined as the deviation from the starting value.

This measure is independent of heart rate and whatever happens during the recovery phase. ST depression is the most widely used parameter for evaluating ischemic coronary artery disease during exercise test. At Uppsala university hospital (Swedish: "Akademiska sjukhuset") the diagnostic criteria are:

- ST depression $>$ 1 mm among lateral leads (V5-V6) or
- ST depression $>$ 1.5 mm among inferior leads (aVF-III)

ST depression is related to the level of cardiac effort. This makes the ST depression criteria depending on the patient's achievement during the test. If he/she doesn't reach maximum effort the ST depression will be underestimated. To roughly deal with this problem physicists use to evaluate the double product.

$$\text{Double product} = \text{maximum heart rate} * \text{maximum systolic blood pressure} \tag{1}$$

This is an indicator of the heart's work load. If the value is less than 25 000 the exercise test will be considered unreliable.

Previous research. ST depression has been extensively analysed and the results varies between studies. A huge meta-analysis [Gianrossi, Detrano, Mulvihil, Lehmann, Dubach, Colombo, McArthur, and Froelicher, 1989] found that the mean sensitivity was 65 % (range 23-100 %, SD 16 %) and the mean specificity was 77 % (range 17-100 %, SD 17 %) among 147 published reports. In later years an extensive swedish report [Kronander, Fischer-Colbrie, Nowak, Brodin, and Elmqvist, 2010] showed similar numbers.[1]

Final parameter(s):

- The greatest ST depression among all leads (aVL and V1 excluded). (stDep)

2.2 ST/HR index

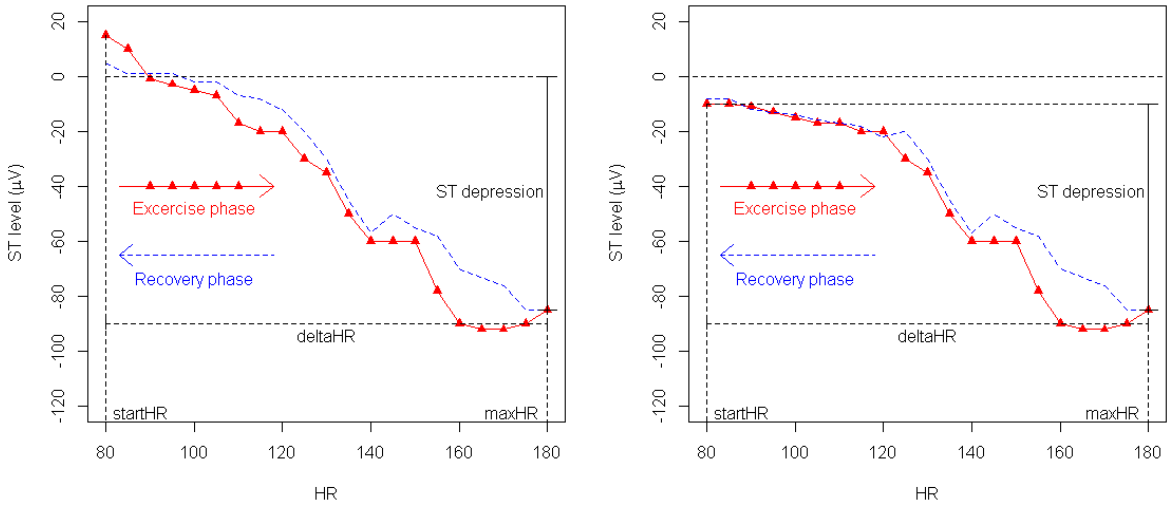
During the years scientists tried to find more robust parameters that are less dependent of the patient's achievement, early in this process "ST/HR index" (unit $\mu V/bpm$) was introduced [Detrano, Salcedo, Passalacqua, and Friis, 1986].

Definition: *(also see figure 3 on page 6)*:

$$\text{ST/HR index} = \frac{\text{ST depression}}{\Delta\text{HR}} \tag{2}$$

In this way a ST depression during a low effort exercise test (i.e. minor heart rate rise) will be given more significance.

[1] Sensitivity/Specificity (%/%): 64/83 (men), 65/71 (women)



(a) Positive ST at the beginning

(b) Positive ST at the beginning

Figure 3: ST depression and ST/HR index

Previous research: A review article [Okin, Kligfield, Johansson, Pahlm, and Brudin, 1995] concluded that the ST/HR index appeared to approach the performance of the ST/HR slope (described in section 2.3) for identification of coronary disease only, not for assessment of its anatomic severity. Kronander et al. [2010] found a better diagnostic performance [2] compared to ST depression. The partition value in these studies is about 1.6 $\mu V/bpm$.
Final parameter(s):

- The ST depression divided by the heart rate change. (sthrInd)

2.3 ST/HR slope

ST/HR index evolved as a simplification of the ST/HR slope (unit $\mu V/bpm$) which was introduced a couple of years before [Elamin, Mary, Smith, and Linden, 1980].
Definition: The ST/HR slope is the minimum (steepest negative) slope including the end part of the work phase which has a significant correlation.
Method of calculation (as described by Okin et al. [1995]):

1. Begin with the last 3 data points of the work phase.
2. Compute a regression line (least squares method) and calculate its slope.
3. Calculate the correlation coefficient and its statistical significance. (*see equation 36 and 37, on page 43*).
4. Repeatedly add one data point towards the beginning of the work phase (i.e. last 4 points followed by last 5 points etc.) until the whole work phase is included (*see figure 4 on page 7*). Calculate the slope and significance of correlation for each data set.

[2] Sensitivity/Specificity (%/%): 78/78 (men), 72/72 (women)

5. The minimum slope (steepest negative) with statistically significant correlation ($p < 0.05$) is chosen as the ST/HR slope for that lead.

6. This is done for each lead.

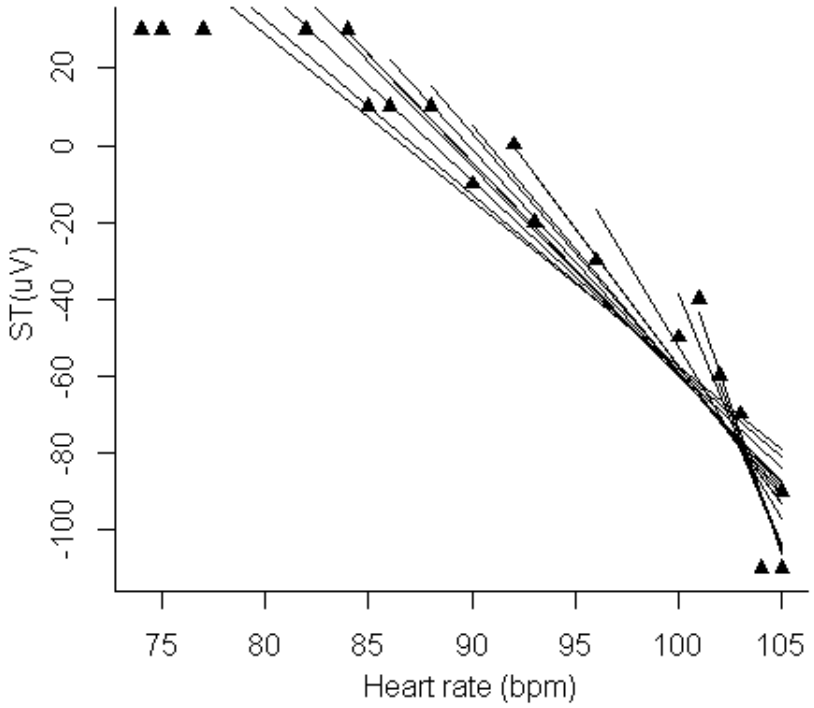


Figure 4: ST/HR slope. One data point added for each regression line. The steepest negative slope with statistic significant correlation is the "ST/HR slope"

Previous research: In previous studies the partition value has been about 2.4-2.5 $\mu V/bpm$. Kronander et al. [2010] showed better diagnostic values [3] compared to the parameters described above.

Final parameter(s):

- Minimum ST/HR slope among all leads (aVL and V1 excluded). (sthrSlope)
- Minimum ST/HR slope among extremity leads (aVL excluded). (sthrSlope_e)
- Minimum ST/HR slope among chest leads (V1 excluded). (sthrSlope_c)

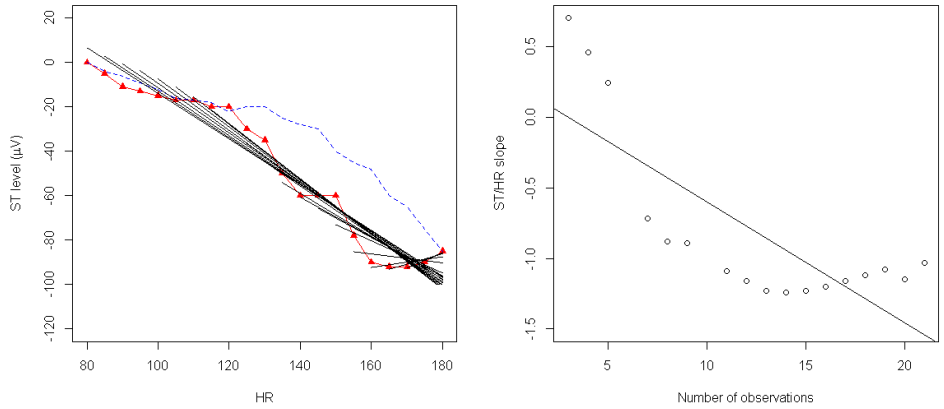
2.3.1 ST/HR recovery slope

Definition: The ST/HR recovery slope is a similar parameter for the recovery phase, which measures the slope of the curve at the beginning of the recovery phase. During this phase can a positive slope be interpreted as a decreasing ST level (since the heart rate decreases). To calculate the recovery slope:

1. Begin with the first 3 data points of the recovery phase.

2. Compute a regression line (least squares method) and calculate its slope.

[3] Sensitivity/Specificity (%/%): 82/82 (men), 75/75 (women)



(a) ST/HR diagram (b) ST/HR acceleration factor

Figure 5: Example of ST/HR acceleration factor. (a): ST depression arises early and then the curve flattens out. That makes the ST/HR slope close to zero when only the last observations are included in the regression. When more observations are included the slope becomes steeper. The slopes and number of data points included are plotted in (b). The slope of the regression line represents the ST/HR acceleration factor, which in this example is negative.

3. Calculate the correlation coefficient and its statistical significance. If $p < 0.05$ the slope is chosen as the ST/HR recovery slope.
4. If the correlation isn't significant you continue and add one data point towards the end of the recovery phase. This is done repeatedly until you find a significant correlation.
5. This is done for each lead.

Final parameter(s):

- Maximum ST/HR recovery slope among all leads (aVL and V1 excluded). (sthrRecSlopeMax)
- Minimum ST/HR recovery slope among all leads (aVL and V1 excluded). (sthrRecSlopeMin)

2.4 ST/HR acceleration factor

The ST/HR slope acceleration factor is a parameter to further quantify the shape of the ST/HR curve. During the calculations for the ST/HR slope in section 2.3 we began with the last 3 data points and then added one data point until the whole work phase was included. The minimum value with a significant correlation was chosen as the ST/HR slope. In this part we're going to use all those calculated slopes (if the correlation is significant).

Method of calculation:

- Plot the number of data points included on the x-axis and the corresponding slope on the y-axis (use the values calculated before). *See figure 5 on page 8*
- Make a regression line (least squares method). The slope of this line is the acceleration factor. The p-value for the correlation will be controlled in this case as-well.

Objectives: A positive value can be interpreted as an accelerating ST depression towards the end of the work phase. If the ST depression would appear early during the exercise the

acceleration factor will be less positive (or negative) even if the ST/HR slope is big. Since this haven't been studied before (what I know about) I would like to investigate however this additional information can increase the diagnostic value of the ST/HR slope. I will analyse the minimum and maximum values among extremity and chest leads.
Final parameter(s):

- Minimum acceleration factor among extremity leads (avL excluded). (AccFacE_{min})
- Maximum acceleration factor among extremity leads (avL excluded). (AccFacE_{max})
- Minimum acceleration factor among chest leads (V1 excluded). (AccFacC_{min})
- Maximum acceleration factor among chest leads (V1 excluded). (AccFacC_{max})

2.4.1 ST/HR recovery acceleration factor

This is the equivalent parameter for the recovery phase. A positive value indicates and early recovery. My theory is that a healthy subject will have a faster recovery and thereby a larger value than someone with CAD. The parameter is exemplified in figure 6.
Final parameter(s):

- Minimum recovery acceleration factor among extremity leads (avL excluded). (recAccFacE_{min})
- Maximum recovery acceleration factor among extremity leads (avL excluded). (recAccFacE_{max})
- Minimum recovery acceleration factor among chest leads (V1 excluded). (recAccFacC_{min})
- Maximum recovery acceleration factor among chest leads (V1 excluded). (recAccFacC_{max})

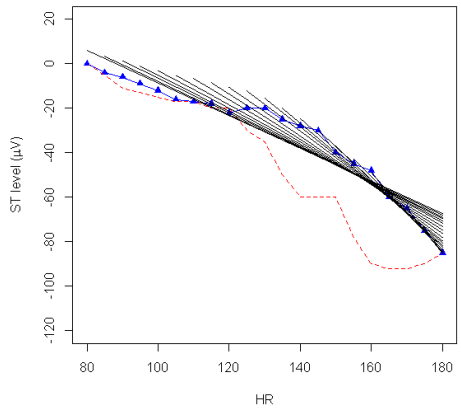
2.5 ST/HR hysteresis

The ST/HR hysteresis will not be analysed in this study. But I'll describe it shortly to give a more complete picture of ST/HR knowledge. All traditional parameters above concentrate on just the work phase of the exercise test. None of them tells us anything about the recovery phase. Taking the recovery phase into account has been shown to increase the diagnostic value [Lehtinen, Sievnen, Vii, Turjanmaa, Niemel, and Malmivuo, 1996], [Svensbergh, Johansson, Pahlm, and Brudin, 2004], [Kronander et al., 2008], [Kronander et al., 2010]. The ST/HR hysteresis was introduced by Lehtinen et al. [1996].
Definition: The ST/HR hysteresis is obtained by integrating the distance between the work phase curve and the recovery phase curve during the first 3 minutes of recovery.
Previous research: This was shown to be the parameter with best prognostic value by Kronander et al. [2010][4]

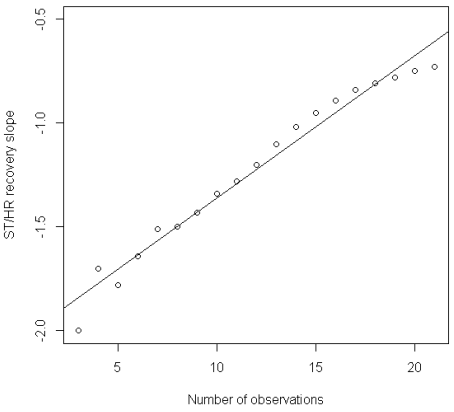
2.6 Normalized Area (NA)

The normalized area is i similar measure to the ST/HR hysteresis. It was introduced by Svensbergh et al. [2004]. The main difference is that the included data subset is variable and relates to the heart rate, compared to the ST/HR hysteresis where it always comprises 3 constant minutes.
Definition: The normalized area is the area between the work phase curve and the recovery phase curve, divided by the heart rate interval. The loop area is defined as positive if the loop is

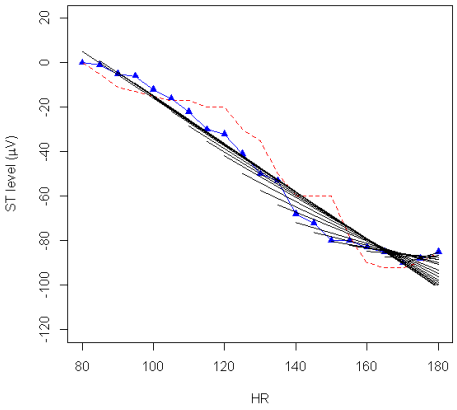
[4]Sensitivity/Specificity (%/%): 83/83 (men), 79/81 (women)



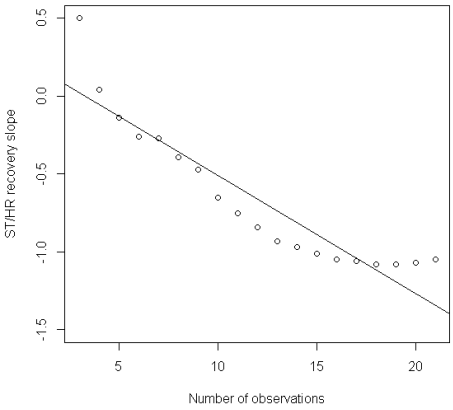
(a) ST/HR diagram, example 1



(b) ST/HR recovery factor, example 1



(c) ST/HR diagram, example 2



(d) ST/HR recovery factor, example 2

Figure 6: Examples of ST/HR recovery acceleration factor. (a): Early recovery which decreases towards the end. With few data point included this gives a steep negative slope which flattens out when more points are included in the regression. The slopes and number of data points included are plotted in (b). The slope of the regression line represents the ST/HR recovery acceleration factor, which in this example is positive. (c and d): Another example with slow initial recovery which leads to a negative recovery acceleration factor.

counter clockwise and negative if it is clockwise. See figure 7 on page 11 which shows a positive NA loop.
The following equations are used in order to calculate the normalized area:

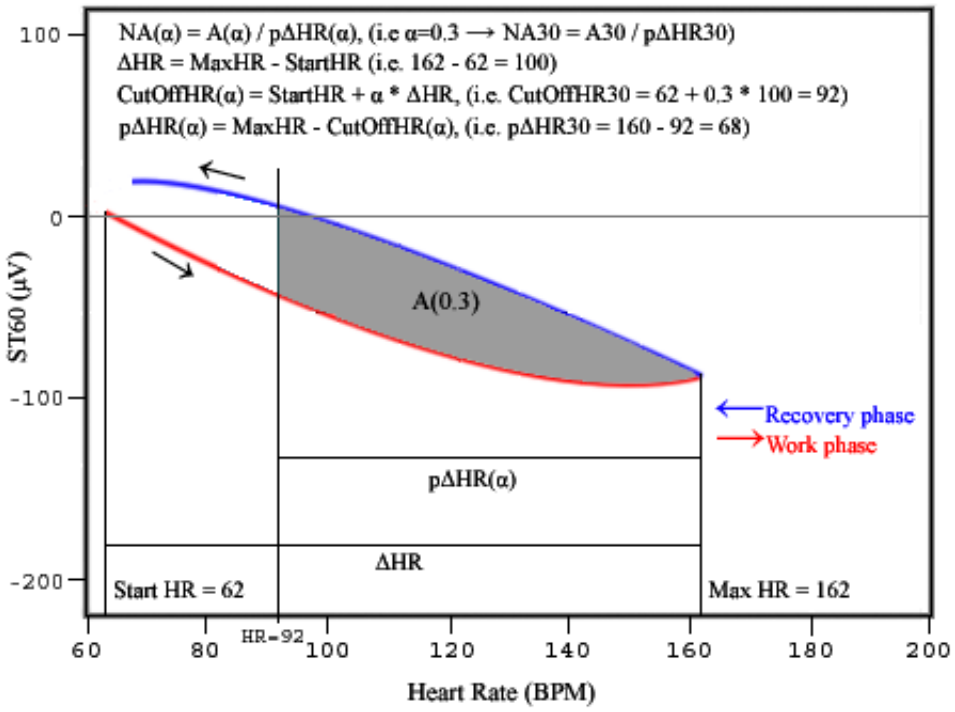


Figure 7: Normalized area (NA).

$$NA(\alpha) = \frac{A(\alpha)}{p\Delta HR(\alpha)} \tag{3}$$

$$p\Delta HR(\alpha) = (1-\alpha) * \Delta HR \tag{4}$$

$$\Delta HR = MaxHR - StartHR \tag{5}$$

$$CutOffHR(\alpha) = StartHR + \alpha * \Delta HR \tag{6}$$

Previous research: Svensbergh et al. [2004] tried to optimize the normalized area by finding the best α-value using discriminant analysis. The result was that $\alpha = 0.3$ is the optimal choice. By combining the minimum value from the extremity leads and the chest leads (i.e. $NA_{0.3}(e+c)$) they reached good diagnostic levels[5] which where even better when further combined with the maximum ST slope [6].
In this report I've chosen to concentrate on the normalized area instead of the ST/HR hysteresis. The main reasons for this is:

- We have a close cooperation with one of the authors (Brudin) of the previous article [Svensbergh et al., 2004] which helps us with new perspectives. I'm also provided with the software developed for their research.
- The algorithm behind the software seems robust.

[5]Sensitivity/Specificity (%/%): 89/77 (men), 83/76 (women)
[6]Sensitivity/Specificity (%/%): 84/88 (men), 83/85 (women)

NPA(α, β)	α	β
NPA(0.2, 0.3)	0.2	0.3
NPA(0.3, 0.4)	0.3	0.4
NPA(0.4, 0.5)	0.4	0.5
NPA(0.5, 0.6)	0.5	0.6
NPA(0.6, 0.7)	0.6	0.7
NPA(0.7, 0.8)	0.7	0.8
NPA(0.8, 0.9)	0.8	0.9
NPA(0.9, 1.0)	0.9	1.0

Table 2: NPAs used in this report

- None of the articles I've read about the ST/HR hysteresis has given an exact explanation of how it is obtained.

Objectives: The idea is to further develop this parameter by approaching it in new ways. I will also perform a thoroughgoing analysis of the NA in order to reproduce the previous results [Svensbergh et al., 2004].
Final parameter(s):

- The minimum (greatest negative) loop area among all leads (aVL, III and V1 excluded) for 8 different α-values (0.3, 0.4,..., 0.9). ($NA(\alpha)$)

2.7 Normalized Partial Area (NPA)

Definition: The normalized partial area is the area of the ST/HR loop between 2 heart rate levels, divided by the heart rate interval. The cut-offs are decided by α and β in the same way as in section 2.6 (where $\beta = 1$). *See figure 8 on page 13.*
In other words this is a measure where the normalized area is cut into pieces. This is done for one extremity lead and one chest lead. The chosen leads are those with the minimum (i.e. most negative) value of NA($\alpha = 30$), (aVL, III and V1 are still excluded).
To be able to calculate the NPAs I need the areas in absolute numbers (de-normalized):

$$NA(\alpha) = \frac{A(\alpha)}{(1-\alpha) * \Delta HR} \rightarrow A(\alpha) = (1-\alpha) * \Delta HR * NA(\alpha) \tag{7}$$

Then I can use the areas in the following equation:

$$NPA(\alpha, \beta) = \frac{A(\alpha) - A(\beta)}{(\beta - \alpha) * \Delta HR} \tag{8}$$

In this report the loop area is cut into 8 pieces according to table 2 on page 12.
Previous research: As far as I know this parameter hasn't been introduced before.
Objective: The idea behind this parameter is finding a way to characterize the ST/HR loop in new ways. I hope to find patterns in the loop that will help us to increase the diagnostic value of exercise tests for patients with suspected ischemic coronary heart disease. In order to further characterize the ST/HR loop I'll introduce a couple of new parameters connected to the NPAs (*figure 26 on page 42 shows an example for a patient*):

2.7.1 Minimum value

Definition: The minimum value among the NPAs.
This tells us how negative the ST/HR loop is as most.

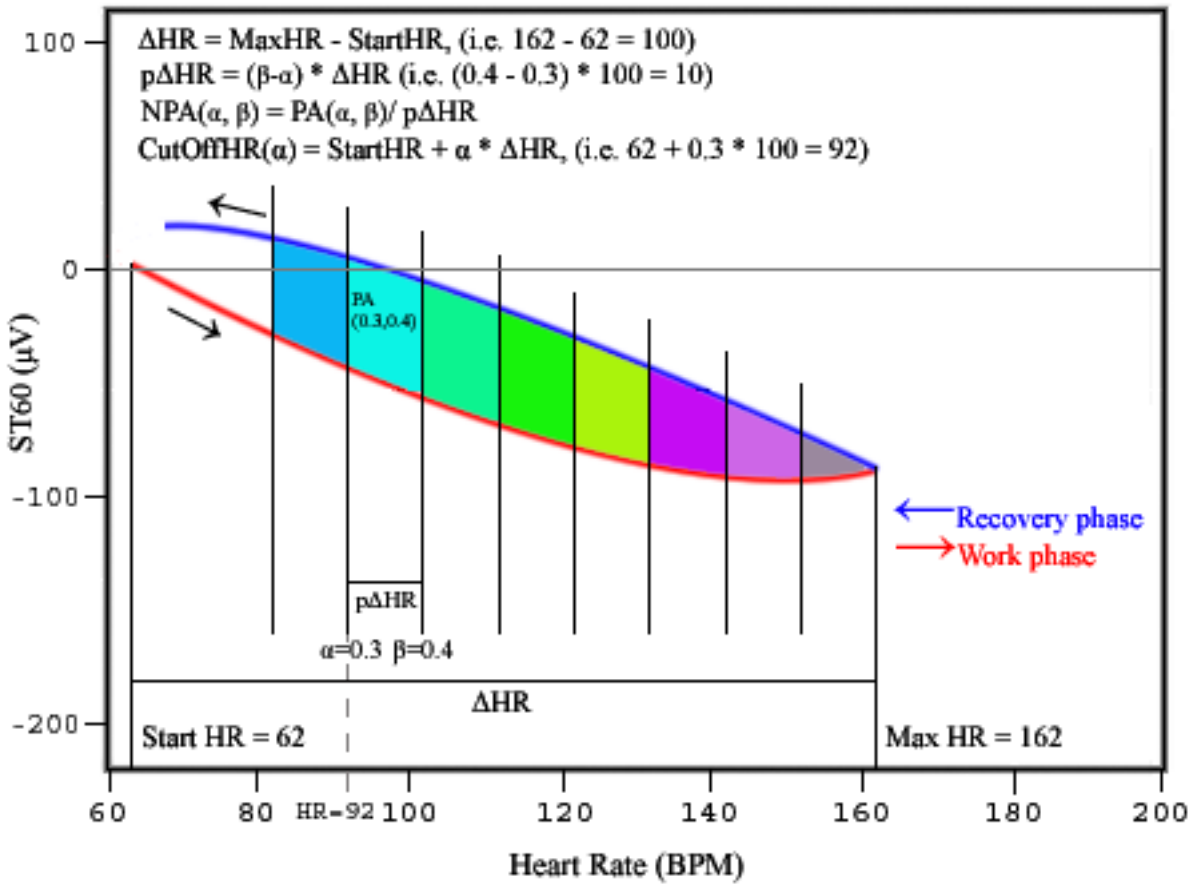


Figure 8: Normalized partial area

- The extremity lead (aVL excluded) and the chest lead (V1 excluded) with minimum NA($\alpha = 0.3$) are chosen. The reason for this is that NA($\alpha = 0.3$) is the best discriminating area measure according to previous research [Svensbergh et al., 2004].
- Each lead contain 8 NPA values and the minimum of these 16 values is chosen as the minimum NPA parameter.

2.7.2 NPA slope

Definition: The NPA slope is the slope of the regression line (least squares method) for the NPA values (see figure 9 on page 14). This is also calculated for the 2 leads with minimum NA($\alpha = 0.3$).
A positive NPA slope means that the ST/HR loop tend to be larger (less negative) in the region close to maximum heart rate compared to the region with slow heart rate.
Final parameter(s):

- The minimum (greatest negative) normalized partial loop area among extremity all leads (aVL, III and V1 excluded) for 8 different α-values (0.3, 0.4,..., 0.9). ($NA(\alpha)$)
- Minimum NPA. (npaMin)
- NPA slope chest. (npaSlopeC)
- NPA slope extremity. (npaSlopeE)

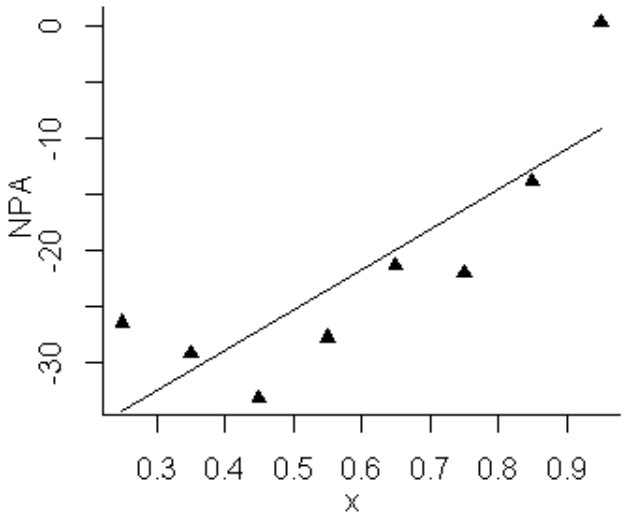


Figure 9: NPA slope. NPA values are plotted on the y axis and $x = \frac{\alpha+\beta}{2}$ on the x axis. The NPA slope is the slope of the regression line. I.e MinNPA = -33.3, PeakLevel = 0.45, NPAslope = 36.3, NPAsd = 10.7

2.8 ST/HR Cross

Definition: This is a parameter describing if and where the ST/HR curves (excercise and recovery) cross. If they never cross the recovery phase ST level will be above (positive no cross) or below (negative no cross) the work phase ST level during the whole test. If it's below the normalized area certainly will be negative. If they cross the normalized area can be either positive or negative. One theory is that the crossing point could be an important parameter for predicting CAD. If the curves cross multiple times the last cross (i.e. the lowest cross level) is chosen as parameter for the analysis. The curve is analysed for the chest lead (V1 excluded) which has the minimum normalized area for $\alpha = 0.3$. This is because it's considered reasonable for both sexes (see section 4). *See figure 10.*

$$\text{crossLevel} = \frac{\text{crossHR} - \text{startHR}}{\text{deltaHR}} \tag{9}$$

Final parameter(s):

- Positive no cross (yes/no). (posCross)
- Negative no cross (yes/no). (negCross)
- Cross level. (crossLevel)

2.9 Summery of parameters

This is a summery of the parameters that are calculated and analysed:

- ST depression
  - i The greatest ST depression among all leads (aVL and V1 excluded). (STdep)
- ST/HR index
  - i The lowest value of ST/HR index among all leads (aVL and V1 excluded). (sthrIndex)

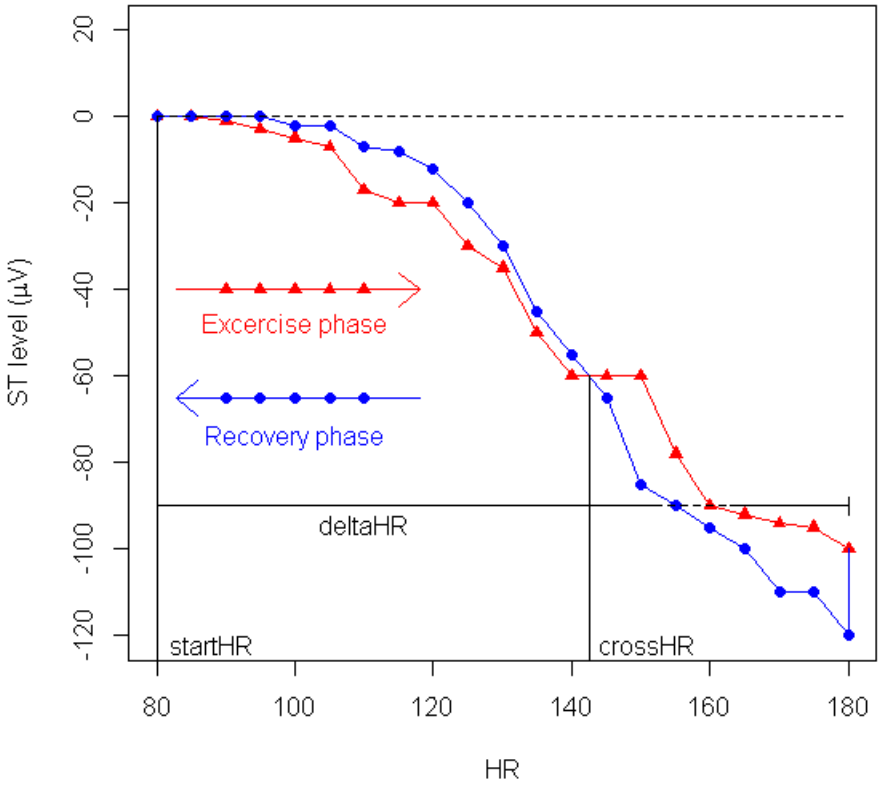


Figure 10: ST/HR curves for exercise and recovery phases. I.e. crossLevel = (143-80)/100 = 0.63

- ST/HR slope
  - i Minimum ST/HR slope among extremity leads (aVL excluded). (sthrSlope$_e$)
  - ii Minimum ST/HR slope among chest leads (V1 excluded). (sthrSlope$_c$)
  - iii Minimum ST/HR slope among all leads (aVL and V1 excluded). (sthrSlope)
  - iv Maximum ST/HR recovery slope among all leads (aVL and V1 excluded). (sthrRecSlopeMax)
  - v Minimum ST/HR recovery slope among all leads (aVL and V1 excluded). (sthrRecSlopeMin)
- ST/HR acceleration factor
  - i Minimum acceleration factor among extremity leads (avL excluded). (AccFacE$_{min}$)
  - ii Maximum acceleration factor among extremity leads (avL excluded). (AccFacE$_{max}$)
  - iii Minimum acceleration factor among chest leads (V1 excluded). (AccFacC$_{min}$)
  - iv Maximum acceleration factor among chest leads (V1 excluded). (AccFacC$_{max}$)
- ST/HR recovery acceleration factor
  - i Minimum recovery acceleration factor among extremity leads (avL excluded). (recAccFacE$_{min}$)
  - ii Maximum recovery acceleration factor among extremity leads (avL excluded). (recAccFacE$_{max}$)

iii Minimum recovery acceleration factor among chest leads (V1 excluded). (recAccFacC$_{min}$)

iv Maximum recovery acceleration factor among chest leads (V1 excluded). (recAccFacC$_{max}$)

- Normalized loop area

  i The minimum (greatest negative) loop area among all leads (aVL, III and V1 excluded) for 8 different α-values (0.3, 0.4,..., 0.9). ($NA(\alpha)$)

- Normalized partial loop area

  i The minimum (greatest negative) normalized partial loop area among extremity all leads (aVL, III and V1 excluded) for 8 different α-values (0.3, 0.4,..., 0.9). ($NA(\alpha)$)

  ii Minimum NPA. (npaMin)

  iii NPA slope chest. (npaSlopeC)

  iv NPA slope extremity. (npaSlopeE)

- ST/HR cross

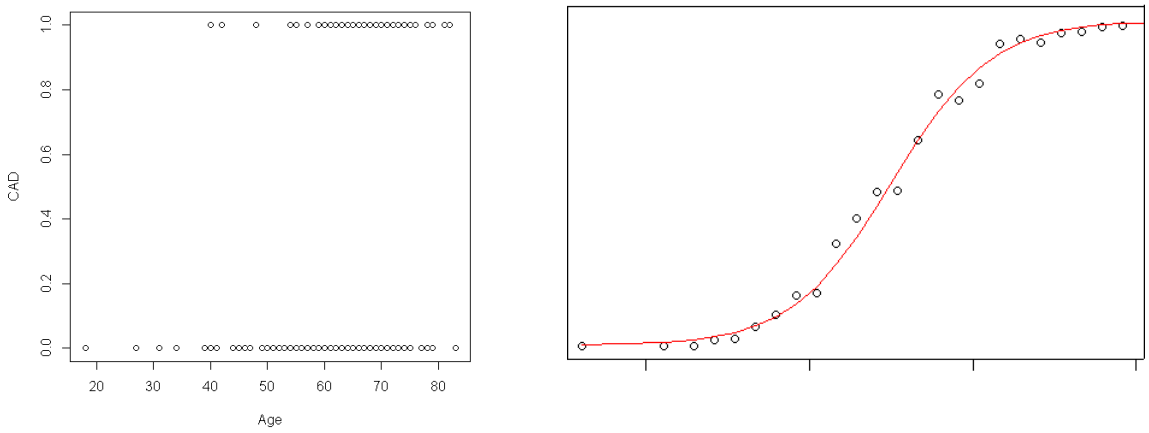
  i Positive no cross (yes/no). (posCross)

  ii Negative no cross (yes/no). (negCross)

  iii Cross level. (crossLevel)

3 Logistic regression

In this study we got a lot of independent variables (covariates) and a dichotomous dependent variable (CAD / no CAD). Then logistic regression is a preferable method to find the best fitting and clinically interpretable model. I'll explain this method below.
If you make a regular scatter plot from a data set with a dichotomous output variable it'll look something like figure 11a where I plotted CAD as a function of age. Fitting a regression line to this would result in a nearly pointless model with enormous residuals.



(a) (b)

Figure 11

It's possible to divide the continuous independent variable into smaller intervals and calculate the CAD fraction in each age group. This fraction can be plotted against the mean age of the group, see figure 11b. The Y axis will now represent the risk for CAD in each age group. If you connect the dots it'll typically result in a s-shaped curve if the change in risk is progressively smaller when you approaches the end points (risk 0 and 1). A regression line will make more sense in this setting compared to the initial situation. But in many cases it'll be better if you make a logit transformation:
Instead of calculation the percentage of CAD in each age group you begin by calculation the odds for CAD:

$$\text{Odds} = \frac{\text{Number of CAD}}{\text{Number of no CAD}} = \frac{\pi(x)}{1-\pi(x)}, \text{ where } \pi(x) = P(Y=1|x) \tag{10}$$

$$\text{logit} = g(x) = ln(\text{odds}) = ln\frac{\pi(x)}{1-\pi(x)} \tag{11}$$

In this transformation you can fit a regression line as for example:

$$g(x) = \beta_0 + \beta_1 x \tag{12}$$

When the coefficients are estimated ($\hat{\beta}_0$, $\hat{\beta}_1$) we can simply get the estimated probability as:

$$\hat{P}(Y=1|x) = \frac{e^{\hat{\beta}_0+\hat{\beta}_1 x}}{1+e^{\hat{\beta}_0+\hat{\beta}_1 x}} \tag{13}$$

Estimation of parameters by the usual least squares method is suboptimal for models with binary outcome [Hosmer, Lemeshow, and Sturdivant, 2013]. The maximum likelihood estimation is a more general method which has lower demands on data properties. This makes it the primary method for parameter estimation for logistic regression. The chosen parameters are those who maximizes the likelihood function. For simple linear regression this is equal to the least squares but for logistic regression you will obtain non-linear equations which needs iterative methods to solve. However, these calculations are carried out by computer software, so we don't have to worry further about the details of parameter estimation.

3.1 Multiple logistic regression

Consider a collection of p independent variables denoted by the vector $\mathbf{x} = (x_1, x_2, ..., x_p)$. The conditional probability is denoted by $\pi(\mathbf{x}) = P(Y = 1|\mathbf{x})$. Logit is produced in the same way as for the univariable case:

$$g(x) = ln\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p \tag{14}$$

One strength of logistic regression is the ability to handle variables on different measurement scales in the same model. Discrete variables (nominal and ordinal scale) have to be exchanged by design variables. In some cases though we can allow an ordinal scale variable i the model if it can be considered as approximately linear.
A discrete variable with k levels will be exchanged by $k-1$ dichotomous design (dummy) variables. One level will be represented by the case where all dummies equals 0, and the other levels are represented by cases where one dummy equals 1 and the rest of the are 0.

3.2 Evaluation of the esimated parameters significance

In linear regression we use SSE (sum of squared errors) as a measure of fit for the model. The corresponding measure for the likelihood estimation is deviance, D:

$$D = -2 * ln\frac{\text{likelihood for the fitted model}}{\text{likelihood for the saturated model}} \tag{15}$$

The saturated (full) model is a model with a parameter for every observation so that the data are fitted exactly. However, in a setting where the output variable is either 0 or 1 this term can be shown to equal 1. In this case the deviance is:

$$D = -2 * ln(\text{likelihood for the fitted model}) \tag{16}$$

The deviance can be used for hypothesis testing by performing a likelihood ratio test. For this purpose we introduce the difference of deviance between two models:

$$G = D(\text{model without the parameter}) - D(\text{model with the parameter}) \tag{17}$$

G will always be positive since the log-likelihood will increase when a new parameter i added (as long as the former model isn't a perfect fit). But a positive G doesn't mean that the new model is significantly better than the previous model. To investigate this we use the fact that G approximately follows the chi-square distribution with $df2 - df1$ degrees of freedom ($df1$ and $df2$ represents the number of independent parameters in the models compared). For example we may have a model with p parameters and its calculated deviance D_p. $G = D_p - D_0$, where D_0 is the deviance of the basic model without any parameters (just an intercept). Then G will follow the chi-square distribution with p degrees of freedom under the null-hypothesis that all β_i

equals zero. If the p-value indicates that we should reject the null-hypothesis we can conclude that at least one of the parameters is significant.
We can also compare two models in the same way. $G = D_p - D_{p-x}$ is the difference in deviance between the model with p parameters and a reduced model with $p-x$ parameters. The difference in degrees of freedom is x. Under the null hypothesis that the models are equally good (i.e. the coefficients of the x removed parameters are zero), G will follow the chi-square distribution with x degrees of freedom. This is called *the partial likelihood test*. If $p < 0.05$ the bigger model is significantly better and should be kept over the reduced model.
Another commonly used measure to compare models with different numbers of parameters is the *Akaike information criterion (AIC)*:

$$\text{AIC} = -2 * L + 2 * (p + 1) = D + 2 * (p + 1) \tag{18}$$

Where L is the log-likelihood, D is the deviance and p the number of parameters. The model with lowest AIC is considered as the best.
If a discrete variable is included (or excluded) all design variables should be included or removed. Keeping a subset of the design variables would mean that we have recoded the levels of the variable. The discrete variable will contribute to the model with one degree of freedom for each design variable.
Another way to consider the significance of each parameter is to check if the corresponding β_i is significantly different from zero. This can be done by calculating a confidence interval for each β_i. The standard errors, $\hat{SE}(\hat{\beta}_i)$, for the those are derived from the second derivatives of the likelihood function and their values are provided by any statistical software. First, as an complement to the likelihood ratio test, we can use the standard error to perform a Wald test:

$$W = \frac{\hat{\beta}_i}{\hat{SE}(\hat{\beta}_i)} \tag{19}$$

Under some sample size assumptions W will be normally distributed and a two-tailed p-value can be obtained from tabular data as $P(|z| > W)$, where z is a random variable following the standard normal distribution.
We can also compute the Wald-based confidence interval as:

$$\hat{\beta}_i \pm z_{1-\alpha/2} * \hat{SE}(\hat{\beta}_i), \text{ where the confidence level is } 1 - \alpha. \tag{20}$$

You can also calculate a confidence interval for $g(x)$ which standard error is the positive square root of the estimated variance of $g(x)$. By transformation of these endpoint you'll get a confidence interval for $P(Y = 1|x)$.
Note that these Wald-based confidence intervals can be sensitive if the sample size or the number of events $(Y = 1)$ is small. In those cases you'll need to compute likelihood-based confidence interval, which I won't cover here.

3.3 Adjustment

In a multivariable model a covariate can adjust the effect of another covariate. This means that the level of the parameters coefficient, β_i, is increased or decreased. Note that this is different from modulation where the effect of one covariate is varying over the levels of another covariate. Mathematically the adjustment can be written as:

- $y = \beta_0 + \beta_1 x_1$, when only one parameter is in the model.
- $y = \beta_0 + (\beta_1 + u)x_1 + \beta_2 x_2$, when a second covariate is included.

Where u denotes the adjustment of covariate 1 by covariate 2. To measure this effect we have to relate u to the initial coefficient.

$$\Delta\hat{\beta}_1\% = 100 * \frac{u}{\hat{\beta}_1} \tag{21}$$

The rule of thumb is that covariate 2 should be included in the model if $\Delta\hat{\beta}\% > 20$ even if its main effect isn't significant.

3.4 Modulation

Modulation, or interaction, is when one covariate is varying over the levels of another covariate. The simplest and most commonly used model for including statistical interaction is to include a term $x_i * x_j$, which represents the interaction between parameter i and j. For a model with two independent variables logit i formulated as:

$$g(x) = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_1 * x_2 \tag{22}$$

There is an remarkable interaction if β_3 is significantly different from zero. If an interaction term is included in the model, both main effects should be kept even if they aren't significant.

3.5 Check the assumption about linearity

The common equation for logit, $g(x) = \beta_0 + \sum \beta_i x_i$, represents a linear model that requires that the logit increases/decreases linearly as a function of each continuous variable. This is rarely totally true and we have to find those cases where the linearity assumption is unacceptable. There are a number of methods to check the eligibility of this assumption.

3.5.1 Smoothed scatter plot

Checking the linearity with a smoothed scatter plot is easy and helpful. The advantage is that when the plot looks linear, then the logit probably is linear in the covariate. But when it isn't it's hard guess how the function differ from linearity.

3.5.2 Fractional polynomials

This is a method developed by Royston and Altman [1994] which tries a set of common non-linear models and investigates which of these that makes the best fit. The logit for a univariable model can generally be described as:

$$g(x, \beta) = \beta_0 + \sum_{j=1}^{J} \beta_j * F_j(x) \tag{23}$$

Where the functions F_j are particular power functions describing x^p and J is the total number of terms (except β_0). Fractional polynomials method focuses on a subset of the infinite number of possible models equation 23 could give us. It's restricting the powers to $p \in \rho = (-2, -1, -0.5, 0, 0.5, 1, 2, 3)$, where the power 0 represents ln. This gives us:

$$F_j(x) = \begin{cases} x^{p_j} \\ x^{p_j} * ln(x), \text{ if } p_j = p_{j-1} \end{cases} \tag{24}$$

We could let the covariate enter the model with any number of terms, J, but we'll restrict this to $J \in (1, 2)$ to get a reasonably simple model. This gives us 8 possible models with $J = 1$ and

36 possible models when $J = 2$ (all distinct pairs of $p_1 \in \rho$ and $p_2 \in \rho$). Most of the calculations are performed by the statistical software which will present the best model (lowest deviance) for $J = 1$ and the best model for $J = 2$ (I will refer to these as M_1 and M_2 in the rest of this section), which leaves us the question if either of these are better than the linear model.
Before answering that question we have to consider the degrees of freedom. Royston and Altman (1994) suggests that each term in the fractional polynomials contributes with 2 degrees of freedom (i.e. 2 when $J = 1$ and 4 when $J = 2$). Thus when comparing the models (under the null hypothesis that the logit is linear in x) through a partial likelihood ratio test the difference in deviance is distributed as chi-square with:

- 1 degree of freedom comparing the linear model to M_1.
- 3 degrees of freedom comparing the linear model to M_2.
- 2 degrees of freedom comparing M_1 to M_2.

First we compare the linear model to M_2. If the test isn't significant at a 0.05 level we simply use the linear model without considering M_1. But if M_2 is significantly better than the linear model we move on and compare this model (M_2) to M_1. If this M_2 is significantly better than M_1 too we use M_2, else we chose M_1.
After selecting the 'winning' model we have to make some clinical consideration. Does this model make sense and is it simple enough to be used in real life? Sometimes it's better to use a model with slightly inferior mathematical properties in favour of a better clinical usability.

3.5.3 Spline functions

One alternative to the fractional polynomials is spline functions. I wont describe this method in detail since I decided that evaluating the fractional polynomials will suffice for this study. The primary purpose for this study is to investigate new parameters and find interesting results for future research. Evaluating both fractional polynomials and splines to find a slightly better fit would be to overdo the analysis of this material.

3.6 Assessing the fit of the model

In the previous sections I've described some important details needed to find the best possible model. When that part is done we've a chosen model assumed to be the best. But how good is it? Before adjudging the model as the final model we've the run some tests and evaluate plots to assess the fit of the model. In this section I'll describe some ways to do this.

3.6.1 Hosmer-Lemeshow Test

This is a test where data is grouped based on the estimated probabilities. The proposed technique is to use the percentiles to form ten groups (i.e. group 1 will consist of the 10 % of data with the lowest estimated probabilities). In each group you'll have to calculate the estimated number of counts where $y = 1$ (i.e. patient with observed CAD). This value is calculated as the sum of the estimated probabilities for all subjects within the group. In addition to this you also calculate the expected number of counts where $y = 0$.

$$\hat{e}_1 = \sum_{j=1}^{k} \hat{\pi}_j \text{ where k is the number of subjects within the group.} \tag{25}$$

$$\hat{e}_0 = \sum_{j=1}^{k}(1 - \hat{\pi}_j) \tag{26}$$

$$\hat{e}_1 + \hat{e}_0 = k \tag{27}$$

This is then compared to the observed number counts as (O_1 denotes observed number subjects where $y = 1$ and O_0 where $y = 0$).

$$c_i = \frac{(O_{1,i} - \hat{e}_{1,i})^2}{\hat{e}_{1,i}} + \frac{(O_{0,i} - \hat{e}_{0,i})^2}{\hat{e}_{0,i}} \text{ where i denotes group number i.} \tag{28}$$

This is finally summed for all groups:

$$\hat{C} = \sum_{i=1}^{g} c_i \text{ where g normally is 10.} \tag{29}$$

If the model is a good fit $\hat{C}$ will follow the chi-square distribution with $g-2$ degrees of freedom. This holds if the number of covariate patterns is close to the number of subject (which normally is true if the model includes at least one continuous covariate) and the expected number of counts should be at least 5 in each group.

3.6.2 Classification tables

Classification tables could especially be used if the purpose of the regression model is to predict whether a subject will have $y = 1$ or $y = 0$. You set a cut-off and classify each subject as 0 or 1 (i.e. each subject with estimated probability > 0.5 is predicted as 1 and subject with lower estimated probability as 0). These are presented in a table showing the observed and the predicted values.

	Observed		
Predicted	y=1	y=0	total
y=1	a	b	a+b
y=0	c	d	c+d
Total	a+c	b+d	

Table 3: Classification table

It's possible to calculate several interesting values from this table:

- Sensitivity $= a/(a + c)$
- Specificity $= d/(b + d)$
- Positive predictive value $= a/(a + b)$
- Negative predictive value $= d/(c + d)$
- Overall rate of correct prediction $= (a + d)/(b + c)$

3.6.3 Receiver operating characteristics (ROC)

Basically this is a plot of the sensitivity against 1-specificity for all possible cut-offs. The area under the ROC curve (AUC) is a good measure of the predictive ability of the model. Rule of thumb:

- $AUC = 0.5$: No discrimination. Flip a coin would be equally good.
- $0,5 < AUC < 0.7$: Poor discrimination.
- $0,7 < AUC < 0.8$: Acceptable discrimination.
- $0,8 < AUC < 0.9$: Excellent discrimination.
- $AUC > 0.9$: Outstanding discrimination.

3.6.4 Assessment via external validation

A good method for validation is if it's possible to divide the data into two subsets: One for modelling and one for validation. This means that you build your model entirely based on a subset of the data and then you run the assessment tests on the other subset.

3.7 Building the model

To start with we got a lot of potential parameters to include in our model, but these have to be reduced to a final set of parameters which will provide the best possible model. This part requires both clinical and statistical considerations. Ideally each parameter should be clinically considered as meaningful and then statistically confirmed as significant. If you clinically consider a parameter as important you probably should include it even if it doesn't reach the preferred level of significance. And vice versa you should refuse significant variables if they don't make any clinical sense. Remember that overall purpose of the study is to find a model that will provide a extra tool for the clinic.
There are several ways and methods to build the model, of which some are provided in statistical software. In the following sections I'll describe some of the techniques available and then I'll decide how to handle my data.

3.7.1 Purposeful selection

This method is described by Hosmer et al. [2013] and it's the most manual of those described here. I consider it educative the understand this procedure and it favours clinical considerations. In the more automated methods described below the model is chosen for statistical reasons only (obviously the computer can't make clinical considerations). The purposeful selection consists of 7 steps which are described below.

- Step 1. Univariable analysis. Making a series of univariable analysis is usually not very good since adjustment and interaction isn't taken into account. But as a screening instrument we use it in the first step of purposeful selection. For categorical variables this can be done through a 2*k contingency table, where 2 represents the levels of the outcome variable and k the levels of the independent categorical variable. This is followed by a chi-square (Likelihood ratio chi-square and Pearson's chi square test will give similar results). For continuous variables you can fit a univariable logistic regression and check the Wald-statistics. Alternatively you can run a simple t-test and get statistics for the difference of the means comparing the two levels of the outcome variable.

All variables with a p-value less than 0.25 will be included in the initial multivariable model.

- Step 2. Fit the multivariable logistic regression model containing all variables identified for inclusion in step 1. In this step we only include main effects (no interaction terms) and we analyse the p-value of each variable's Wald-statistics. All variables with p-values greater than 0.05 should now be eliminated and the new model compared to the old using a partial likelihood ratio test.

- Step 3. Check for adjustment. In this step we have to compare coefficients in the new smaller model to those in the larger model at the beginning of step 2. We want to investigate whether one or more of the excluded variables provided an important adjustment of some of the other variables. In any of the remaining variables got a $\Delta\hat{\beta} > 20\%$ we have to bring back the excluded variables and remove them one by one to decide which of them we should keep in the model due to their adjustment of other variables.

- Step 4. Add the variables excluded in step 1 (the univariable screening) to the model, one by one, to ensure that they still aren't significant in a multivariable environment.
  The model at the end of step 4 is referred to as the *preliminary main effects model*.

- Step 5. Check for linearity. In this step we check that the logit increases/decreases linearly as a function of each continuous variable in the model. Methods described in section 3.5.
  The model at the end of step 5 is referred to as the *main effects model*.

- Step 6. In this step we'll check for possible interactions among the covariates in the main effects model. Initially we should consider which interactions that possibly could make any clinical sense. These are then tested, one by one, by adding them to the main effects model and evaluating their Wald statistics. If we've a categorical variable with 2 or more dummy variables we add the interaction terms of another covariate and all dummy variables at once.
  When all possible interactions are tested, those who turned out to be significant (at a 0.05 level) are added together to the main effects model. If some of the interactions no longer is significant they will be removed. In this situation you could exclude and/or include terms one by one until you find the best possible model where all interactions are significant or included by well motivated clinical reasons. Note that we won't consider any main effects in this step. At this time, we'll keep them even if they are no longer significant. The model at the end of step 6 is referred to as the *preliminary final model*.

- Step 7. In this step we've to assess the fit of the model. Methods described in section 3.6.

3.7.2 Backward selection

In backward selection we start with the full model (all possible parameters included) and then we reduce the model until it can't be better. The comparison between models are based on the *Akaike information criterion (AIC)*, see section 3.2.

- To start with we got the AIC for the initial model. For each covariate the program calculates what the AIC would be if that parameter was excluded from the model.

- The exclusion that would result in the lowest AIC is performed. The AIC for the new reduced model is now compared to the possible models where another covariate is excluded.

- This procedure is repeated until no possible exclusion would result in a lower AIC. The remaining parameters are fitted in a preliminary model.

3.7.3 Forward selection

This is similar to backward selection but we're starting with an empty model and comparing its AIC to the possible models where a covariate is added to the model. This goes on until we can't add another parameter without increasing the AIC.

3.7.4 Stepwise selection

Stepwise is an combination where parameters can be both added and removed. In the initial model we can include either all or none of the parameters. The model the compares the AIC to all possible models where either one of the included parameters are excluded or one of the excluded parameter is included. This goes on until no parameter can be included or excluded without increasing the AIC.

3.8 Modelling for this study

I'll try to run a combination of the techniques described above:

- Step A. Univariable analysis. A few totally new parameters are included in this study and I want to get some preliminary results of their value. I also find the univariable results interesting by themselves. Most of the previous research is presented as comparison of variables in univariable analysis.
  All parameters with $p < 0.25$ will be included in the model in step C.

- Step B. Check the linearity assumption for each continuous parameter by using the fractional polynomials, se section 3.5.2. If a non-linear model is significantly better at the level $p = 0.10$ I'll consider it later in the procedure.

- Step C. Fit the multivariable logistic regression model containing all variables identified for inclusion in step 1. Run a stepwise selection to reduce this model. Make manual corrections based on clinical and statistical considerations to form a preliminary main effects model.

- Step D. Make considerations about the parameters excluded in step C. Is any of the excluded parameters contributing with an important adjustment effect on another covariate. Check if $\Delta\hat{\beta} > 20\%$ for any of the included parameters (comparing the latest model to the model at the beginning of step C). If yes, add the excluded covariates one by one to be able to include those who are contributing with important adjustment.

- Step E. Add the variables excluded in step 1 (the univariable screening) to the model, one by one, to ensure that they still aren't significant in a multivariable environment.

- Step F. If any parameters turned out as non-linear in step B these will be tested for now. The proposed non-linear transformations are added to the model one by one (exchange the linear variant if it's already in the model). Perform a partial likelihood test to find out it the new model is significantly better. At the end of step F the model will be my main effects model.

- Step G. Checking for interactions, following step 6 of purposeful selection.

- Step H. Assessment of fit. I've a rather small sample and will probably have an expected number of CAD less than five in several percentile groups. This makes the Hosmer-Lemeshow test a bit unreliable. I'll run the test anyway but will have this uncertainty in

mind when interpreting the result.
Further I'll make some classification tables, plots and calculate the area under the ROC curve.
I would like to divide my data to be able to perform a external validation, but my sample is to small for this.

Outliers

In the following analysis outliers are removed in every single calculation. For detecting outliers I'll use the interquartile range (IQR) method. $ICQ = Q3 - Q1$. Outliers are defined as values $< Q1 - 3 * IQR$ or $> Q3 + 3 * IQR$. This is necessary since the data contains same extreme values which probably are incorrect.
In the multivariable analysis only subject with complete data sets will be included. Since there's no previous research on a few of the parameters it's hard to replace the outliers in a safe way. Unfortunately I'll lose some subject for the multivariable analysis but I prefer this over including unsure data.

4 Results

Now I'll follow the method described in section 3.8. Previous research and clinical experiences suggests that there is a significant difference between men and women when it comes to ECG patterns for CAD patients. The univariable analysis (step A) is done for both sexes and then I'll do the multivariable analysis for men. Unfortunately we had to few female subject with CAD ($y = 1$) to be able to perform a meaningful multivariable analyse.

Step A: Univariate analysis

In this section I will go through the parameters and analyse them one by one. According to step 1 of the purposeful selection I'll calculate the p-values to decide which parameters to include in the first model, this is done by a simple t-test.
In this subject most of the previous research is univariable comparison between parameters. Thus, in addition to the model building I'm interested in each parameter's discriminant power (to be able to compare to previous research). In this purpose I've performed univariable logistic regression for some parameters. The results are presented as area under the ROC curve (AUC) and sensitivity/specificity for the variable at an arbitrary chosen partition value.
The results of the univariable analysis are presented in table 4 (men) and table 5 (women). Below I'll give some comments to the result of each parameter.

	P value	Partition value	Sensitivity	Specificity	AUC (95 % CI)
ST depression	0.022	119	0.70	0.56	0.64 (±0.10)
ST/HR Index	4.55e-6	2.01	0.66	0.67	0.75 (±0.09)
ST/HR Slope	0.003	5.75	0.68	0.67	0.71 (±0.10)
ST/HR Slope chest	3.99e-06	4.05	0.73	0.66	0.76 (±0.09)
ST/HR Slope extremity	0.20	*	*	*	*
ST/HR rec slope min	0.22	*	*	*	*
ST/HR rec slope max	0.64	*	*	*	*
ST/HR Acc fac max ext	0.15	*	*	*	*
ST/HR Acc fac min ext	0.16	*	*	*	*
ST/HR Acc fac max chest	0.02	*	*	*	*
ST/HR Acc fac min chest	0.83	*	*	*	*
Rec acc fac max ext	0.05	*	*	*	*
Rec acc fac min ext	0.21	*	*	*	*
Rec acc fac max chest	0.10	2.18	0.56	0.52	0.63 (±0.10)
Rec acc fac min chest	0.0002	0.34	0.71	0.63	0.71 (±0.10)
Normalized area ($\alpha = 0.4$)	2.801e-08	-12.91	0.76	0.79	0.80 (±0.08)
Norm. part. area ($\alpha = 0.5$)	5.606e-07	-11.39	0.69	0.69	0.77 (±0.09)
NPA minimum value	0.03	-50	0.63	0.62	0.63 (±0.10)
NPA slope chest	0.91	*	*	*	*
NPA slope extremity	0.58	*	*	*	*
Positive no cross	0.51	*	*	*	*
Negative no cross	0.0003	*	*	*	*
Cross level	0.73	*	*	*	*

Table 4: Men. Results univariable analysis.

	P value	Partition value	Sensitivity	Specificity	AUC (95 % CI)
ST depression	0.56	90	0.50	0.60	0.54 (±0.16)
ST/HR Index	0.35	1.96	0.54	0.69	0.58 (±0.19)
ST/HR Slope	0.04	5.14	0.64	0.65	0.69 (±0.17)
ST/HR Slope chest	0.22	3.78	0.64	0.65	0.64 (±0.21)
ST/HR Slope extremity	0.32	*	*	*	*
ST/HR rec slope min	0.10	*	*	*	*
ST/HR rec slope max	0.78	*	*	*	*
ST/HR Acc fac max ext	0.64	*	*	*	*
ST/HR Acc fac min ext	0.39	*	*	*	*
ST/HR Acc fac max chest	0.98	*	*	*	*
ST/HR Acc fac min chest	0.60	*	*	*	*
Rec acc fac max ext	0.19	*	*	*	*
Rec acc fac min ext	0.17	*	*	*	*
Rec acc fac max chest	0.01	2.15	0.62	0.61	0.71 (±0.16)
Rec acc fac min chest	0.03	-0.6	0.83	0.83	0.81 (±0.16)
Normalized area ($\alpha = 0.2$)	0.006	-13.59	0.64	0.66	0.72 (±0.15)
Norm. part. area ($\alpha = 0.5$)	0.005	-4.38	0.71	0.65	0.73 (±0.14)
NPA minimum value	0.02	-20	0.93	0.41	0.72 (±0.15)
NPA slope chest	0.07	*	*	*	*
NPA slope extremity	0.06	*	*	*	*
Positive no cross	1	*	*	*	*
Negative no cross	0.18	*	*	*	*
Cross level	0.13	*	*	*	*

Table 5: Women. Results univariable analysis.

ST depression

ST depression is just as expected a weak parameter for our study since the patients where chosen because their exercise test were inconclusive. But I've chosen, from clinical considerations, to still include it in my preliminary multivariable analyses even if $p > 0.25$ for women. I think this makes sense because ST depression is an important parameter and it's likely that it still contributes with adjustment and/or interaction effects.

ST/HR Index

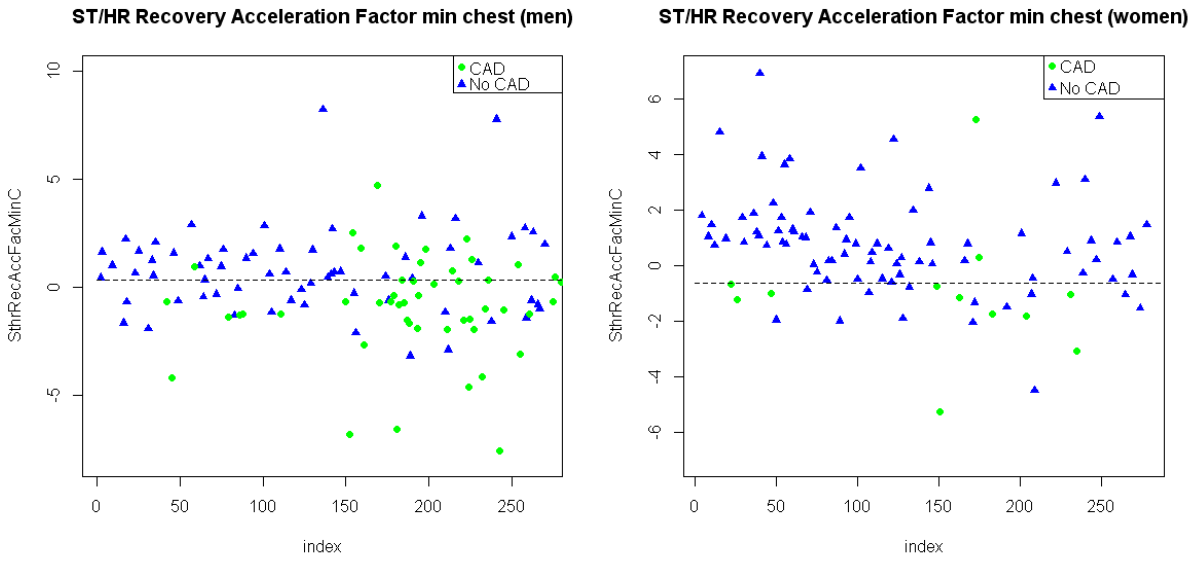
This is a significantly better parameter for men, but for women it's still insignificant. I'll include this parameter in the multivariable for both sexes by the same reason as ST depression.

ST/HR Slope

Normally ST/HR Slope is chosen as the minimum value among all leads (excluding aVL and V1). I made that calculation but also tried to analyse the chest and extremity leads separately. Once again it's a poor result for women. The relationship is barely significant and the confidence interval for AUC is huge. The separate analysis including only chest leads give an interestingly strong significance for men.

ST/HR recovery slope

CAD subjects tend to have a larger ST/HR recovery slope but the significance is very low for both sexes.



(a) men (b) women

Figure 12: Scatter plots for the minimum chest lead value for each subject. Dashed lines are proposed partition values.

ST/HR acceleration factor

Since this is a totally new parameter I've chosen to screen a few variants for significance. They include minimum and maximum values for extremity and chest leads. The same procedure is done for the recovery acceleration factor. The ST/HR acceleration factor gives quite good significance for men but not for women.
But some of the values for ST/HR recovery acceleration factor show surprisingly low p-values. Particularly for women the minimum value among chest leads has a strong discriminating power (especially a good negative predictive value of 97 %). Figure 12 shows a scatter plot for this parameter.

Normalized Area

For the normalized area we got several values for each patient. The minimum value among extremity and chest leads (aVL, V1 and III excluded) are chosen for 8 different cut offs (α-values). To decide which parameter that holds the best discriminating power we calculate the general distances (GD), which is defined as:

$$\text{GD} = \frac{(\text{mean of a}) - (\text{mean of b})}{\text{pooled standard deviation of a and b}} \tag{30}$$

$$\text{pooled SD of a and b} = \sqrt{\frac{Sda^2 * (na-1) + Sdb^2 * (nb-1)}{na+nb-2}} \tag{31}$$

Where a is the population with CAD and b the population without CAD. Sda/b is the standard deviation and na/b is the population size for each group. GD is exemplified in figure 13. The general distance is a measure of the parameter's discriminating power. It's calculated for all combinations mentioned above. I also included an extra parameter which is a combination of the minimum value among extremity leads and the minimum among chest leads. This improved

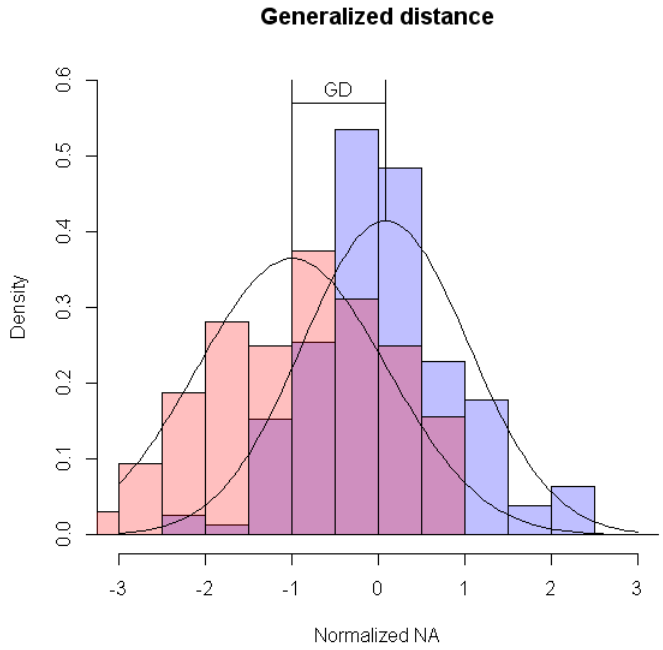


Figure 13: General distance for minimum value among chest leads, $\alpha = 0.3$. Both men and women included. Pink bars represents the population with CAD and blue bars the population without CAD.

the discriminating power in a previous report [Svensbergh et al., 2004]. The result is presented in figure 14. As you can see the GD is greater for men for all values of α and the chest parameter is better than the extremity and the combined parameters. The best value of α is 0.2 for women and 0.4 for men. In the previous report [Svensbergh et al., 2004] they showed similar results and argued for using $\alpha = 0.3$ for both sexes. But since I decided to handle men and women separately I'll continue to work with the best parameter for each sex (α 0.2 for women and 0.4 for men). I suppose that several values of α would result in $p < 0.25$, but to get a reasonable number of parameters in my preliminary multivariable model I've chosen to only use the variants mentioned above.

Normalized partial area

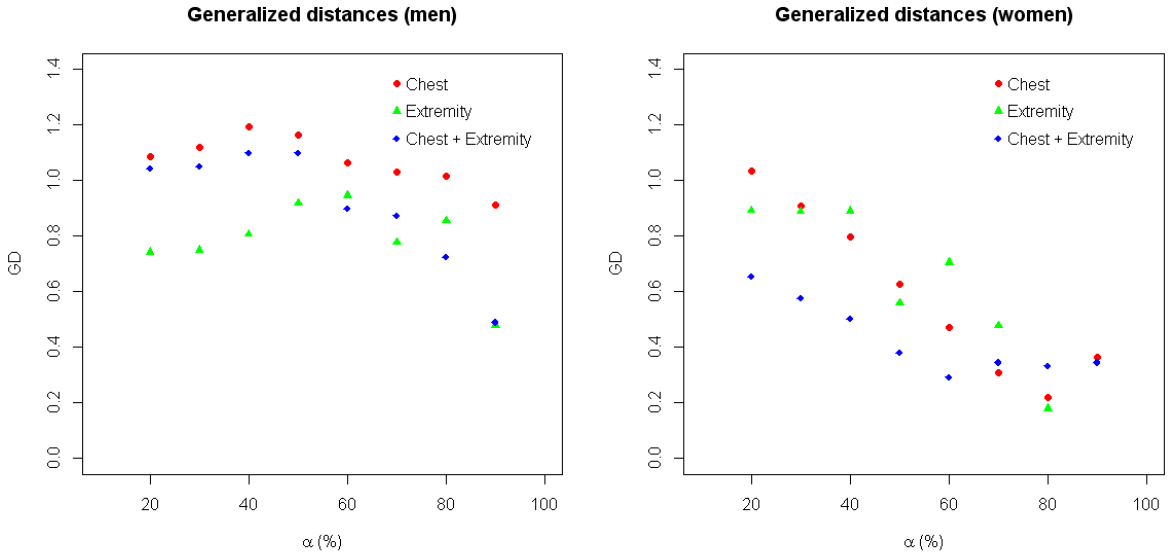
The same analysis as in the previous section (4) for the normalized area gives the generalized areas presented i figure 15. These tell us that this can be an interesting parameter but it's possible that the normalized area is slightly better. Notably is that both men and women has a maximum GD at $\alpha = 0.5$ for chest leads.

NPA minimum value

This is another interesting parameter, especially for women. As you can see in the scatter plot (figure 16) it's possible to obtain a good negative predictive value (0.97).

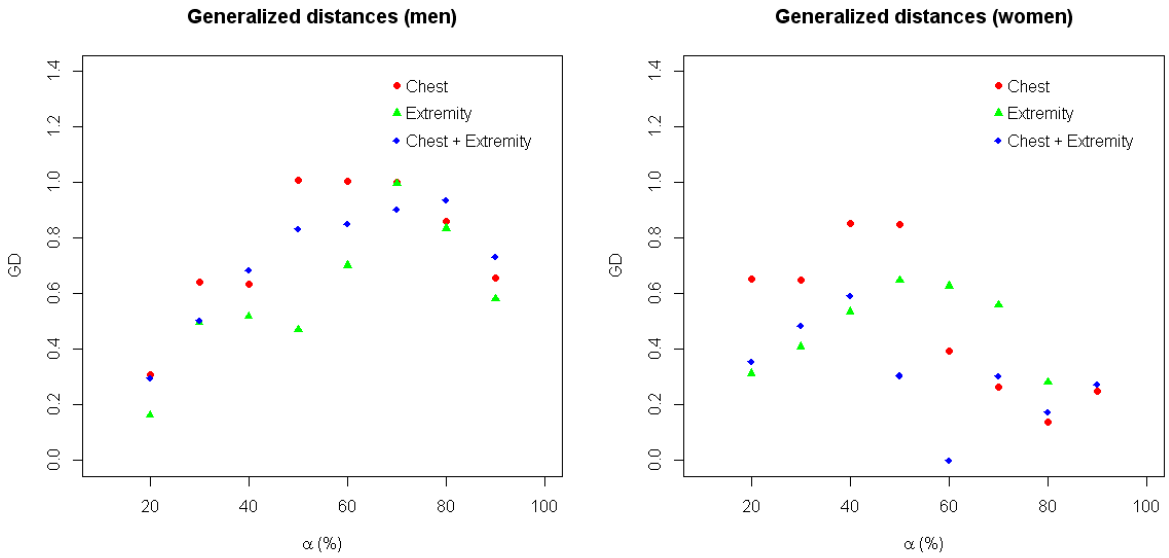
NPA slope

These parameters showed a slightly significant result for women but not for men.



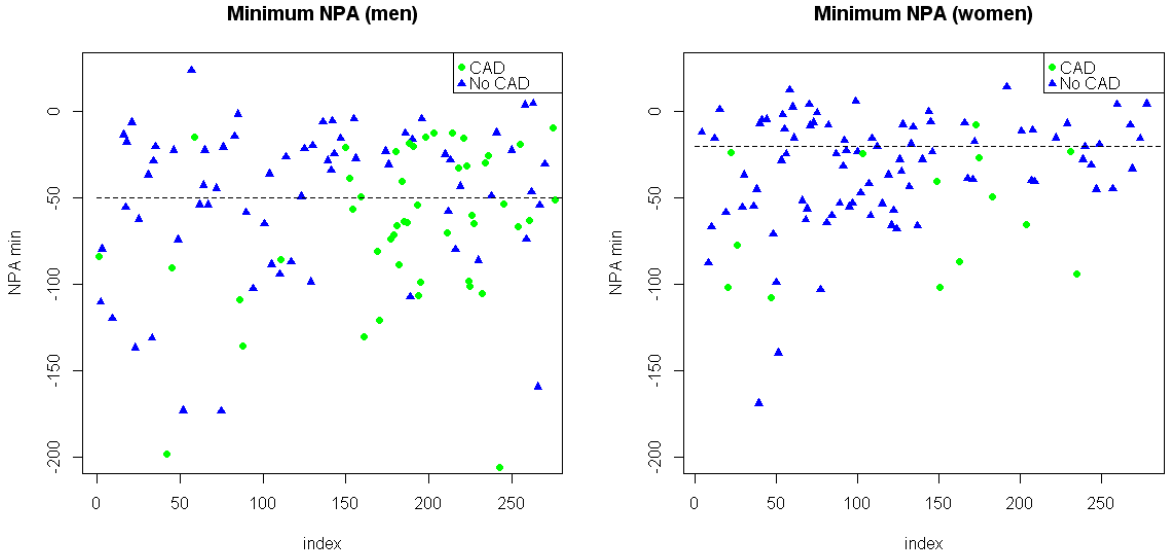
(a) men (b) women

Figure 14: Normalized area. Generalized distances for different α values.



(a) men (b) women

Figure 15: Normalized partial area. Generalized distances for different α values.



(a) men

(b) women

Figure 16: Scatter plots for the minimum NPA for each subject. Dashed lines are proposed partition values.

ST/HR Cross

First I'll analyse those who never cross. They can be positive (recovery curve above exercise curve) or negative (recovery curve below exercise curve). In this case both variables are category, which makes the chi-squared test suitable. The results are presented in contingency tables (table 6 and 7). Notably is that a negatively non crossing curve seems to be a strongly significant factor for men. In the next part I'll analyse those that do cross to see if the level is important. This was weakly significant for women.

Positive	CAD	No CAD
Yes	2	6
No	48	62

Negative	CAD	No CAD
Yes	36	25
No	14	43

Table 6: Men. The first table represents the relationship between having a positive non crossing curve (recovery phase curve above exercise during the whole test) and CAD. In the second table negative non crossing curves are analysed. P value for this correlations are 0.51 and 0.0003.

Positive	CAD	No CAD
Yes	1	10
No	13	79

Negative	CAD	No CAD
Yes	7	25
No	7	64

Table 7: Women. The first table represents the relationship between having a positive non crossing curve (recovery phase curve above exercise during the whole test) and CAD. In the second table negative non crossing curves are analysed. P value for this correlations are 1 and 0.18.

Step B

From this point only male data is included. I evaluated the fractional polynomials for all continuous parameters. The only parameter that turned out as significantly non-linear ($p < 0.1$) was ST/HR slope. The purposed transformation was sthrSlope + sthrSlope^2.

Step C

The result of the model including all parameters selected in step A is presented in figure 17. Note that I also added BMI and the maximum double product as interesting parameters. This is then reduced using stepwise selection and the resulting model is presented in figure 18. I think that it's possible to exclude SthrAccFacMaxC since it's a complex variable with low significance. After the removal the models are compared using a partial likelihood ratio test. The reduced model is presented in figure 19. The difference in deviance is $78.827 - 76.445 = 2.382$. At one degree of freedom the test tells us that the bigger model isn't significantly better than the reduced ($p = 0.12$).

```
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)       -2.263e+00  3.840e+00  -0.589   0.5557
maxDP             -3.031e-05  8.861e-05  -0.342   0.7323
BMI               -4.032e-02  9.959e-02  -0.405   0.6856
stDep              4.660e+00  1.077e+01   0.433   0.6653
sthrInd            1.375e+02  5.780e+02   0.238   0.8119
sthrSlope         -8.482e-01  4.795e-01  -1.769   0.0769 .
sthrSlopeC        -2.710e-01  2.345e-01  -1.155   0.2479
sthrSlopeE         8.763e-01  4.174e-01   2.099   0.0358 *
SthrSlopeMinRec   -3.164e-01  1.925e-01  -1.644   0.1002
SthrAccFacMaxE     1.286e-01  1.528e-01   0.842   0.3999
SthrAccFacMinE    -1.690e-02  1.347e-01  -0.125   0.9002
SthrAccFacMaxC    -2.120e-01  1.408e-01  -1.505   0.1322
SthrAccFacMaxRecE -1.826e-01  1.736e-01  -1.052   0.2929
SthrAccFacMinRecE  2.321e-01  3.519e-01   0.659   0.5096
SthrAccFacMaxRecC  1.562e-01  1.497e-01   1.043   0.2968
SthrAccFacMinRecC -3.971e-01  2.629e-01  -1.510   0.1309
MinNAC40          -2.374e-02  2.653e-02  -0.895   0.3708
MinNPAC50          1.278e-02  2.339e-02   0.547   0.5846
npaMin            -9.503e-03  1.036e-02  -0.917   0.3590
allNegCyes         1.645e+00  7.390e-01   2.225   0.0260 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 121.263  on 90  degrees of freedom
Residual deviance:  68.676  on 71  degrees of freedom
AIC: 108.68
```

Figure 17: First main effects model for men

Step D

$\Delta\hat{\beta} \approx 25\%$ for two of the parameters. I tried to add the excluded parameters one by one to see if any would contribute with an important adjustment effect, but I couldn't find a significantly better fit.

Step E

I included the parameters excluded in step A one by one. Actually I found that the minimum acceleration factor among chest leads was now significant. The inclusion of this parameter made

```
                Estimate Std. Error z value Pr(>|z|)
(Intercept)     -3.452768   0.898864  -3.841 0.000122 ***
sthrSlope       -0.824245   0.317669  -2.595 0.009468 **
sthrSlopeE       0.751289   0.289010   2.600 0.009335 **
SthrSlopeMinRec -0.219875   0.106842  -2.058 0.039596 *
SthrAccFacMaxC  -0.147147   0.095566  -1.540 0.123622
MinNAC40        -0.024705   0.008643  -2.858 0.004259 **
allNegCyes       1.721559   0.634903   2.712 0.006697 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 121.263  on 90  degrees of freedom
Residual deviance:  76.445  on 84  degrees of freedom
AIC: 90.445
```

Figure 18: Result of stepwise selection

```
                Estimate Std. Error z value Pr(>|z|)
(Intercept)     -3.210931   0.851088  -3.773 0.000161 ***
sthrSlope       -0.670108   0.288853  -2.320 0.020347 *
sthrSlopeE       0.637479   0.268176   2.377 0.017450 *
SthrSlopeMinRec -0.224790   0.106533  -2.110 0.034854 *
MinNAC40        -0.023253   0.008558  -2.717 0.006584 **
allNegCyes       1.633958   0.613018   2.665 0.007689 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 121.263  on 90  degrees of freedom
Residual deviance:  78.827  on 85  degrees of freedom
AIC: 90.827
```

Figure 19: Reduced model

the minimum recovery slope insignificant. The preliminary main effects model is presented in figure 20.

```
                Estimate Std. Error z value Pr(>|z|)
(Intercept)     -3.210931   0.851088  -3.773 0.000161 ***
sthrSlope       -0.670108   0.288853  -2.320 0.020347 *
sthrSlopeE       0.637479   0.268176   2.377 0.017450 *
SthrSlopeMinRec -0.224790   0.106533  -2.110 0.034854 *
MinNAC40        -0.023253   0.008558  -2.717 0.006584 **
allNegCyes       1.633958   0.613018   2.665 0.007689 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 121.263  on 90  degrees of freedom
Residual deviance:  78.827  on 85  degrees of freedom
AIC: 90.827
```

Figure 20: Preliminary main effects model

Step F

In step B sthrSlope turned out to be significantly non-linear at level $p < 0.10$ with the purposed transformation sthrSlope+sthrSlope^2. This makes no clinical sense to me but for mathematical purposes I tried to add the second degree term to my preliminary main effects model. This term was barely significant but a partial likelihood test with 2 degrees of freedom wasn't significant. Thus I consider my previous model as my final main effects model.

Step G

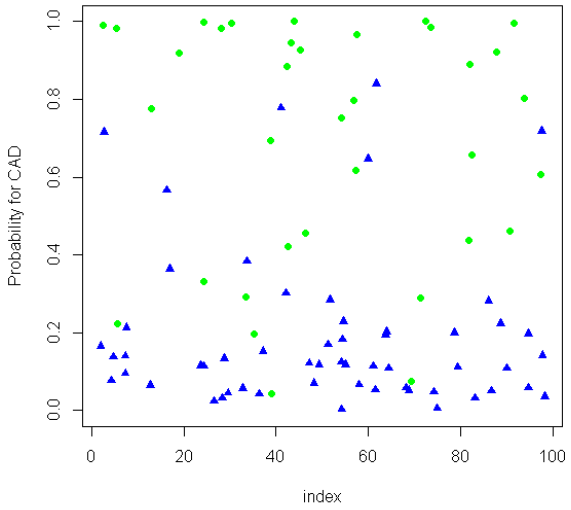
I added the possible interactions one by one and two of them were significant. I added both to my main effects model and performed a partial likelihood ratio test. The new model including the interaction terms was a significantly better fit ($p = 0.006$). The preliminary final model is presented in figure 21.

```
                            Estimate Std. Error z value Pr(>|z|)
(Intercept)                -4.29812    1.23176  -3.489 0.000484 ***
sthrSlope                  -1.32052    0.42316  -3.121 0.001805 **
sthrSlopeE                  0.74313    0.37031   2.007 0.044772 *
SthrAccFacMinC             -0.02566    0.09142  -0.281 0.778963
allNegCyes                  0.24885    0.91151   0.273 0.784844
MinNAC40                   -0.02601    0.01004  -2.590 0.009587 **
sthrSlope:sthrSlopeE       -0.03110    0.01437  -2.164 0.030485 *
SthrAccFacMinC:allNegCyes -0.37028    0.17318  -2.138 0.032510 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 121.263  on 90  degrees of freedom
Residual deviance:  67.909  on 83  degrees of freedom
AIC: 83.909
```

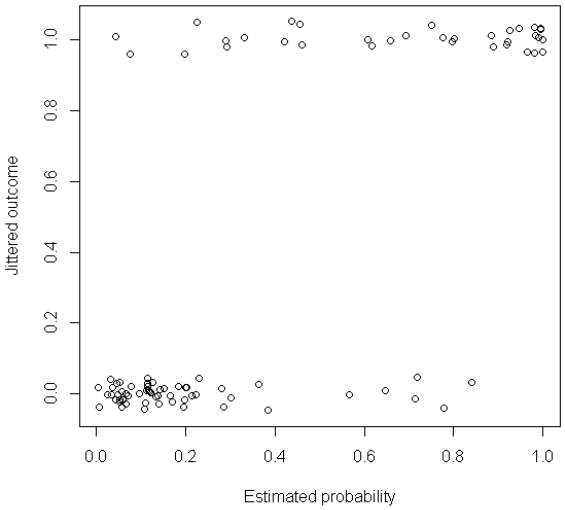
Figure 21: Preliminary final model

Step H

Assessments of the fit. To start with I'll present some plots of the estimated probabilities and observed outcomes, see figure 22 and 23. As can be seen both the scatter plots and the histograms indicates good discrimination.
To compute classification tables I need to decide a reasonable cut-off. Using $p = 0.5$ as cut-off gives a good specificity but a weak sensitivity, see table 8. If we prefer a good sensitivity I'd propose $p = 0.2$ as cut-off. The classification table for this is presented in table 9. A cut-off trying to maximize both sensitivity and specificity would be $p = 0.3$, see table 10.
The ROC curve is presented in figure 24. The area under the curve is : 0.90 ± 0.07 (95 % CI), which can be interpreted as excellent to outstanding discrimination.
The Hosmer-Lemeshow test results in $\hat{C} = 7.9$, which according to the hypothesis that $\hat{C}$ follows the chi-square distribution with 8 degrees of freedom gives $p = 0.44$. This indicates that the fit is good, but as previously discussed (in section 3.8) this isn't totally reliable for my data.

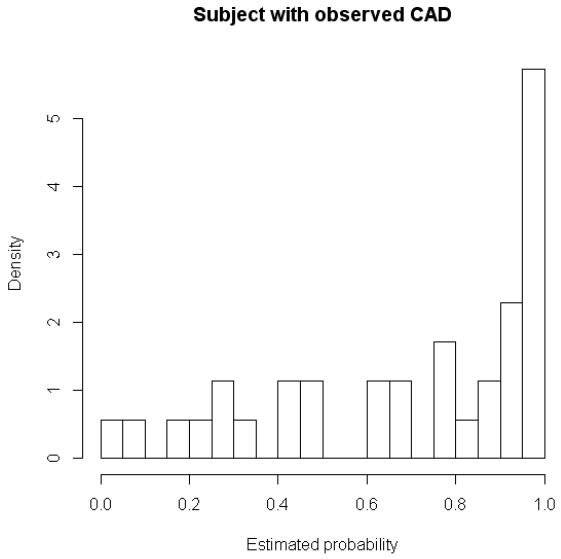


(a) Green dots represented observed CAD ($y = 1$), blue triangles represented observed no CAD ($y = 0$).

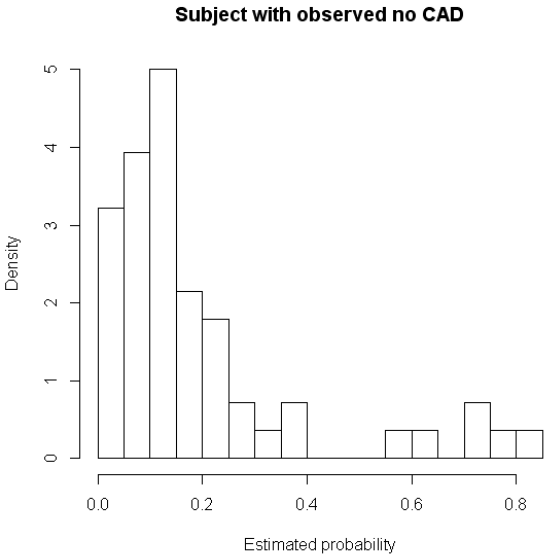


(b) Jittered outcomes = observed outcome $+u_i$, where u_i is generated from the uniform distribution (-0.05,0.05).

Figure 22



(a)



(b)

Figure 23: Histograms for subjects with observed CAD and no CAD

	Observed		
Predicted	y=1	y=0	total
y=1	24	6	30
y=0	11	50	61
Total	35	56	

Table 8: Classification table. Sensitivity = 24/35 = 0.69, specificity = 50/56 = 0.89

	Observed		
Predicted	y=1	y=0	total
y=1	32	16	48
y=0	3	40	43
Total	35	56	

Table 9: Classification table. Sensitivity = 32/35 = 0.91, specificity = 40/56 = 0.71

	Observed		
Predicted	y=1	y=0	total
y=1	29	9	38
y=0	6	47	53
Total	35	56	

Table 10: Classification table. Sensitivity = 29/35 = 0.83, specificity = 47/56 = 0.84

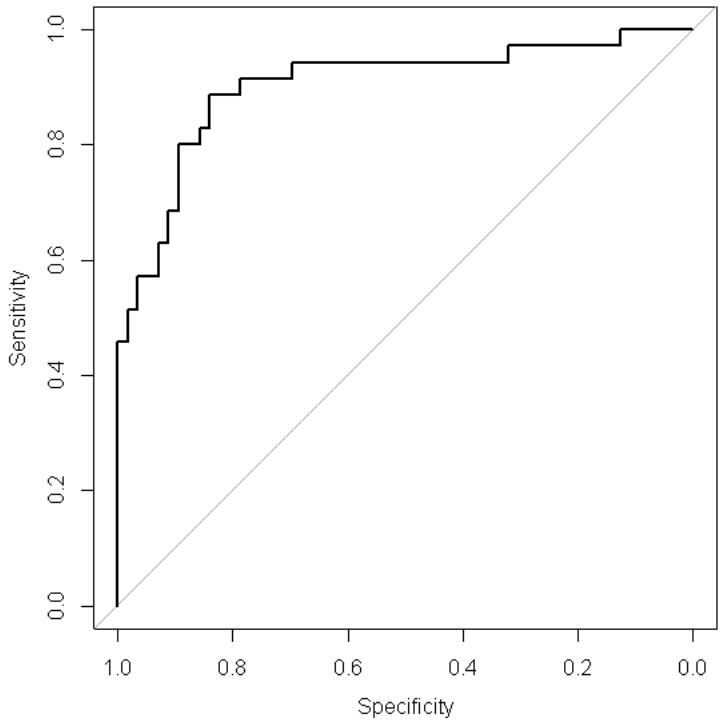


Figure 24: ROC curve

5 Conclusions

First I've to consider the study population. Men and women followed the same inclusion criteria (inconclusive exercise test) but the CAD proportion differs a lot. Of the radiologically confirmed CAD / no CAD 42 % of men and only 14 % of women actually had CAD. The null hypothesis that traditional evaluation criteria are equally good for both sexes can thereby be rejected at $p = 0.000005$. Previously research also proves that commonly used methods discriminates significantly better for men. We need better diagnostic criteria for women! Unfortunately the small number of women with observed CAD in this study made it impossible to make a meaningful multivariable analyse. On the positive side we found some interesting univariable results for women. Especially the recovery acceleration factor showed a surprisingly good discriminating power with a negative predictive value of 97 %. Further the values for the partial normalized areas were partly better for women than men. However I'm cautious in drawing conclusion from these results since the significance is not that strong. But I still consider them as very interesting for further examination.

For men I was able to compute model, using logistic regression, that discriminates way better than the traditional measures for the patients included in this study. One conclusion from the model is that both the work phase and the recovery phase matters.

It would be interesting to examine how well my model discriminates for another population. Remember that this study only includes subjects with inconclusive exercise tests. It wouldn't be fair to compare the discriminating power to previous models where patients who are easy clinically classified are included as well. The discriminating power of my model is thereby probably underestimated.

Finally I wish that these results in some way can contribute to future research and better evaluation criteria. Avoiding extra examinations favours the patients as-well as the community.

A Data handling

This part will describe the data handling process from raw data to data ready for analysing. Data is exported from the clinical exercise test computer program at the hospital as an XML file. This file contains data for the patient (i.e. age, sex, medical treatment, systolic blood pressure and heart rate at rest) and for the exercise test (i.e. heart rate and ST level for all leads recorded several times every minute). It also contains measures of blood pressure and estimated levels of chest pain and effort.

Basic conditions

In the beginning of this project I was given the xml files for 267 patients and a document containing the categorized results from their radiological examinations. I was also provided with 2 applications/scripts (described below) for basic data handling. These where used by Svensbergh et al. [2004].

Mats Johansson's program (sthr.exe) will load the xml file and calculate the normalized area (NA(α)) and the ST/HR slope(α) for all 12 leads. Default settings will calculate these for $\alpha = 0.3$. In order to calculate the values with other α-values you'll have to make changes in a configuration file (sthr.ini) and restart the program. Output from this program is an ascii file containing the calculated values in addition to most of the data from the xml file. It will also produce a pdf file with a graphical presentation of the loops and slopes.

Lars Brudin's visual basic script operates in Microsoft Excel. It extracts data from the ascii file and produces a more user friendly presentation. It also makes calculations which include finding the maximum ST depression for each lead.

Work plan

In order to analyse the parameters described in section 2 I'll need values calculated for several different α-values (0, 0.1,..., 0.9 for slopes and 0.2, 0.3,..,0.9 for NAs), which the tools described above aren't designed for. Basically I would have to run the whole procedure 10 times for each patient. Then manually copy the requested values from the xls files to a master file for each patient. The estimated time for this procedure is at least about an hour per patient and the risk for copying some incorrect values is almost inevitable. I considered this administrative work load as unreasonable for this project. I decided to instead develop some scripts to automatically calculate and extract the same information in much less time. In the following part I'll describe the actions of these scripts.

Development of scripts

I developed one script (runPatient) in the autohotkey language to make the computing automated and thereby faster. My second script (calculatePatient) is a visual basic script for data extraction and calculations. My third script will finally make a summery for all patients in one single file.

Autohotkey script

This script will take the xml file as input and firstly run the sthr application with default settings ($\alpha = 0.3$ for both NAs and slopes). It will produce the ascii and the pdf file as described above. When this is done runPatient will make this iterate another 9 times with different configurations. In total this will produce 10 ascii according to table 11 on page 40. File 1 will then be processed by Brudin's script. This will produce an excel file containing basic data about the patient and

File	α for NA	α for slopes
File 1	0.3	0.3
File 2	0.4	0.4
File 3	0.5	0.5
File 4	0.6	0.6
File 5	0.7	0.7
File 6	0.8	0.8
File 7	0.9	0.9
File 8	0.2	0.2
File 9	-	0.1
File 10	-	0

Table 11: Ascii files produced by STHR application

the exercise test. It will also make calculations (I won't use the most of the calculated values). In the next step runPatient will open the excel file produced by Brudin's script and run my excel macro (script), described below.

Finally runPatient will save the file and arrange the other files in a user friendly folder named as the patient's personal identity number. RunPatient is a semi-automatic script which requires surveillance. It pauses 2 times to let me handle possible notifications which may appear. The whole process runs in less than a minute for each patient.

Visual basic scripts (macros)

In this section I'll describe the actions of my visual basic macros in detail. It operates in Microsoft Excel and requires the excel file produced by Brudin's script and the 10 ascii files produced by Johansson's application.

1. The first action of the macro is to create a new sheet and copy some data from sheet 1 (the output from Brudin's script):
   (a) Personal registration number
   (b) Date of exercise test
   (c) Sex
   (d) Age
   (e) Length
   (f) Weight
2. Calculate file names. The ascii files will automatically be given names by the STHR application according to a specified pattern:

   mmddxx.yaz, where

   mmdd = Date of birth.

   xx = Day of exercise test (01-31)

   y = Month of exercise test (Jan=1, Feb=2,..., Sep=9, Oct=A, Nov=B, Dec=C.

   a = Control character.

   z = Year of exercise test (just last digit. 0-9).

The control character will be calculated in the following way:

X = Digit 10 and 11 in the personal identity number $(= 00 - 99)$
$Z = X \bmod 25$ $(= 0 - 24)$
$Y = Z + 65$ $(= 65 - 89)$

The control character will be the character which ascii code is Y. If a file with that name already exists the new file will be given the next character in the ascii table ($Y = Y + 1$) as control character. If $Y > 89$ it will be recalculated as $Y = Y - 42$.
These names are calculated by my macro to be able to automatically import data from the files.

3. Import data from the ascii files. It will copy all data from the first file and paste it in the bottom part of the excel sheet. This first file will give me:
   (a) Raw data including time, heart rate, effect and ST levels for all leads.
   (b) Calculated normalized areas and slopes for the default configuration ($\alpha = 0.3$).

   Then it will copy data from file 2 to 10 (excluding the raw data which is unaffected by different configurations).
4. Produce tables for the areas and slopes for each configuration and lead.
5. Find the minimum values for the normalized area (greatest negative) among extremity and chest leads (aVL, III, V1 excluded). See figure 25

Minimum NA values		
Alpha	Extremity	Chest
0.2	-64	-173
0.3	-62	-168
0.4	-58	-157
0.5	-50	-139
0.6	-47	-122
0.7	-42	-98
0.8	-38	-63
0.9	-21	-31

Figure 25: Table of NAs

6. Calculate the normalized partial areas and their associated parameters:
   (a) Find the extremity and chest leads with minimum NA($\alpha = 0.3$), aVL, III and V1 excluded.
   (b) De-normalize the areas for these leads.

$$NA(\alpha) = \frac{A(\alpha)}{(1-\alpha) * \Delta HR} \rightarrow A(\alpha) = (1-\alpha) * \Delta HR * NA(\alpha) \tag{32}$$

   (c) Calculate the partial areas:

$$A(\alpha, \beta) = A(\alpha) - A(\beta) \tag{33}$$

(d) Normalize the partial areas:

$$NPA(\alpha, \beta) = \frac{A(\alpha, \beta)}{(\beta - \alpha) * \Delta HR} \tag{34}$$

(e) Plot $\frac{\alpha+\beta}{2}$ on the x-axis and NPA on the y-axis (see figure 9 on page 14). Calculate the slope of the regression line as the "NPA slope".

(f) Calculate the standard deviation among the NPAs.

$$\sigma = Standard\ deviation = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2},\ where\ \mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{35}$$

(g) Find the minimum NPA values and their associated levels (= $\frac{\alpha+\beta}{2}$)

(h) Put it together in a table, *see figure 26 on page 42*

NPA values		
Level	Extremity	Chest
25	-26.52	-70.72
35	-29.24	-79.56
45	-33.32	-83.98
55	-27.88	-70.38
65	-21.42	-65.96
75	-22.1	-57.12
85	-13.94	-32.64
95	0.34	-10.2
Slope	0.36	0.88
SD	10.69	25.19
MinNpa	-33.32	-83.98
MinLevel	45	45

Figure 26: Table of NPAs and associated parameters.

7. Re-calculate slopes. I wasn't really satisfied with the slope values produced by the STHR application. I noticed an irregular time interval between the data points in our raw data (varies from 1 second to 15 seconds). These "extra" data points will make the calculation of areas better, but will give misleading information when working with regression lines. Therefore the script will go through all data points and check the time interval. If the interval is less than 9 seconds one of the data points will be removed. No data point will be removed if it can't be done without leaving an interval less than 15 seconds. I.e. If the interval between X and $X + 1$ is less than 9 seconds: Then it will calculate the new interval if X is removed ($X - 1$ to $X + 1$) and if $X + 1$ is removed (X to $X + 2$). If neither of these is less than 15 seconds no point will be removed. If both are less than 15 seconds the one which will leave the shortest interval exceeding 9 seconds will be removed. This will give an updated table of data points where most of the time intervals will be between 9 and 15 seconds.
In the next step it will calculate some basic heart rate values:

(a) HR_{max}: Maximum heart rate during work phase (defined as the data points where the load is > 0 and rising).

(b) HR_{start}: Heart rate at the beginning of the work phase (defined as the last observation during rest phase).

(c) Heart rate interval during work phase: $\Delta HR = HR_{max} - HR_{start}$.

(d) Corresponding values during recovery phase: $\Delta HR_{rec} = HR_{StartRec} - HR_{EndRec}$.

From these values the cut-off heart rates will be calculated as $HR_{start} + \alpha * \Delta HR$, where α is the proportion of the heart rate interval that will be below the cut-off. The cut-off will be calculated for $\alpha = 0, 0.1, 0.2, ..., 0.9$ during work phase and for $\alpha = 0.3, 0.6, 1$ during recovery phase.
During work phase $\alpha = 0$ will include all values and during recovery phase $\alpha = 1$ will include all values. Now ST/HR slopes and correlation will be calculated for each cut-off and lead. It will also calculate the statistical significance (p-value) of the correlation:

$$r = correlation\ coefficient = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{36}$$

$$t = r * \frac{\sqrt{n-2}}{\sqrt{1-r^2}} \tag{37}$$

The p-value will then be obtained from the student's t distribution with $n - 2$ degrees of freedom.
Only slopes with a statistically significant correlation ($p < 0.05$) will be presented in the final table (see figure 27 on page 43).

Slope Table (calculated)												
Cut off	aVL	I	-aVR	II	aVF	III	V1	V2	V3	V4	V5	V6
W0	Non sign	-0.93396	-1.04296	-0.98605	-0.67616	Non sign	-0.42156	-4.3193	-4.46864	-4.95699	-4.03184	-1.56043
W10	Non sign	Non sign	-0.95935	-1.04675	-0.88415	Non sign	-0.52439	-4.60163	-4.81098	-5.35569	-4.38211	-1.79065
W20	Non sign	Non sign	-1.11056	-1.41658	-1.48322	-1.09329	-0.91066	-5.43435	-5.68115	-6.41905	-5.31589	-2.416091
W30	Non sign	Non sign	-1.11056	-1.41658	-1.48322	-1.09329	-0.91066	-5.43435	-5.68115	-6.41905	-5.31589	-2.416091
W40	Non sign	Non sign	-1.4585	-1.73993	-1.86113	-1.37017	-1.01274	-5.55053	-6.00657	-6.73377	-5.682	-2.59244
W50	Non sign	Non sign	-1.86787	-1.98799	-1.93994	Non sign	-1.12312	-6.21021	-6.53453	-7.3994	-6.03604	-2.930931
W60	Non sign	Non sign	-2.47487	-2.48923	Non sign	Non sign	-1.57013	-7.14217	-6.90282	-7.91766	-5.99809	-3.489708
W70	Non sign	Non sign	-4.21875	Non sign	Non sign	Non sign	Non sign	-8.90625	-8.125	-9.0625	-6.875	-5
W80	Non sign	Non sign	-6.9375	-8.9375	-6.75	Non sign	Non sign	-12.9375	-13.75	-13.375	-11	-9.125
W90	Non sign	Non sign	Non sign	-12.0588	-10.2941	Non sign	Non sign	Non sign	Non sign	Non sign	-12.0588	Non sign
R30	Non sign	Non sign	Non sign	Non sign	Non sign	Non sign	Non sign	Non sign	Non sign	Non sign	Non sign	Non sign
R60	Non sign	Non sign	Non sign	4.714795	4.393939	4.456328	-2.96791	Non sign	3.190731	5.338681	Non sign	-2.308378
R100	-1.45444	Non sign	Non sign	2.135331	2.197808	2.217928	-1.31254	Non sign	Non sign	Non sign	-1.132	-1.335839

Figure 27: Table with calculated slopes. Non significant values are excluded

8. Calculate the ST/HR slope which was introduced by Elamin et al. [1980], described in section 2.3.
   This is found by calculating the slopes and corresponding p-values of the correlation for all possible combinations including the last part of work. Start with the last 3 data points, then the 4 last and continue until all data points during the work phase are included. The steepest negative slope with $p < 0.05$ from each lead will represent the ST/HR slope. The same procedure is done for the recovery phase data (but then maximum slope will be chosen since a decreasing ST levels will give a positive slope during recovery). Finally the minimum value (steepest negative) for each cut-off (aVL and V1 excluded) will be chosen, *see figure 28 on page 44*

9. Calculate the ST/HR acceleration factor.
   Now the macro will use all the slopes calculated in 8.

   (a) Values with a significant correlation will be plotted on the y-axis and the associated number om included data points on the x-axis, *see figure 5 on page 8.*

(b) The slope of the regression line is the acceleration factor for that lead.

(c) The minimum and maximum acceleration factor among extremity leads (aVL excluded) and among chest leads (V1 excluded) will be decided.

(d) The same procedure is done for the recovery phase data.

(e) Values are put together in a table. *See figure 28 on page 44*

	Extremity	Chest
ST/HR Slope	-12.06	-16.5
Max accelatation factor work	15.29	12.49
Min accelatation factor work	4.95	10.15
Avg accelatation factor work	10.02	11.3
ST/HR Recovery Slope	4.71	5.94
Max accelatation factor recovery	-5.51	3.36
Min accelatation factor recovery	-6.12	-12.09
Avg accelatation factor recovery	-5.73	-4.36

Figure 28: Table ST/HR slope and minimum/maximum/average acceleration factor for work and recovery phase.

10. Extract and calculate some single values:

    (a) BMI ($\frac{Weight}{length^2}$)

    (b) Maximum Effort

    (c) Maximum breast pain

    (d) Maximum heart rate.

    (e) Percent of predicted maximum heart rate.

    (f) Systolic blood pressure at rest

    (g) Maximum systolic blood pressure

    (h) Double product (maximum heart rate $*$ maximum systolic blood pressure)

    (i) Maximum effect (load).

    (j) Percent of predicted maximum effect.

11. Finally it will produce a single row summery of all parameters. This can later be copied into one file representing all patients. It will also produce a third sheet with explanations to all parameters.

	A	B	C	D	E	F	G	H	I	J	K	L	
1	PCode	Date	Sex	Age	Length	Weight	ProbNum	ProbNumS	BMI	MaxEffort	MaxChestPain	HRRest	M
2	4	####	M	81	172	75	0.02	-2.0	25.4	13	0	71	

Figure 29: First part of summery for each patient

My third and final script (createSummey) also operates in Microsoft Excel. It will iterate through every patients xls file and fetches the data do make a single file which includes everything I need. In this file each row is represented by one patient. This file will be the input for statistical software where the analysing will take place.

References

R Detrano, E Salcedo, M Passalacqua, and R. Friis. Exercise electrocardiographic variables: a critical appraisal. *J Am Coll Cardiol.*, 8(4):836–847, 1986.

MS Elamin, DA Mary, DR Smith, and RJ. Linden. Prediction of severity of coronary artery disease using slope of submaximal st segment/heart rate relationship. *Cardiovasc Res.*, 14(12): 681–691, 1980.

R. Gianrossi, R. Detrano, D. Mulvihil, K. Lehmann, P. Dubach, A. Colombo, D. McArthur, and V. Froelicher. Exercise-induced st depression in the diagnosis of coronary artery disease. a meta-analysis. *Circulation*, 80(1):87–89, 1989.

D. Hosmer, S. Lemeshow, and R. Sturdivant. *Applied logistic regression.* John Wiley and Sons, Inc., Hoboken, New Jersey, third edition, 2013. ISBN 978-0-470-58247-3.

H. Kronander, W. Fischer-Colbrie, J. Nowak, LA. Brodin, and H. Elmqvist. Exercise electrocardiography for diagnosis of coronary artery disease: impact of sampling rate on the diagnostic performance of st/hr-loop based on data from early recovery phase. *Clin Physiol Funct Imaging*, 28(2):96–100, 2008.

H. Kronander, W. Fischer-Colbrie, J. Nowak, LA. Brodin, and H. Elmqvist. Diagnostic performance and partition values of exercise electrocardiographic variables in the detection of coronary artery disease–improved accuracy by using st/hr hysteresis. *Clin Physiol Funct Imaging*, 30(2):98–106, 2010.

R. Lehtinen, H. Sievnen, J. Vii, V. Turjanmaa, K. Niemel, and J. Malmivuo. Accurate detection of coronary artery disease by integrated analysis of the st-segment depression/heart rate patterns during the exercise and recovery phases of the exercise electrocardiography test. *Am J Cardiol.*, 78(9):1002–1006, 1996.

P. Okin, A. Kligfield, P.Svensbergh, M. Johansson, O. Pahlm, and LH. Brudin. Gender-specific criteria and performance of the exercise electrocardiogram. *JACC*, 25(7):1726–1735, 1995.

P. Royston and D. Altman. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *JRSSA*, 43(3):429–467, 1994.

A. Svensbergh, M. Johansson, O. Pahlm, and LH. Brudin. St-recovery loop of exercise-induced st deviation in the identification of coronary artery disease: which parameters should we measure? *J Electrocardiol.*, 37(4):275–283, 2004.