A Comparison of Functional and Anatomical Investigations with Angina Symptoms: Results from a Randomized Controlled Trial

P. McKavanagh1,2*, G.M. Walls1, F. Denny1, L. Lusk1, P.A. Ball1, T.R. Trinick1, E. Duly1, R.M. Verghis3, C. McCune1, M.T. Harbinson2, P.M. Donnelly1

1Ulster Hospital, South Eastern Health and Social Care Trust, Upper Newtownards Road, Dundonald, Belfast BT16 1RH, UK
2Centre for Experimental Medicine, Queen’s University Belfast, Lisburn Road, Belfast BT9 7BL, UK
3The Northern Ireland Clinical Trials Unit, Education and Research Centre, Royal Hospitals, Belfast BT12 6BA, UK

*Corresponding author: McKavanagh, P. M.B., M Med Sc, Ulster Hospital, South Eastern Health and Social Care Trust, Upper Newtownards Road, Dundonald, Belfast BT16 1RH, UK, E-mail: mckavanagh@doctors.org.uk

Abstract
The presence of angina is a potential source of reduced quality of life and may be a predictor of adverse cardiovascular outcomes. Recent guidelines suggest that management of angina should revolve around the use of the Diamond Forrester (DF) score to determine which tests should be used for an individual patient. The choice of test can be broadly categorized into anatomical or functional. Little is known about how either type of test correlates with the DF score. This study was part of a larger randomized Controlled Trial to compare the use of cardiac Computerized Tomography (CT) and exercise Electrocardiogram Stress Test (EST) in patients with stable angina. This analysis compared the results from EST and CT against the DF scores. Overall all results showed that there is a poor correlation between DF and both tests. R-squared value from the linear model showed correlation between EST and DF scores to be 0.315, with the correlation between CT and DF scores 0.415. This calls into question the useful of symptoms to drive investigations.

Background
Despite declining incidences of Myocardial Infarction (MI), the prevalence of angina remains high[1-3], with large epidemiological studies showing that it is not a benign condition[3,4]. It is due to a temporary imbalance between myocardial perfusion and oxygen demand[5], although unlike MI, it has no universal definition. The diagnostic investigations for patients with suspected angina can be classified as either functional or anatomical. Functional tests assess the haemodynamic consequences of obstructive Coronary Artery Disease (CAD), whereas anatomical tests directly visualise the coronary arteries for stenoses[6]. The choice of test often depends on local expertise, availability and cost. Recent guidelines have suggested that the investigation used should depend on likelihood of disease, as defined by symptoms[7-9].

The Diamond Forrester (DF) score is a pre-test probability model that assesses the likelihood of having significant CAD in symptomatic patients[10]. It uses the patient’s symptoms, sex, age and risk factors to give a score that predicts disease likelihood. Recent work has suggested that it can over-estimate the presence of CAD[11-13], but it remains the recommended scoring system in national guidelines[7-9]. Little is known however about its correlation with different types of chest pain investigations.

This study is part of the ongoing analysis of the Cardiac CT for the Assessment of Pain and Plaque (CAPP) randomised controlled trial, which is designed to compare the economic and clinical outcomes of using cardiac Computerized Tomography (CT) or exercise stress electrocardiogram test (EST) for patients with suspected stable CAD. [ISRCTN52480460]. This analysis will explore the relationship of the DF with clinical investigations.

Received Date: April 13, 2016
Accepted Date: May 20, 2016
Published Date: May 26, 2016


DOI: 10.15436/2378-6914.16.025

Copyrights: © 2016 McKavanagh, P. This is an Open access article distributed under the terms of Creative Commons Attribution 4.0 International License.
Method

The method of this trial has been described elsewhere\(^{(14)}\), with a synopsis provided below. The study protocol was approved by the Office for Research Ethics Committee Northern Ireland (ORECNI) and the South Eastern Health and Social Care Trust (SE-HSCT) Research and Development Committee. The study was also supported by the Northern Ireland Medical Physics Agency, the Northern Ireland Cardiovascular Research Network and the Northern Ireland Clinical Trials Unit. The trial process relevant to this sub-study is described below.

Patients

The CAPP trial prospectively evenly randomized 500 patients who were referred to clinic with symptoms of stable chest pain to EST or cardiac CT. Stable chest pain was defined as troponin negative without symptoms suggestive of unstable angina. Exclusion criteria have been cited elsewhere.

All patients had their age, gender, risk factors and chest pain characteristics documented. Pain type was classified as typical, atypical or non-anginal using criteria defined\(^{(15)}\). The probability of significant CAD was calculated using the DF score with the method described elsewhere. In keeping with national guidelines\(^{(7)}\), risk factors associated with high cardiovascular risk included diabetes mellitus, and/or smoking and/or total serum cholesterol level greater than 6.47 mmol/l.

EST

EST was performed using a standard Bruce protocol treadmill with continuous 12 lead ECG monitoring and registration at 1 min intervals. Manual blood pressure monitoring was performed every 2 min.

Criteria for discontinuation the test were ST changes of depression > 0.3 mV or elevation > 0.1 mV; blood pressure changes of systolic > 230 mmHg or diastolic > 130 mmHg or a > 10 mmHg systolic blood pressure drop; arrhythmias such as sustained ventricular tachycardia, increasing frequency of polymorphic ventricular complexes, or altered atrioventricular or intraventricular conduction; and patient symptoms of exhaustion, extreme dyspnoea, or angina. EST results were classified as negative, positive, or inconclusive as defined by previous criteria\(^{(16)}\).

CT image acquisition

Patients in the CT arm underwent both a Calcium Score (CS) and subsequent Computerized Tomography Coronary Angiogram (CTCA), regardless of initial CS. These were performed on a 64-detector platform (Philips Brilliance 64 Cleveland, Ohio, USA). As per departmental policy, both oral and intravenous beta blockers were used for heart rate control pre scan. A heart rate of 65 bpm or below was considered optimal for imaging, although CTCA would have been performed below the level of 70 bpm. Patients with heart rates above 70 bpm despite intravenous beta blockers were rescheduled for a later date with larger doses of pre-procedure oral beta blockade. A non-contrast enhanced, axial prospective triggered CS was performed at 3-mm slice thickness with milli amperes optimized per patient and a standard 120 kV. CTCA contrast medium enhancement was achieved with a biphasic injection protocol. An aliquot of 80 – 100 mL of iodinated contrast material (Ioversal, Optiray 350 mgI/mL, Coviden, Hampshire, UK) was administered through an 18-gauge intravenous antecubital fossa cannula at a flow rate of 6 mL/s followed by a saline chaser of 50 mL at 3.5 mL/s. An automated bolus tracking technique prospectively monitored contrast arrival in the descending aorta. A threshold of 110 Hounsfield Units (HU) was used to initiate the scan. Image acquisition was in the cranio-caudal direction, and the scan field of view was refined using precise anatomical landmarks cross referenced from the coronary CS. Images from CTCA were reconstructed at 0.8-mm slice thickness with the standard reconstruction filter. Scan parameters (kV, mAs) were optimized by the imaging clinician and were patient specific. The choice of a retrospective gating, or an axial prospective triggering algorithm, was at the discretion of the clinician. For all retrospective examinations, ECG dose modulation was applied (DoseRight Cardiac, Philips Healthcare, Cleveland, Ohio, USA).

Analysis of CT images

CT was assessed using a semi-automated analysis package (Heartbeat CS, Philips Cleveland). Areas of calcification within the coronary arteries were identified from a transaxial image stack, and a validated algorithm was used to calculate the total Agatston score\(^{(17)}\). CTCAs were evaluated for the presence of coronary artery stenosis by an experienced cardiologist and radiologist using the anatomical American Heart Association 15-point score\(^{(18)}\). A coronary diameter stenosis > 50% in a major epicardial artery > 2 mm was considered significant, which is in keeping with previous work\(^{(19-22)}\). Such results were deemed positive and could involve either calcified or non-calcified disease. Diameter stenosis < 50% but > 0%, was considered non-significant and together with the presence of no disease (stenosis = 0%) was deemed negative. If any large vessels ( > 2 mm) in diameter could not be evaluated due to excessive calcifications, motion artefacts, or inadequate contrast filling, the study was deemed inconclusive.

Study endpoints

The primary outcome measure of the trial was the difference in the percentage change in scores within the Seattle Angina Questionnaire (SAQ) domains from baseline to 3 months between the two cohorts. This analysis involves the comparison of investigations against DF scores in terms of correlation and agreement with the final diagnosis.

www.ommegaonline.org
Statistical Analysis

There were several aims of this study. Firstly, the authors wished to find the correlation of DF scores with CT and EST in classifying significant and non-significant CAD in patients with angina. Secondly to determine the agreement of CT and EST findings in terms of a high/low DF score. Thirdly to determine the accuracy of DF classification with CT and EST in classifying significant and non-significant CAD according to the diagnosis established at the conclusion of the investigative pathway.

Demographics

Descriptive statistics were used to summarize demographic variables by group. In general, mean, standard deviation, range (minimum to maximum) were employed for continuous variables, whilst frequency and percentage were used for categorical variables. To assess the differences between groups, a Fisher’s Exact test was performed.

Sample Size

The sample size was calculated in relation to the original trial primary outcome, i.e., the change of the SAQ score at 3 months from baseline. Details of this have been previously described[14].

DF Scores

To determine if there was a relationship between DF scores and the EST and CT classification, box plots illustrating the distribution of the scores were initially created. A linear model with Analysis Of Variance (ANOVA) was performed to further examine differences of DF scores among the positive, negative and inconclusive groups for each of EST and CT. A two sample t-test was used to determine if there was a difference in DF scores between EST and CT tests.

DF Classification

DF scores were classified in two groups to help identify those who were categorized as high risk and low risk. Those scores above 50 were categorized as ‘high’, whilst those equal to or below 50 were categorized as ‘low’. To determine if there were differences between the classified DF scores and EST and CT classification, a Chi-squared test with Yate’s continuity correction was employed. To examine if the classification threshold of 50 was appropriate, true positive and negative classifications with 95% confidence intervals were derived.

Results

Demographics

Of the 500 patients with suspected CAD who were randomized, 9 did not receive the allocated test and 3 others withdrew after 3 months. This meant that 245 patients in the EST cohort and 243 in the CT cohort (see Figure 1). Baseline demographics are seen in Table 1.

Table 1: Baseline Demographics of each cohort

<table>
<thead>
<tr>
<th>Category</th>
<th>EST arm Mean ± SD (n = 245)</th>
<th>The CT arm Mean ± SD (n = 243)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age</td>
<td>58.9 ± 10.2</td>
<td>57.8 ± 10.0</td>
<td>0.22</td>
</tr>
<tr>
<td>Mean BMI (kg/m²)</td>
<td>28.0 ± 3.6</td>
<td>27.8 ± 3.6</td>
<td>0.53</td>
</tr>
<tr>
<td>Sex</td>
<td>131 males</td>
<td>138 male</td>
<td>0.44</td>
</tr>
<tr>
<td>Smoker</td>
<td>47</td>
<td>46</td>
<td>0.93</td>
</tr>
<tr>
<td>Diabetic</td>
<td>12</td>
<td>14</td>
<td>0.68</td>
</tr>
<tr>
<td>Known Hypertensive</td>
<td>73</td>
<td>77</td>
<td>0.53</td>
</tr>
<tr>
<td>Mean Cholesterol (mmol/L)</td>
<td>5.4 ± 1.1</td>
<td>5.3 ± 1.1</td>
<td>0.49</td>
</tr>
<tr>
<td>Mean Diamond Forrester scores (%)</td>
<td>44.9 ± 30.2</td>
<td>47.8 ± 31.7</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Character of Pain

<table>
<thead>
<tr>
<th>Character of Pain</th>
<th>EST arm</th>
<th>The CT arm</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; Non-angina</td>
<td>156 (63.7%)</td>
<td>143 (58.8%)</td>
<td>0.28</td>
</tr>
<tr>
<td>&gt; Atypical</td>
<td>20 (8.2%)</td>
<td>16 (6.6%)</td>
<td>0.47</td>
</tr>
<tr>
<td>&gt; Typical</td>
<td>68 (27.8%)</td>
<td>84 (34.6%)</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Result of index investigation

<table>
<thead>
<tr>
<th>Result of index investigation</th>
<th>EST arm</th>
<th>The CT arm</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; Positive</td>
<td>47</td>
<td>73</td>
<td>0.0054</td>
</tr>
<tr>
<td>&gt; Negative</td>
<td>132</td>
<td>164</td>
<td>0.0021</td>
</tr>
<tr>
<td>&gt; Inconclusive</td>
<td>66</td>
<td>6</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

EST = Exercise Stress Electrocardiogram Test
CT = Computerised Tomography
The mean DF score was 46.3%. There was no statistically significant difference between the cohorts, as seen in Table 1. Figure 2 and Table 2 shows the distribution of DF scores against EST and CT results. There was highly statistically significant difference noted in DF scores between the EST categories of positive, negative and inconclusive ($p \leq 0.05$). Likewise highly statistically significant differences were also noted in DF scores between the CT categories of positive, negative and inconclusive ($p \leq 0.05$). After removing the inconclusive categories, the difference in DF scores between the EST categories remained highly statistically significant. Positive values were on average 21.23 higher than negative values. This pattern was also seen after removing the inconclusive categories from the CT results, with the difference in DF scores between the CT categories remaining highly statistically significant. Positive values were on average 28.33 higher than negative values.

**Table 2:** DF scores as per test result

<table>
<thead>
<tr>
<th></th>
<th>Non-significant result from initial test</th>
<th>Significant result from initial test</th>
<th>Inconclusive result from initial test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative EST (n = 132)</td>
<td>No CAD/non-obstructive on CT (n = 164)</td>
<td>Positive EST (n = 47)</td>
<td>Significant CAD on CT (n = 73)</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>56.8</td>
<td>63.7</td>
</tr>
<tr>
<td></td>
<td>Number of males</td>
<td>54.8</td>
<td>64.6</td>
</tr>
<tr>
<td></td>
<td>Average Diamond Forrester, %</td>
<td>76 (57.6%)</td>
<td>26 (55.3%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>84 (51.2%)</td>
<td>52 (71.2%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>39.3</td>
<td>58.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>37.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>32.0</td>
</tr>
</tbody>
</table>

**Results of EST and CT**

In the EST arm the mean exercise duration was 7 minutes 35 seconds (ranging from 1 min 33 sec to 16 min 20 sec). Overall, 47 were positive, 132 negative and 66 inconclusive (see Table 1). In the CT arm, the mean CS was 172.2 ± 490.1 (ranging from 0 to 3901) with a median of 0. One hundred and twenty-six patients had a CS of zero; 92 had a score of 1 – 400, and 25 had a CS > 400. CTCA demonstrated that 73 had significant disease, 58 had non-obstructive disease, 106 had no disease, and 6 were inconclusive (see Table 1).
With inconclusive categories included, there was no statistically significant difference found between the test (EST or CT) and DF scores (p-value = 0.300). This result remained true with inconclusive values removed (p-value = 0.074). A Chi-square test suggested that there was no statistically significant difference between positive EST/CT and high/low DF scores (p-value = 0.524). Similarly, there was no statistically significant difference between negative EST/CT and high/low DF scores (p-value = 0.662).

Correlation between investigation results and DF

To derive a correlation against the EST or CT classification, the R-squared value from the linear model was examined. The overall correlation between EST and DF scores was estimated to be 0.315, with the correlation between CT and DF scores was estimated to be 0.415.

Discussion

The current study involves a prospective randomized cohort of patients presenting with chest pain of suspected CAD origin. It compares their DF scores and the results of their allocated investigation, EST or CT, along with the final diagnosis following further investigation where required. The main finding of the study was that the overall correlation between DF and tests results was poor. Previous studies have suggested that there is a weak correlation between myocardial ischemia and the burden of angina. This paper has added that direct visualization of CAD stenoses with an anatomical test is no more accurate than functional testing in predicting high likelihood of disease as suggested by DF score. To our knowledge this has never been described before.

Approximately 25% of people experience chest pain in their life time, with it accountable for approximately 1% of visits to the general practitioner, 5% of Accident and Emergency Department (A & E) attendances and up to 25% of unplanned hospital admissions. However it is estimated that only 11 to 39% of patients presenting with chest pain are eventually found to have CAD. Patients with chest pain and no evidence of significant CAD can continue to have physical limitations and a poor quality of life. The pathophysiology of such pain is unknown, although prior studies have suggested possible mechanisms such as endothelial and microvascular coronary dysfunction as well as psychological factors. Traditionally patients with symptomatic non obstructive disease are a difficult cohort to manage as symptoms are often indistinguishable from those with significant CAD.

The overall poor correlation between DF score and both the EST and CT results means that this study adds to the body of evidence that suggests that symptom driven stratification does not permit accurate discrimination in the detection of significant CAD. Character of pain in a patient’s history alone is not a good discriminator of risk of significant CAD, with previous work showing patients with non-anginal pain having a yearly risk of cardiac death of 0.4% compared to 0.9% of those presenting with typical pain. Furthermore myocardial ischemia does not always equate to angina with it estimated that 25% to 45% of patients with CAD experience episodes of asymptomatic ischemia. Asymptomatic ischemia is associated with poor prognosis, whereas the converse of angina in the absence of ischemia is not. It has also been suggested that relief of ischemia rather than relief of angina should be the aim of therapy. In an ideal world both anatomy and functional significance of disease would be determined before deciding on a management strategy. However in everyday clinical practice both anatomical and functional tests are used in chest pain assessment, with the choice depending on local policy, availability, and clinician expertise.

Cardiac CT work has also highlighted the weakness with risk stratification by history as the presence of non-significant disease on CTCA is associated with adverse CAD events. Asymptomatic individuals with disease on CT also have a less favorable prognosis. Cardiac CT has the ability to exclusion CAD in patients with a low-to-intermediate cardiovascular risk profile, with scanners able to produce good quality scans without compromising image quality. Patients with chest pain yet normal coronary arteries have a good prognosis. Recent work has suggested that the use of cardiac CT can prevent mislabeling of patients with the diagnosis of angina. This present study on stable chest pain is in keeping with previous published work that CT helps stratify patients better than the history scoring systems alone.

It is somewhat predictable that in those patients with low DF scores, negative results were more accurately identified by the EST than the CT. Due to the anatomical nature of CT it is possible to have the presence of obstructive disease detected before the onset of symptoms. Anatomical disease can be present without functional significance or symptoms, whereas in contrast the EST attempts to detect ischemia, with pain being part of the ischemic cascade.

There are a number of limitations in this study. Firstly it was a single centre study. Secondly due to exclusion criteria of BMI and renal function there were low numbers of diabetics in this study, whose symptoms may be less ‘typical’ than standard patients. Thirdly EST is not as accurate as stress imaging modalities, which may have produced different results. However the aim of the CAPP study was to contrast CT with the present standard of care, which in the UK is the EST. Fourthly the study was powered to evaluate the difference in SAQ domains over time, not for this analysis of contrasting SAQ results against test results. Furthermore, in this study cardiac CT was used for all patients regardless of pre-test probability. Guidelines suggest that it should be only used for low and intermediate risk patients. In addition, the final diagnosis of significant or non-significant CAD was influenced by the discretion of clinicians ordering additional investigations. This was not a standardized process and may introduce bias.

Conclusion

In conclusion, DF score is a poor tool for clinicians making the decision to proceed to either CTCA or EST, as there it has poor correlation with both of these tests.
Investigations with Angina Symptoms

References

Investigations with Angina Symptoms


