Statistical Analysis Plan

Study title	A prospective, randomized, controlled, parallel-group, multicentre trial to examine the cost-effectiveness and safety of adding the CADScor®System as a rule-out test in patients referred with symptoms suggestive of stable coronary artery disease.
Short title	The FILTER-SCAD trial
ClinicalTrials-gov ID	NCT04121949
SAP Status and version	Draft, version 1.0, 02 November 2022
Protocol Version	5.0, 15-Dec-2020
Trial Principal Investigators	Eva Prescott, MD, DMSc, FESC Søren Galatius, MD, DMSc, FESC Department of Cardiology Bispebjerg Frederiksberg University Hospital Copenhagen, Denmark
Authors	Louise Hougesen Bjerking, MD Kim Wadt Hansen, MD PhD Department of Cardiology Bispebjerg Frederiksberg University Hospital Copenhagen, Denmark
External statistical consultant	Samuel Emil Schmidt, MSc PhD Department of Health Science and Technology, Biomedical Engineering & Informatics, Aalborg University

This statistical analysis plan will be finalized before database closure.

List of Abbreviations

CAD	Coronary Artery Disease
CCS	Chronic coronary syndrome
CCTA	Coronary computed tomographic angiography
DF-score	Diamond-Forrester Score
ECG	Electrocardiogram
FAS	Full Analysis Set
ICA	Invasive Coronary Angiographies
MPI	Myocardial perfusion imaging
NIT	Non-invasive Test
PTP	Pre-test Probability
SAP	Statistical Analysis Plan
SAQ	Seattle Angina Questionnaire
SAS	Screened Analysis Set
SD	Standard Deviation
QOL	Quality of Life

Table of Contents

1. Background and Study Design

The FILTER-SCAD trial is a randomized, controlled, multicentre, superiority trial with two parallel groups investigating the use of the acoustic measured risk score (CAD-score) for obstructive coronary artery disease (CAD) for ruling out CAD as compared to current standard diagnostic strategy in patients with symptoms suggestive of CAD.

Approximately 2000 patients with symptoms suggestive of CAD referred for outpatient assessment are expected to be enrolled at 5 study sites in Denmark and 1 in Sweden. The subjects will be randomized 1:1 to either 1) standard diagnostic strategy i.e. pre-test probability (PTP) stratification followed by non-invasive test (NIT), if indicated (control group), or (2) PTP plus CAD-score stratification followed by NIT, if indicated (intervention group). Randomization is stratified according to study site and PTP. Given the stratified design, all analyzes are carried out stratified by site and PTP.

2. Analysis Datasets

All patients from whom written informed consent are obtained will be included in the **full analysis set (FAS)**. Subjects will be analysed according to the allocated randomization group in accordance with the intention-to-treat principles.

The **per protocol set (PPS)** will include the subjects from the FAS who did *not* have one of the following major protocol violations:

- inclusion criteria not met
- exclusion criteria met
- no PTP calculation
- no CAD-score measurement (intervention group only).

Patients who did not receive the randomly allocated CAD-score measurement will be included and analysed in the control group (per protocol analysis).

The **screened analysis set (SAS)** will include all screened patients who receive a subject number.

3. Demographics and Baseline Characteristics

Continuous variables will be presented as mean and standard deviation for variables following a normal distribution or as median with interquartile range if showing a skewed distribution. Categorical variables will be presented as frequencies and percentages. Categorical data will be compared using Fisher's exact

test and continuous variables using T-test if normal distributed data or the Mann-Whitney U test. The analyses are performed on **FAS, PPS**, and **SAS**.

4. Primary Objective

The primary objective is to investigate if a combined PTP plus CAD-score guided rule-out strategy is superior in terms of reducing overall number of diagnostic procedures as compared to a standard PTP guided strategy when selecting patients with suspected stable CAD for NIT.

4.1. Primary endpoint

Difference between the two treatment groups in cumulative number of non-invasive and invasive diagnostic procedures, assessed at 1 year after randomization. NITs include exercise ECG, CCTA, Rb-PET CT, MPI, CMRi and stress echocardiography. Invasive procedures include invasive coronary angiographies (ICA) only.

4.2. Statistical Hypothesis, Model and Method of Analysis

Statistical Hypothesis

The statistical null hypothesis is that there is no difference between a PTP plus CAD-score guided rule-out strategy and a PTP guided strategy alone at one year in reducing diagnostic procedures (π) in patients with symptoms suggestive of stable CAD.

$$H_0: X/\pi_{PTP+CAD-score} = X/\pi_{PTP}$$

The alternate hypothesis is that a PTP plus CAD-score guided rule-out strategy is superior to a PTP guided strategy alone at one year in reducing diagnostic procedures (π) in patients with symptoms suggestive of stable CAD.

$$H_A: X/\pi_{PTP+CAD-score} \neq X/\pi_{PTP}$$

Statistical Model and Method of Analysis

The cumulative numbers of diagnostic test will be assessed using Poisson based test and visually by Nelson-Aalen nonparametric estimator.

Total numbers of patients undergoing diagnostic procedures will be presented as total numbers divided into categories of numbers of procedures: Categories: [0, 1, 2, 3+ diagnostic procedures].

The analyses are performed on **FAS** in accordance with the intention-to-treat principles.

A reduction of 15 % or more in the primary endpoint is regarded as clinically significant.

Sample size

A sample size of 521 provides 80 % power with an alpha significance level of 0.05 for testing superiority in terms of numbers of NIT between the control and intervention group.
The power calculation is based on the study design applicable when the study was initiated (Protocol version 3.0) which reflected the European guidelines from 2013 (Montalescot G *et al.* 2013). Protocol version 4.0 and 5.0 reflects the updated European guidelines from 2019 (Knuuti *et al.* 2019), however, the power calculation remains unchanged and a note provides an explanation for each section below.
The main analysis will be based on intention to treat. The study allows for some degree of cross-overs where patients ruled out instead are crossed-over to a NIT. The power calculation simulates a worst-case scenario where the two groups have the lowest expected difference in number of diagnostic procedures; 0% of patients with DF <15% crossed-over to NIT in the control group and 20% of patients with a CAD-score ruling out obstructive CAD in the intervention crossed-over to NIT.
A reduction of 15% or more in the primary endpoint is regarded as clinically significant. Assuming an alpha significance level of 0.05, a statistical power of 80% and an expected number of NIT/ICA of 0.94 per patient in the standard care group, a sample size of 314 patients in each arm is needed to ascertain superiority of the intervention, in this case an absolute reduction in the primary endpoint of 0.17 NIT/ICA per patient, i.e. number of NIT/ICA.
The expected number of NIT/ICA of 0.94 in the control group and 0.77 in the intervention group is based on the following sequence of assumptions:

- No. of first NITs: n=760 in the control group (all patients with DF-score 15-85%) and n= 610 in the intervention group (all patients with a CAD-score not ruling out CAD + 20% of patients with a CAD-score ruling out CAD).
- No. of second NITs: 10% of first NIT; n=76 in the control group and n=61 in the intervention group.
- No. of ICAs: 10% of total population; n=100 in each group.

Note: the patient pathway is expected to change only slightly due to the updated ESC 2019 guidelines (Knuuti *et al.* 2019). Current data indicates that the number of patients in each arm will increase to 521 for the primary endpoint; which stays within the total number of 2000 patients as described below.
Given the abovementioned assumptions, the mean number of investigations (NITs and ICAs combined) per patient will be 0.94 in group 1 and 0.77 in group 2. A sample size of 314 patients in each arm will then provide 80% power for testing superiority in terms of number of diagnostic procedures for the group with CAD-score stratification compared to the group with DF-score only stratification, at an alpha significance level of 0.05.

The study allows for some degree of cross-overs where patients with PTP or CAD-score ruling out CAD instead are crossed-over to a NIT. The power calculation simulates a worst-case scenario where the two

groups (control and intervention group) have the lowest expected difference in number of diagnostic procedures; 0% of patients with low PTP ruling out CAD (Diamond-Forrester(DF) <15 %) crossed-over to NIT in the control and 20% of patients with a CAD-score ruling out CAD in the intervention group crossed-over to NIT.

4.3. Subgroup Analysis

We will investigate the effect of a PTP plus CAD-score guided strategy versus a PTP guided strategy across the following pre-specified subgroups:

- PTP (≤5% vs. >5-15 % vs. >15%)
- PTP (≤5% vs. >5 %)
- PTP (≤5% vs. >5-15 %)
- Age (<65 y vs. ≥65 y).
- Sex (female vs. male).
- Hypertension (yes vs. no).
- Dyslipidaemia (yes vs. no).
- Diabetes mellitus (yes vs. no).
- Current/former smoker (yes vs. no).
- Family history of CAD (yes vs. no).
- Body mass index (<30 kg/m_2 vs. ≥30 kg/m_2).

Poisson-regression will be used in an interaction analysis. The interaction between the randomization variable and the subgroup variable will be investigated. The subgroup analysis will be made using both FAS and PPS.

5. Key Secondary Objective

The key secondary objective is to investigate if a combined PTP plus CAD-score guided rule-out strategy is non-inferior in terms of major adverse cardiac events (MACE) as compared to a standard PTP guided strategy when selecting patients with suspected stable CAD for NIT.

5.1. Key Secondary Endpoint

Difference between the two treatment groups in proportion of MACE, assessed at 1 year after randomization. MACE is a composite of all-cause death, non-fatal myocardial infarction, unstable angina pectoris, heart failure, ischaemic stroke, and major complication of cardiovascular procedures or diagnostic testing within 72 hours.

5.2. Statistical Hypothesis, Model and Method of Analysis

Statistical Hypothesis

The statistical null hypothesis is a PTP plus CAD-score guided rule-out strategy is inferior in terms of safety outcomes MACE (π) in patients with symptoms suggestive of stable CAD at one year when using a pre-specified non-inferiority margin (δ) as compared to a standard PTP guided strategy.

$$H_0 = \pi_{PTP+CAD-score} - \pi_{PTP} \geq \delta$$

The alternate hypothesis is that a PTP plus CAD-score guided rule-out strategy is non-inferior to a standard PTP-guided strategy in terms of MACE in patients with symptoms suggestive of stable CAD at one year when using a pre-specified non-inferiority margin (δ).

$$H_A = \pi_{PTP+CAD-score} - \pi_{PTP} < \delta$$

Statistical Model and Method of Analysis

The absolute risk difference between treatment groups will be compared to a non-inferiority margin (δ) of 1.5% using a two-sided 95% confidence interval applying a continuity-corrected modification of the Wilson's score method. Based on previous studies (Therming *et al* 2018) the obstructive CAD prevalence is expected to be 6.5% and the event rate in each group is expected to be 1.3% at 1-year FU.

Moreover, data will be analyzed using stratified proportional hazard Cox regression and visually by Kaplan Meier curves.

Sample size

A sample size of 1914 provides 90% power with an alpha significance level of 0.05 for testing non-inferiority in terms of MACE between the control and intervention group.
This sample size calculation is based on the hypothesis of non-inferiority in terms of MACE in the group of patients with DF-score and CAD-score stratification prior to NIT compared with the group of patients with standard care strategy of DF-score stratification prior to NIT. It is assumed that in the control group, 50% of the patients with a DF-score <15% are ruled out of obstructive CAD without further testing, and 50% are referred for further testing. Based on previous studies (Therming *et al* 2018) the obstructive CAD prevalence is expected to be 6.5% and the event rate in each group is expected to be 1.3% at 1-year FU. The non-inferiority margin is set to 1.5%.

The assumed annual MACE event rates for different patient categories are:

- Clean arteries (CACS=0 and no stenosis): 0.4%

- Non-obstructive CAD: 1.2%
- Diagnosed CAD: 6.9%
- Unobserved CAD: 10.0%

The diagnostic performance of the CAD-score and the DF-score is assumed to be 88.7% and 96.7%, respectively, for sensitivity and 41.5% and 17.9%, respectively, for specificity (Schmidt *at al* 2019). The weighted average for sensitivity and specificity of the combined NIT battery is assumed to be 85.6% and 80.2%, respectively, assuming a distribution of exercise ECG, stress echo, MPI and CCTA as in Therming *et al* 2018 and diagnostic performance for each test as in the ESC 2013 guidelines (Montalescot *et al* 2013).

With the abovementioned assumptions a sample size of 1914 provides 90% power for testing non-inferiority in terms of MACE between the two testing strategies, at an alpha significance level of 0.05.

Note: the MACE event rate is not expected to change due to the updated ESC 2019 guidelines (Knuuti *et al.* 2019), thus the power calculation related to the MACE endpoint remains unchanged.

6. Other Secondary Objectives

Other secondary objectives are to compare the two treatment arms with respect to:

6.1. Symptoms

6.1.1 Objective

To investigates if chest pain differs in the two treatment arms

6.1.2 Endpoint

Change in chest pain from baseline to 3 months, and 1 year, respectively, after randomization. Chest pain symptoms will be assessed by the Seattle Angina Questionnaire (SAQ).

6.1.3 Statistical Hypothesis, Model and Method of Analysis

The statistical null hypothesis is that there is no difference between the two treatment arms with regards to symptoms. Within treatment group comparisons will be made with Wilcoxon signed rank test. Between treatment group comparisons will be made with Mann-Whitney U test.

6.2. Investigations

6.2.1 objectives

A. To investigate if the number of first NIT differ in the two treatment arms 1 year after randomization.
B. To investigate if the number of downstream tests differ in the two treatment arms 1 year after randomization. Downstream tests defined as all NITs and ICA to detect CAD performed after a) PTP and first standard non-invasive testing and b) PTP, CAD-score and first non-invasive testing,

respectively, in the two groups. The term PTP refers to the DF-score for patients enrolled according to protocol version 3.0 and to the PTP in the ESC 2019 guidelines for patients enrolled according to protocol version 4.0 and 5.0, respectively.

C. To investigate if the number of ICA differ in the two treatment arms 1 year after randomization.
D. To investigate if the number of negative ICA differ in the two treatment arms 1 year after randomization. Classification of ICA results as positive, negative, or inconclusive will be done at each individual site according to local criteria/guidelines.
E. To investigate if the number of repeat referrals differ in the two treatment arms 1 year after randomization.
F. To investigate if the number of NIT performed after revascularization in the two treatment arms 1 year after randomization.

6.2.2 endpoints

A. Difference between the two treatment groups in the number of first NIT, assessed at 1 year after randomization.
B. Difference between the two treatment groups in the number of downstream tests, assessed at 1 year after randomization.
C. Difference between the two treatment groups in the number of invasive coronary angiographies, assessed at 1 year after randomization.
D. Difference between the two treatment groups in the number of negative invasive coronary angiographies, assessed at 1 year after randomization.
E. Difference between the two treatment groups in the number of repeat referrals, assessed at 1 year after randomization.
F. Difference between the two treatment groups in numbers of NIT performed after revascularization, assessed at 1 year after randomization.

6.2.3 Statistical Hypothesis, Model and Method of Analysis

A. The statistical null hypothesis is that there is no difference between the two treatment arms with regards to the number of first NIT differ in the two treatment arms 1 year after randomization.
B. The statistical null hypothesis is that there is no difference between the two treatment arms with regards to the number of downstream tests differ in the two treatment arms 1 year after randomization.

C. The statistical null hypothesis is that there is no difference between the two treatment arms with regards to the number of invasive coronary angiographies differ in the two treatment arms 1 year after randomization.
D. The statistical null hypothesis is that there is no difference between the two treatment arms with regards to the number of negative invasive coronary angiographies differ in the two treatment arms 1 year after randomization.
E. The statistical null hypothesis is that there is no difference between the two treatment arms with regards to the number of repeat referrals differ in the two treatment arms 1 year after randomization.
F. The statistical null hypothesis is that there is no difference between the two treatment arms with regards to the numbers of NIT performed after revascularization 1 year after randomization.

Comparisons of categorical data will be done using Fisher's exact test.

6.3. Diagnosis

6.3.1 Objectives

A. To investigate if the time to rule-out CAD differ in the two treatment arms 1 year after randomization.
B. To investigate if the time to diagnosis of CAD differ in the two treatment arms 1 year after randomization.

6.3.2 Endpoints

A. Difference between the two treatment groups in time to rule-out CAD, assessed at 1 year after randomisation. Time to rule-out is assessed as the period from randomisation until the first test that rules out stable CAD.
B. Difference between the two treatment groups in time to diagnosis of CAD, assessed at 1 year after randomisation. Time to diagnosis is assessed as the period from randomisation until the first test that leads to the final diagnosis.

6.3.3 Statistical Hypothesis, Model and Method of Analysis

A. The statistical null hypothesis is that there is no difference between the two treatment arms with regards to the time to rule-out CAD differ 1 year after randomization.
B. The statistical null hypothesis is that there is no difference between the two treatment arms with regards to the time to diagnosis of CAD differ 1 year after randomization.

Time-to-event data will be analysed using Nelson-Aalen nonparametric estimator and stratified proportional hazards Cox regression.

6.4. Treatment

6.4.1 Objectives

A. To investigate if change in basic lifestyle measures from baseline to 3 months, and 1 year, respectively, after randomization differ in the two treatment arms.
B. To investigate if the proportion of patients initiating and adhering to optimal medical treatment (event prevention, antianginal therapy), differ in the two treatment arms 1 year after randomization.

6.4.2 Endpoints

A. Change in basic lifestyle measures from baseline to 3 months, and 1 year, respectively, after randomization. (Basic lifestyle measures will be assessed by the HeartDiet questionnaire).
B. Difference between the two treatment groups in the proportion of patients initiating and adhering to optimal medical treatment (event prevention, antianginal therapy), assessed at 1 year after randomization.

6.4.3 Statistical Hypothesis, Model and Method of Analysis

A. The statistical null hypothesis is that there is no difference in basic lifestyle measures from baseline to 3 months, and 1 year, respectively, after randomization between the two treatment groups.
B. The statistical null hypothesis is that there is no difference in the proportion of patients initiating and adhering to optimal medical treatment (event prevention, antianginal therapy), assessed at 1 year after randomization between the two treatment groups.

Comparisons between treatment group will be made with Mann-Whitney U test.

6.5. Safety

6.5.1 Objectives

A. To investigate the proportion of each of the individual components of major adverse cardiac events (all-cause death, myocardial infarction, unstable angina pectoris, heart failure, ischaemic stroke, major complication of cardiovascular procedures or diagnostic testing within 72 hours) in the two treatment groups at 1 year after randomization.
B. To investigate the difference between the two treatment groups in the cumulative contrast dose, assessed at 1 year after randomization.
C. To investigate the difference between the two treatment groups in the cumulative radiation dose, assessed at 1 year after randomization.
D. To investigate the difference between the two treatment groups in the proportion of bleedings requiring hospitalization, assessed at 1 year after randomization.
E. To assess the adverse events related to the CADScor®System in the intervention group.

6.5.2 Endpoints

A. Difference between the two treatment groups in the proportion of each of the individual components of major adverse cardiac events (all-cause death, myocardial infarction, unstable angina pectoris, heart failure, ischaemic stroke, major complication of cardiovascular procedures or diagnostic testing within 72 hours), assessed at 1 year after randomization.
B. Difference between the two treatment groups in the cumulative contrast dose, assessed at 1 year after randomization.
C. Difference between the two treatment groups in the cumulative radiation dose, assessed at 1 year after randomization.
D. Difference between the two treatment groups in the proportion of bleedings requiring hospitalization, assessed at 1 year after randomization.
E. Adverse events related to the CADScor®System in the intervention group.

6.5.3 Statistical Hypothesis, Model and Method of Analysis

A. The statistical null hypothesis is that there is no difference in the proportion of each of the individual components of major adverse cardiac events (all-cause death, myocardial infarction, unstable angina pectoris, heart failure, ischaemic stroke, major complication of cardiovascular procedures or diagnostic testing within 72 hours) between the two treatment groups, assessed at 1 year after randomization.
B. The statistical null hypothesis is that there is no difference in the cumulative contrast dose, assessed at 1 year after randomization between the two treatment groups. Comparisons between treatment group will be made with Mann-Whitney U test.
C. The statistical null hypothesis is that there is no difference in the cumulative radiation dose, assessed at 1 year after randomization between the two treatment groups.
D. The statistical null hypothesis is that there is no difference in the proportion of bleedings requiring hospitalization, assessed at 1 year after randomization between the two treatment groups.
E. The statistical null hypothesis is that there is no difference in adverse events related to the CADScor®System between the two treatment arms.

Comparisons between treatment group will be made with Mann-Whitney U test. Comparisons of categorical data will be done using Fisher's exact test.

6.6. Quality of Life

6.6.1 Objective

To investigate if change in quality of life (QOL) from baseline to 3 months, and 1 year, respectively, after randomization differ in the two treatment groups.

6.6.2 Endpoint

Change in quality of life (QOL) from baseline to 3 months, and 1 year, respectively, after randomization (Quality of life will be assessed by the EuroQoL-5D).

6.6.3 Statistical Hypothesis, Model and Method of Analysis

The statistical null hypothesis is that there is no change in quality of life (QOL) from baseline to 3 months, and 1 year, respectively, after randomization (Quality of life will be assessed by the EuroQoL-5D).

Within and between treatment group comparisons will be made with Wilcoxon signed rank test and Mann-Whitney U test, respectively.

7. Sensitivity Analysis

A sensitivity analysis on all endpoints will be performed on patients enrolled according to protocol version 4.0 and 5.0 to investigate any potential differences in study results related to the change in study design due to the update of the European guidelines. The analyses will be performed using both PPS and SAS.

Heterogeneity in the treatment effect between sites and PTP groups will be investigated. An interaction analysis between the randomization variable and site will be made.

8. Missing Data

Data will be analysed for potential extreme outliers. If present, each outlier will be investigated. However, outliers will be included in the analysis, except if a clinical cause can be excluded.

9. References

Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. Eur Heart J. 2019 Aug 31

Task Force Members, Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J. 2013 Oct;34(38):2949–3003

Therming C, Galatius S, Heitmann M, Højberg S, Sørum C, Bech J, et al. Low diagnostic yield of non-invasive testing in patients with suspected coronary artery disease: results from a large unselected hospital-based sample. Eur Heart J Qual Care Clin Outcomes. 2018 01;4(4):301–8