Angina pectoris: Has the cornerstone of symptomatic coronary artery disease been moved?

Renee C Hessian MD, Rob SB Beanlands MD

But there is a disorder of the breast marked with strong and peculiar symptoms, considerable for the kind of danger belonging to it, and not extremely rare, which deserves to be mentioned more at length. The seat of it and sense of strangling, and anxiety with which it is attended, may make it not improperly be called angina pectoris. They who are afflicted with it are seized while they are walking (more especially if it be up hill, and soon after eating) with a painful and most disagreeable sensation in the breast, which seems as if it would extinguish life, if it were to increase or continue; but the moment they stand still this uneasiness vanishes.

William Heberden 1768 (1-3)

Classical history of angina pectoris has long been considered to be a symptom of myocardial ischemia and, as such, a predictor of the presence of significant coronary artery disease. The work by Højlund-Carlsen et al (pages 641-647 of the current issue of the *Journal*) challenges this cornerstone of care.

This work provides an interesting window into the correlation of clinical history with the presence of angiographic disease. This study was a substudy of a larger prospective study that looked at myocardial perfusion scintigraphy (MPS) as a gatekeeper for coronary angiography (CA) in patients referred for angiography with angina pectoris and possible coronary artery disease.

The data collected for this study are unique. At the time of the study, MPS was not a routine part of the work-up of patients with suspected coronary artery disease in Denmark. The patients were isolated on the island of Funen. Patients were sent directly for angiography from local general practitioners or through local hospitals. Before angiography, they received MPS, and only after the angiogram were the results available to the investigator or clinician. This local practice variation enabled a comparison of MPS with CA in a group of patients who would not normally be selected for angiography based on the results of the MPS study, thus eliminating the effect of 'referral' or 'work-up' bias on the correlation of MPS and CA. Work-up bias has been examined in the literature, and the overall effect is that a greater number of angiograms are performed in patients with abnormal scans. This increases the sensitivity and reduces the specificity of the test being evaluated – in this case, the MPS (4,5).

This study has a number of companion publications (6-8); the purpose of the original study was to compare patient findings on MPS with those found on CA. In this population, the overall sensitivity of MPS was found to be 75% and the specificity was 79% (8). At publication, these results were attributed to the absence of a referral or work-up bias.

The focus of the current publication was the correlation of clinical features to the angiogram. Three clinical assessment scales were used to classify patients undergoing angiography. The subgroups of patients who were thought to have a high probability of disease were then segregated, and the correlation to significant angiographic coronary artery disease was determined. The first grouping was based on the chest pain classification from the Diamond and Forrester scale (9), which separates chest pain into three groups based on patient history. The chest pain types include typical angina, atypical angina and noncardiac chest pain. For the analysis, the group with typical angina was chosen as the high-risk group. The second grouping involved classifying anginal pain based on CCS class and further classifying patients into groups with a pain score lower than grade 2 and a pain score grade 2 or higher. The last classification was based on the complete Diamond and Forrester criteria, estimating the overall probability of coronary artery disease. This scale uses the above classification of chest pain, as well as the sex and age of the patient. The group chosen included patients who were thought to have a greater than 85% probability of disease.

The sensitivity and specificity of the symptom of typical angina, compared with the presence of angiographic lesions of at least 50%, were 71% and 58%, respectively. The remainder of the results were presented as scores combined from the risk classifications. Given the interdependence of these scores, a composite analysis might not have been reliable and might not necessarily have had added predictive ability. The data are, however, presented so that sensitivity and specificity of chest pain of CCS grade 2 or higher may be calculated. This symptom is a more specific manifestation of coronary artery disease that resulted in a fall in sensitivity of this clinical feature to 64% and a similar specificity of 59%. For the finding of high pretest likelihood, based on the Diamond and Forrester criteria, the sensitivity again fell to 62% and the specificity fell to 72%. The authors' analysis suggested that the clinical classification failed to classify patients into the risk group that was eventually provided by their angiograms, even when these interdependent scores were combined.

The characteristics of performance of a symptom, sign or test are described by sensitivity and specificity. In this way, the test's performance characteristics can be compared with an established 'gold standard' and with each other. Traditional teaching suggests that the sensitivity and specificity of a test is a test characteristic, and that it should not change based on the patient population to which it is applied (10). There are a number of reasons why the sensitivity and specificity of the authors' clinical measures did not perform as well as one may

Division of Cardiology, University of Ottawa Heart Institute, Ottawa, Ontario

Correspondence and reprints: Dr Renee C Hessian, Ottawa Heart Institute, 40 Ruskin Street, Ottawa, Ontario K1Y 4W7.

Telephone 613-761-4377, fax 613-761-5022, e-mail rhessian@ottawaheart.ca

Received for publication April 24, 2007. Accepted May 13, 2007

have expected. These should be examined before concluding, as Høilund-Carlsen et al did, that the clinical examination of patients "appears to be unreliable".

There is a small variation in test performance depending on the specific nuclear protocol used. The sensitivity was 92% (82% to 98%) in the pooled literature using tomographic single photon emission computed tomography technology (4) and the specificity was 68%. The specific protocol used by the authors was a one-day dual-isotope protocol with a stress technetium 99m sestamibi/rest thallium-201 single photon emission computed tomography imaging protocol. With this protocol, Berman et al (11) reported a sensitivity of 91% and specificity of 75%. Although Høilund-Carlsen et al's sensitivity fell short of Berman et al's benchmark, the specificity was similar. The patients were stressed with bicycle ergometry (50%) or by adenosine stress (50%). The authors did not identify any technical issues contributing to their observed reduction in sensitivity and believed that this reduction was entirely attributable to the lack of referral bias in their population. Other authors have agreed that this mechanism has also reduced the sensitivity of stress echocardiography when used as a diagnostic test for the presence of coronary artery disease (12).

The population Høilund-Carlsen et al examined was not comparable with most of the literature to which they contrasted their data. The authors did not examine consecutive patients; instead, they excluded patients who were older than 75 years of age, and had diabetes or known cardiac disease, prior revascularization or infarction, as well as those with the absence of angina and those with a number of noncardiac diagnoses. One hundred nineteen of 476 patients (25%) were excluded from analysis based on the presence of infarction or revascularization. Nonconsecutive inclusion of patients has been shown to increase the estimated diagnostic accuracy of the test (13).

There is no recent literature using the Diamond and Forrester predictive tool, but in past publications, the clinical features of patient history have been compared with angiographic disease. In the publication by Sox et al (14), predictors of coronary artery disease were established from a group of patients who were awaiting angiography at Stanford University (Palo Alto, USA). In this group, the greatest predictors of coronary artery disease were age older than 60 years, typical angina, history of infarction, male sex and a history of smoking. The only exclusions from this group were patients who did not give proper consent. In Pryor et al (15,16), the modelling was completed differently. Patients were questioned after they had a noninvasive work-up and were referred for angiography. Again, no patients were excluded from follow-up, and the historical factors that were of importance in predicting the presence of coronary artery disease included pain type, sex, prior infarction, age, smoking, hyperlipidemia, diabetes, and ST and T wave changes. The most important characteristic was the type of chest pain experienced by the patient. This predictive ability of their model, however, was significantly improved by the addition of other historical and risk factor information.

The patient population explored by testing is very important. The sensitivity and specificity of a test are best reproduced in a patient population that is similar to the one in which the test was developed. On the surface, the population in Høilund-Carlsen et al's study appears to be similar to those of other studies carried out in centres with angiography. There may be differences in how patients are referred for angiography,

but the indication for angiography and the determination of the risks of the test should be fairly similar.

This study presented a population in which many of the higher-risk patients were removed from analysis. The overall normal angiography rate in this population was 65%. It is thus not surprising that the population contained a large population of women (46%) and patients with hypertension (64%). The average age was also young – 57 years. Both women and patients with hypertension have anginal-type symptoms that can occur without the presence of obstructive coronary artery disease. Including women in diagnostic studies decreases the sensitivity and has no effect on the specificity of noninvasive testing (17).

In the examined cohort, there was a low use of medical therapy, especially acetylsalicylic acid, and only one-half of patients with hypercholesterolemia were treated with statins. Only 28 patients of the cohort, or 8%, had three-vessel disease. Overall, this population had mild and noncritical coronary artery disease. This is partly related to the exclusion of certain patients, but the patient population resembled a primary practice population rather than a referral population.

If this population is compared with the community practice population in the publication by Sox et al (14), similarities can be seen. If the historical information that was determined to be important in the preangiography patients, after being formulated into a chest pain score, is applied to the community setting, the likelihood of having coronary artery disease increases significantly, because the chest pain score increases. At low levels of the chest pain score, there is a very low prevalence of coronary artery disease compared with the preangiography population. So, with this comparison, the findings of these authors were not that different from the literature.

A meta-analysis of studies examining factors that affect the sensitivity and specificity of MPS (17,18) suggested that sensitivity is higher in studies that include patients with prior infarctions and higher percentages of men. Most studies in this review included a population in which more than 80% were men. These factors were important in univariate and multivariate analyses. The presence of work-up bias reduced specificity, but it did not affect sensitivity. In a similar publication (18) examining sensitivity of clinical features, but using treadmill testing, the sensitivity of typical chest pain type was very high – 79.6%. Again, a different gold standard was used, but consistency was obtained using the symptom of typical angina in the prediction of the presence of coronary artery disease.

Høilund-Carlsen et al have pointed out a discrepancy in sensitivity and specificity in their results that they found to be unexpected. Other than some of the technical issues that have been discussed, there remain several issues that can alter the sensitivity and specificity of prediction tools. The effect of the spectrum of the disease tested can have a significant effect on the characteristics of the performance of a diagnostic test (19-21).

This issue is relevant in this particular situation because the population that was tested by the authors was different from the literature. The traditional teaching of constant sensitivity and specificity of a particular test may not be the same when the patient populations being compared are dramatically different. This source of bias was first described by Ransohoff and Feinstein in 1978 (22). However, there is disagreement about whether this source of bias exists. Some investigators explain that subgroups can be found in which a test performs differently when used within a population. Some features that are known to cause a

test to perform differently include sex, age, severity of disease and clinical venue from which the patient is referred (20).

Severity of disease is a frequent cause of bias in the diagnosis of disease, especially in cardiac disease. Many of our diagnostic tests have been validated in teaching centres, especially in patient populations undergoing angiography. This group of patients, by definition, has more severe disease than patients who would be in a physician's office but not be referred for angiography. This alone allows an easier diagnosis to be made, and diagnostic tests are known to overperform in this situation. A recent review (5) of the potential bias in diagnostic testing confirms that the sensitivity of testing increases with disease severity. Given the milder nature of disease in the study by Højlund-Carlsen et al, the sensitivity may be lower than expected for the given clinical tools.

Another issue of spectrum bias occurs when the category of the diseased or not diseased is not naturally dichotomized (23). This is a common issue in the continuum of coronary stenoses. An example of this occurred in this publication; the authors needed to decide the level at which coronary stenosis patients would be considered to have disease. A lesion of 50% was chosen, which allows the potential for misclassification for lesions that are near this cut-off point. This has an effect on sensitivity and specificity, depending on the frequency of misclassification and the prevalence of lesions near 50% in the population. In some literature in which the degree of disease is more severe than in this population, misclassification of disease at this level

is not likely to affect the test's performance characteristics. However, in this group, the severity of disease was shifted toward the milder spectrum, allowing a greater possibility of misclassification. This allowed for interaction among disease prevalence, severity, and test characteristics of sensitivity and specificity.

The paper by Højlund-Carlsen et al has provided interesting information with regard to the use of clinical features in the prediction of coronary artery disease and has offered a unique analysis of patients referred for angiography without a referral bias. The results obtained by the authors might have been related to factors that are identified by themselves and by us.

The use of clinical features of angina as a diagnostic tool for coronary artery disease has been challenged. However, limitations related to other biases make it difficult to draw definitive conclusions from the work performed. The study pointed out that not every patient with angina has coronary artery disease. It also pointed out the difficulties in evaluating clinical features and diagnostic tests in a given patient population, which may be biased. Further investigations that minimize bias are needed. It may be appropriate to re-evaluate clinical criteria established in a different era and in a different patient population than we face today. Nevertheless, a good history of angina pectoris, along with other clinical features, remain key aspects in the evaluation and diagnosis of hemodynamically significant coronary artery disease. For now, at least, the cornerstone remains.

REFERENCES

- Heberden W. Some account of a disorder of the breast. *Medical Transactions of The Royal College of Physicians of London*, 1772;2:59-67.
- Heberden W. Pectoris dolor. In: Payne T, ed. *Commentaries on the History and Cure of Diseases*. London, Mews-Gate, 1802:362-8.
- Khan IA, Mehta NJ. Initial historical descriptions of the angina pectoris. *J Emerg Med* 2002;22:295-8.
- Beller GA, Zaret BL. Contributions of nuclear cardiology to diagnosis and prognosis of patients with coronary artery disease. *Circulation* 2000;101:1465-78.
- Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: A systematic review. *Ann Intern Med* 2004;140:189-202.
- Højlund-Carlsen PF, Johansen A, Christensen HW, et al. Potential impact of myocardial perfusion scintigraphy as gatekeeper for invasive examination and treatment in patients with stable angina pectoris: Observational study without post-test referral bias. *Eur Heart J* 2006;27:29-34.
- Johansen A, Højlund-Carlsen PF, Christensen HW, et al. Observer variability in the evaluation of dual-isotope Tl-201/Tc-99m sestamibi rest/stress myocardial perfusion SPECT in men and women with known or suspected stable angina pectoris. *J Nucl Cardiol* 2004;11:710-8.
- Johansen A, Højlund-Carlsen PF, Christensen HW, et al. Diagnostic accuracy of myocardial perfusion imaging in a study population without post-test referral bias. *J Nucl Cardiol* 2005;12:530-7.
- Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. *N Engl J Med* 1979;300:1350-8.
- Rothman KJ. *Epidemiology: An Introduction*. Oxford University Press, 2002:201.
- Berman DS, Kiat H, Friedman JD, et al. Separate acquisition rest thallium-201/stress technetium-99m sestamibi dual-isotope myocardial perfusion single-photon emission computed tomography: A clinical validation study. *J Am Coll Cardiol* 1993;22:1455-64.
- Roger VL, Pellikka PA, Bell MR, Chow CW, Bailey KR, Seward JB. Sex and test verification bias. Impact on the diagnostic value of exercise echocardiography. *Circulation* 1997;95:405-10.
- Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ* 2006;174:469-76.
- Sox HC Jr, Hickam DH, Marton KI, et al. Using the patient's history to estimate the probability of coronary artery disease: A comparison of primary care and referral practices. *Am J Med* 1990;89:7-14. (Erratum in 1990;89:550).
- Pryor DB, Harrell FE Jr, Lee KL, Califf RM, Rosati RA. Estimating the likelihood of significant coronary artery disease. *Am J Med* 1983;75:771-80.
- Pryor DB, Shaw L, McCants CB, et al. Value of the history and physical in identifying patients at increased risk for coronary artery disease. *Ann Intern Med* 1993;118:81-90.
- Detrano R, Janosi A, Lyons KP, Marcondes G, Abbassi N, Froelicher VF. Factors affecting sensitivity and specificity of a diagnostic test: The exercise thallium scintigram. *Am J Med* 1988;84:699-710.
- Hlatky MA, Pryor DB, Harrell FE Jr, Califf RM, Mark DB, Rosati RA. Factors affecting sensitivity and specificity of exercise electrocardiography. Multivariable analysis. *Am J Med* 1984;77:64-71.
- Pewsnar D, Battaglia M, Minder C, Marx A, Bucher HC, Egger M. Ruling a diagnosis in or out with "SpIn" and "SnNOut": A note of caution. *BMJ* 2004;329:209-13.
- Mulherin SA, Miller WC. Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. *Ann Intern Med* 2002;137:598-602.
- Montori VM, Wyer P, Newman TB, Keitz S, Guyatt G; Evidence-Based Medicine Teaching Tips Working Group. Tips for learners of evidence-based medicine: 5. The effect of spectrum of disease on the performance of diagnostic tests. *CMAJ* 2005;173:385-90.
- Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978;299:926-30.
- Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med* 1997;16:981-91.