Effects of chronotropic incompetence and β-blocker use on the exercise treadmill test in men

Andre J. Gauri, MD,[a] Vinod K. Raxwal, MD,[b] Larissa Roux, MD, MPH,[c] William F. Fearon, MD,[a] and Victor F. Froelicher, MD, FACC[b] ***Palo Alto, Calif, and Calgary, Alberta, Canada***

Objective Our purpose was to assess the diagnostic characteristics of the exercise test in patients who fail to reach conventional target heart rates and in patients on β-blockers.

Background Exercise test results are often considered "inadequate" or "nondiagnostic" in patients taking β-blockers and in patients who do not achieve 85% of their age-predicted maximal heart rate.

Methods The results of exercise tests and coronary angiography performed to evaluate chest pain in 1282 male patients without a prior history of myocardial infarction, coronary revascularization, diagnostic Q wave on the baseline electrocardiogram, or previous cardiac catheterization were analyzed with respect to β-blocker exposure and failure to reach 85% age-predicted maximal heart rate. Sensitivity, specificity, and predictive accuracy of exercise testing, as well as area under the curve for the receiver operating characteristic plots were calculated for these subgroups with use of coronary angiography as the reference. The angiographic criterion for significant coronary artery disease was 50% narrowing or greater in one or more major coronary arteries.

Results The population was divided into 4 exclusive groups on the basis of whether they reached their target heart rates and whether they were receiving β-blockers. Sixty to 40 percent of this clinical population failed to reach target heart rate, of which 24% (n = 303) were receiving β-blockers and 40% (n = 518) were not. The group of patients who reached target heart rate and were not taking β-blockers was taken as the reference group (n = 409). The group of patients supposedly β-blocked but who reached the target heart rate (n = 52) had hemodynamic and test characteristics similar to those of the reference group and most likely were not taking their β-blockers or were not adequately dosed. The prevalence of angiographic coronary disease was significantly higher in the 2 groups failing to reach target heart rate, both in the presence and absence of β-blockers, compared with the reference group (68% and 64%, respectively, vs 49%, $P < .01$). Although the areas under the curve of the receiver operating characteristic curves for ST depression of the groups failing to reach target heart rate were not significantly different from the reference group, the predictive accuracy and sensitivity were significantly lower for 1 mm of ST depression in the β-blocked group who did not reach target heart rate (predictive accuracy of 56% vs 67%, sensitivity of 44% vs 58%, $P < .01$). The only way to maintain sensitivity with the standard exercise test in the β-blocker group who failed to reach target heart rate was to use a treadmill score or 0.5-mm ST depression as the criteria for abnormal.

Conclusion Sensitivity and predictive accuracy of standard ST criteria for exercise-induced ST depression are significantly decreased in male patients who are taking β-blockers and do not reach target heart rate. In those who fail to reach target heart rate and are not β-blocked, sensitivity and predictive accuracy are maintained. (Am Heart J 2001;142:136-41.)

See related Editorial on page 1.

All the guidelines relative to ischemic heart disease recommend the standard exercise test as the first choice for the evaluation of the patient who has chest pain and does not have resting electrocardiogram (ECG) abnormalities that affect the interpretation of the exercise ECG response.[1-4] However, the diagnostic accuracy of the standard exercise test in patients taking β-adrenergic blockers or who have not reached an age-predicted target heart rate has been questioned. Often these tests are recorded as indeterminate or nondiagnostic and more expensive or invasive diagnostic testing is performed. Interpretation of test results in these populations as nondiagnostic may not be justified and may affect diagnostic characteristics[5] of the test in general. The purpose of this study was to examine the effect of β-blockers and chronotropic incompetence on the diagnostic accuracy of the exercise test. In particular, it is

From the Divisions of Cardiovascular Medicine, [a]Stanford University Medical Center and the [b]Veterans Affairs Palo Alto Health Care System, Palo Alto, Calif, the [c]Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada.
Submitted Oct 18, 2000; accepted Jan 11, 2001.
Reprint requests: Victor Froelicher, MD, Department of Cardiology 111c, PAVAHCS, Bldg 100, Room E2-441, 3801 Miranda Ave, Palo Alto, CA 94304.
E-mail: vicmd@aol.com, www. cardiology.org
4/1/115788
doi:10.1067/mhj.2001.115788

important to determine whether these groups of patients can safely avoid further diagnostic scrutiny (ie, whether the predictive accuracy of the exercise ECG test in these important subgroups is comparable to that of the general at-risk population).

Methods

Patients

Eight thousand male patients underwent treadmill testing at two Veterans Administration Medical Centers between 1987 and 1998. Of these patients, 3454 were evaluated for chest pain with exercise treadmill testing followed by coronary angiography performed within 3 months. Patients with previous cardiac surgery or angiography, valvular heart disease, left bundle branch block, paced rhythms, or a Wolff-Parkinson-White pattern on the resting ECG, were excluded from the study. Patients with previous myocardial infarction by history or by Q-wave criteria were also excluded from the study to minimize the effects of measurement bias. A total of 1282 patients with chest pain and coronary angiography within 3 months of exercise testing were found to have met the inclusion and exclusion criteria of our study.

Subjects so accrued were classified on the basis of exposure into 4 groups. As noted above, the main exposures (or baseline characteristics) of interest were the use of β-blockers and the inability to reach target heart rates during the entry exercise test. Data were compiled from computerized records and from retrospective chart reviews if computerized records were found to be incomplete.

Exercise testing

Patients underwent treadmill testing with the US Air Force School of Aerospace Medicine[6] protocol or an individualized ramp treadmill protocol.[7] Before ramp testing, patients completed a survey evaluating their current activity levels from presented activities, listed by increasing metabolic equivalent (METS) order. This questionnaire estimated patients' self-reported exercise capacities before the test and thus allowed most patients to reach maximal exercise at approximately 10 minutes.[8] Visual ST-segment deviation was measured at the J junction and corrected for pre-exercise ST-segment depression while standing. ST slope was measured over the following 60 milliseconds and classified as upsloping, horizontal, or downsloping. Slope was coded as 1 for horizontal, 2 for downsloping if there was 0.5 mm or more junctional depression, and 0 or normal slope otherwise. The ST response considered was the most horizontal or downsloping ST segment depression in any lead except aVR during exercise or recovery. An abnormal response was specifically coded as either one half or 1 mm or more of horizontal or downsloping ST-segment depression.

No test was classified as indeterminate, medications were not withheld, and no patients were excluded for inadequate heart rate responses.[9] After maximal heart rate was regressed on age in our male veteran population, the linear regression equation of $210 - 0.8 \times (\text{Age})$ was derived for age-predicted maximal heart rate. This equation serves as a reasonable estimate for the wide range of published regression equations for age-predicted maximal heart rate.[10,11] Inadequate heart rate response or chronotropic incompetence was defined as the inability to reach 85% of age-predicted maximal heart rate. The exercise tests were performed, analyzed, and reported per standard protocol and using a computerized database (EXTRA, Mosby, Chicago).[12]

Coronary angiography

Decisions for cardiac catheterization were consistent with clinical practice. Coronary artery narrowing was visually estimated and expressed as percent lumen diameter stenosis. Patients with a 50% narrowing of the left main, left anterior descending, left circumflex, or right coronary arteries (or their major branches) were considered to have significant angiographic coronary artery disease. The 50% criterion was chosen to be consistent with definitions used by the Coronary Artery Bypass Graft Surgery Trialists' Collaboration.[13]

Scores

The pretest score of Morise et al was calculated to determine whether the exercise test had incremental diagnostic value in the subgroups and the Duke treadmill score was calculated to determine whether it improved the diagnostic characteristics of the exercise test (see Appendix).[14,15] A Duke treadmill score less than 1 was considered abnormal.

Statistical methods

Patients were categorized according to the presence or absence of a β-blocker from their medical regimen and whether they reached 85% age-predicted maximal heart rate. Comparisons among the subgroups on continuous variables were performed with the analysis of variance statistical method. Comparisons between categorical variables were performed with the χ^2 test. With use of angiographic evidence of coronary disease as the reference, area under the curve (AUC) of receiver operating characteristic (ROC) plots was determined for the various subgroups. A significant difference between AUCs for each subgroup was defined as a z-score >1.96, with z-score = $(AUC_1 - AUC_2)/(SE_1^2 + SE_2^2)^{1/2}$. Statistical analysis was performed with the Number Crunching System Software (Salt Lake City, Utah).

Results

Population characteristics

Table I summarizes baseline features of the 1282 subjects in 4 categories according to β-blocker use and target heart rate. Sixty-forty percent of this clinical population failed to reach target heart rate, of which 24% (n = 303) were receiving β-blockers and 40% (n = 518) were not. The group of patients who reached target heart rate and were not taking β-blockers was taken as the reference group (n = 409). The smallest subgroup (n = 52) was those supposedly taking β-blockers but who reached target heart rate. All of their clinical and exercise test responses were similar to the group who reached target heart rate and were not receiving β-blockers, making it appear that either they did not take

Table I. Baseline characteristics

	No β-blocker/ target HR reached	β-Blocked/ target HR reached	No β-blocker/ failed to reach	β-Blocked/ failed to reach target HR
No. (% total)	409 (32%)	52 (4%)	518 (40%)	303 (24%)
Mean age ± SD (y)	58 ± 11	56 ± 13	60 ± 10	60 ± 10
Body mass index ± SD	28 ± 5	28 ± 5	28 ± 5	28 ± 4
CHF (%)	3.2	0	4.1	2.0
HTN (%)	46	58	48	68*
COPD (%)	8.6	0	6.8	4.3*
Claudication (%)	7.8	5.8	9.3	10.2
Diabetes (%)	13	12	16	15
Current smokers (%)	28	21	36*	35*
Resting HR ± SD	86 ± 15	81 ± 14	78 ± 13*	69 ± 13*

HR, Heart rate; *CHF,* congestive heart failure; *HTN,* hypertension; *COPD,* chronic obstructive lung disease.
*$P < .01$ compared with the reference group reaching target heart rate and not β-blocked.

Table II. Hemodynamic, angiographic, and treadmill characteristics

	No β-blocker/ target HR reached	No β-blocker/ failed to reach	β-Blocked/ failed to reach
Angiographic CAD (%)	49	64*	68*
Max HR ± SD	153 ± 15	120 ± 14*	109 ± 16*
Max SBP ± SD	180 ± 26	164 ± 29*	161 ± 29*
METS ± SD	8.8 ± 3.4	6.6 ± 2.6*	7.0 ± 2.7*
1-mm Exercise-induced STD (%)	41	37	36
Sensitivity of 1-mm STD	58	50	44*
Specificity of 1-mm STD	75	86*	81
Predictive accuracy of 1-mm STD	67	63	56*
Sensitivity of Duke treadmill score	55	57	57
Specificity of Duke treadmill score	77	83	76
Predictive accuracy of Duke treadmill score	66	66	63

HR, Heart rate; *CAD,* coronary artery disease; *SBP,* systolic blood pressure; *STD,* ST depression.
*$P < .01$ compared with the reference group reaching target heart rate and not β-blocked.

their prescribed β-blocker or were inadequately dosed. Because of their small number and similar responses to the reference group, they were excluded from further comparison.

The 3 groups remaining for comparison had similar clinical characteristics. There were no statistically significant differences in age, body mass index, or prevalence of congestive heart failure, claudication, or diabetes. The β-blocked group did, however, have a significantly higher prevalence of hypertension (68% vs 46%, $P < .01$) and a significantly lower prevalence of chronic obstructive lung disease (4.3% vs 8.6%, $P < .01$) compared with the reference group. Also, both groups failing to reach target heart rate had more current smokers (36% and 35% vs 28%, $P < .01$) and their resting heart rates were significantly lower (78 ± 13 and 69 ± 13 vs 86 ± 15, $P < .01$).

Angiographic and exercise test characteristics

Hemodynamic responses to exercise and diagnostic accuracy of exercise testing are summarized in Table II. There was no difference in left ventricular ejection fraction determined at angiography in any of the subgroups (average ejection fraction 65%). The 2 groups who failed to reach target heart rate, both in the presence and absence of β-blocker, had a significantly higher prevalence of angiographic disease than did the group who reached target heart rate and were not receiving β-blockers (68% and 64%, respectively, vs 49%, $P < .01$). The 2 groups failing to reach target heart rate also achieved significantly lower maximal heart rates and systolic blood pressures and achieved METS.

The sensitivity (50%) and specificity (80%) of exercise-induced ST depression for all 1282 subjects were similar to those reported elsewhere in the literature.[16,17] One millimeter of exercise-induced ST depression was the least sensitive (44%) and had the lowest predictive accuracy (56%) in the group that was β-blocked and failed to achieve target heart rate in spite of similar AUC for the ROC curves in all groups (Table III). Because of this finding, a less stringent ST-depression criterion of 0.5 mm was tried in this subgroup. The sensitivity and pre-

Table III. Area under the ROC curves in the 3 main groups for the Morise pretest score, exercise-induced ST depression, and the Duke Treadmill score

	No β-blocker/ reached HR target	No β-blocker/ failed to reach target HR	β-Blocked/ failed to reach target HR
Pretest score	0.64	0.62	0.69
Exercise-induced ST depression	0.67	0.70*	0.66
Duke treadmill score	0.72	0.72	0.75†

HR, Heart rate.
*z-score >1.96 compared with AUC of pretest score for same subgroup.
†z-score >1.96 compared with AUC of exercise-induced ST depression for same subgroup.

dictive accuracy in this group improved significantly, with values comparable to results for 1-mm ST depression in the other groups (sensitivity 54%, specificity 76%, and predictive accuracy 61%).

Scores

Based on ROC analysis, exercise-induced ST depression did not provide incremental value in the group on β-blockers who did not achieve target heart rate (AUC 0.69 pretest score vs 0.66 for ST) but did so in the other groups, as demonstrated by the greater AUC (Table III). The Duke treadmill score significantly improved the diagnostic characteristics over the ST response only in the subgroup on β-blockers who failed to reach target heart rate.

Discussion

This study has sought to define the role and utility of exercise testing in the identification of coronary artery disease in patients on β-blockers and in patients who fail to reach the conventional heart rate targets on which this test is based. Controversies about the significance and interpretability of treadmill test results must be placed in the context of the dynamic state of diagnostic cardiology, in which numerous modalities with varying risks and capabilities are available to clinicians. Specifically, the relatively safe and noninvasive treadmill test has been used as a means to select patients who might be more likely to benefit from cardiac catheterization, with its inherent risks and costs, than the general population of cardiac patients.

Unfortunately, the diagnostic properties of exercise testing are not well defined, despite its widespread use, because of the characterization of a large proportion of test results as "nondiagnostic." Although it may be true that in patients being evaluated for coronary heart disease that are considered to have nondiagnostic test results (because of chronotropic incompetence or β-blocker use) are possibly at high enough risk to justify immediate referral for cardiac catheterization, it is still valuable to know the properties of the treadmill test in this group because this test is often the initial diagnostic modality of choice.

With respect to β-blocker use associated with failure to reach target heart rate, our study demonstrates that in these patients the standard 1-mm ST criterion has a lower sensitivity and predictive accuracy than our reference group. It is notable that in our cohort of patients the prevalence of angiographic coronary artery disease is significantly higher in this subgroup. Although the lowered predictive accuracy is explained in part by the higher disease prevalence, it does not explain the lower sensitivity. However, the use of 0.5-mm ST criterion or the Duke treadmill score provided similar test characteristics to that obtained with use of the standard 1-mm criterion in the reference group.

The issue of chronotropic incompetence not associated with β-blockade is somewhat more complex. Inability to reach target heart rate carries prognostic importance in terms of mortality, even when adjustments are made for myocardial perfusion.[18] Maximum-attained heart rate on the treadmill test is therefore an important measurement. However, failure to achieve the standardized target heart rate may not preclude meaningful interpretation of the test with respect to coronary artery disease. Consideration of these tests as nondiagnostic leads to loss of information, alters test characteristics, and may not be justified. Although heart rate targets were initially used in exercise testing because of safety concerns, the many limitations make this practice unwise. For instance, there is a high degree of variability in maximal heart rate. The regression equations used typically have correlation coefficients between –0.4 and –0.6 with an SE of estimate of 10 to 25 beats/min.[10] In our subgroup that failed to reach heart rate target not due to β-blockade, there was no difference in the test characteristics for 1-mm ST depression from the reference group. In addition, this group exhibited the greatest incremental improvement in diagnosis over the pretest score (AUC of 0.62 vs 0.70, z >1.96).

The above findings suggest that the cost of using target heart rate attainment as a condition for test adequacy may be too high when the number of excluded patients and the changes in diagnostic properties of the test are considered. Target heart rates themselves have been noted to be highly arbitrary and to have little support for use in treadmill testing in the literature. A large meta-analysis reviewing 150 study groups and over 24,000 patients found that target heart rate values used varied between 70% and 100% of age-predicted maximum.[19] The number of patients that were able to achieve an adequate heart rate was only reported in 15 of the 150 study groups. Many patients failing to reach an adequate heart rate and having no exercise-induced ST depressions were either excluded from the study or recorded as nondiagnostic. The Coronary Artery Surgery Study,[20] for example, excluded approximately one third of its subjects from analysis because of chronotropic insufficiency (Weiner D, personal communication). Similarly, 40% of the patients in our study fit into this traditionally nondiagnostic group with inadequate heart rate response not due to β-blockade. Our study found that 64% of these subjects had angiographic coronary disease. Exclusion of such a large proportion of patients with high disease prevalence has imposed severe limitations on the interpretability of data from many studies of the diagnostic properties of treadmill testing.

Limitations

In this study sensitivity, specificity, and predictive accuracy of treadmill testing in the diagnosis of coronary artery disease were compared among subgroups of the cohort. Selection of subjects for entry into the cohort was highly prone to bias as only a minority of patients from the original group underwent both exercise testing and angiography in rapid succession—many individuals considered at exercise testing to be at low risk for coronary artery disease were excluded from analysis. Certainly, sicker patients were selected for prescription of β-blockade. Test properties observed in this study therefore cannot be generalized to the entire population of individuals undergoing exercise testing because the test was not evaluated in a large number of patients. Observed differences in test properties may have been dampened by this "work-up bias" because the prior probability of disease may have been significantly different in the study population and the group of patients initially referred for exercise testing. Because our study population was all male, we cannot assume that the same conclusions apply to women undergoing exercise testing for the evaluation of coronary artery disease. Test characteristics may also have been influenced by the significant difference in coronary disease prevalence rates in subgroups.

Conclusions

Patients taking β-blockers and those with chronotropic insufficiency are already at increased risk of adverse cardiac events[21] and may benefit, in many instances, from cardiac catheterization. However, exactly which of these patients will benefit is still not clear. Refinements in exercise testing and other predictive modalities may save cost and morbidity when such patients are being evaluated for coronary disease. This study demonstrates that exercise-induced ST depression can be used in guiding the decision to proceed to angiography in patients who fail to reach target heart rate and are not receiving β-blocker, but that treadmill scores or 0.5-mm ST depression criterion are required in those taking β-blockers who fail to reach target heart rate to maintain the sensitivity and predictive accuracy of the exercise test.

References

1. Fletcher GF, Balady G, Froelicher VF, et al. Exercise standards: a statement for health professionals from the American Heart Association. Circulation 1995;91:580-632.
2. Gibbons RJ, Balady GJ, Beasley JW, et al. ACC/AHA guidelines for exercise testing. J Am Coll Cardiol 1997;30:260-311.
3. O'Rourke RA, Brundage BH, Froelicher V, et al. American College of Cardiology/American Heart Association expert consensus document on electron-beam computed tomography for the diagnosis and prognosis of coronary artery disease. Circulation 2000;102:126-40.
4. Braunwald E, Antman EM, Beasley JW, et al. ACC/AHA guidelines for the management of patients with unstable angina and non-ST-segment elevation myocardial infarction: executive summary and recommendations. Circulation 2000;102:1193-209.
5. Philbrick JT, Horwitz RI, Feinstein AR. Methodologic problems of exercise testing for coronary artery disease: groups, analysis and bias. Am J Cardiol 1980;46:807-12.
6. Wolthuis R, Froelicher VF, Fischer J, et al. New practical treadmill protocol for clinical use. Am J Cardiol 1977;39:697-700.
7. Myers J, Buchanan N, Walsh D, et al. A comparison of the ramp versus standard exercise protocols. J Am Coll Cardiol 1991;17:1334-42.
8. Myers J, Do D, Herbert W, et al. A nomogram to predict exercise capacity from a specific activity questionnaire and clinical data. Am J Cardiol 1994;73:591-6.
9. Reid M, Lachs M, Feinstein A. Use of methodological standards in diagnostic test research. JAMA 1995;274:645-51.
10. Hammond HK, Froelicher VF. Normal and abnormal heart rate responses to exercise. Prog Cardiovasc Dis 1985;27:271-96.
11. Morris CK, Meyers J, Froelicher VF, et al. Normogram based on metabolic equivalents and age for assessing aerobic exercise capacity in men. J Am Coll Cardiol 1993;22:175-82.
12. Shiu P, Froelicher V. Extra: an expert system for exercise reporting. J Noninvasive Test 1998;II-4:21-7.
13. Yusuf S, Zucker D, Peduzzi P, et al. Effect of coronary artery bypass graft surgery on survival: overview of 10-year results from randomized trials by the Coronary Artery Bypass Graft Surgery Trialists Collaboration. Lancet 1994;344:563-70.
14. Morise AP, Haddad J, Beckner D. Development and validation of a clinical score to estimate the probability of coronary artery disease in men and women presenting with suspected coronary artery disease. Am J Med 1997;102:350-6.
15. Shaw LJ, Peterson ED, Shaw LK, et al. Use of a prognostic treadmill

score in identifying diagnostic coronary disease subgroups. Circulation 1998;98:1622-30.
16. Froelicher VF, Lehmann KG, Thomas R, et al. The electrocardiographic exercise test in a population with reduced workup bias: diagnostic performance, computerized interpretation, and multivariable prediction. Ann Intern Med 1998;128:965-74.
17. Gianrossi R, Detrano R, Mulvihill D, et al. Exercise-induced ST depression in the diagnosis of coronary artery disease: a meta-analysis. Circulation 1989;80:87-98.
18. Lauer MS, Paskow F, Cole N, et al. Impaired heart rate response to graded exercise: prognostic implications of chronotropic incompetence in the Framingham Heart Study. Circulation 1996;93:1520-6.
19. Detrano R, Gianrossi R, Froelicher VF. The diagnostic accuracy of the exercise electrocardiogram: a meta-analysis of 22 years of research. Prog Cardiovasc Dis 1989;32:173-206.
20. Weiner DA, Ryan TJ, McCabe CH, et al. Exercise stress testing: correlations among history of angina, ST-segment response and prevalence of coronary-artery disease in the Coronary Artery Surgery Study (CASS). N Engl J Med 1979;301:230-5.
21. Lauer MS, Paskow F, Cole N, et al. Impaired chronotropic response to exercise stress testing as a predictor of mortality. JAMA 1999; 281:524-9.

Appendix

Scores

Morise et al pretest score. Age code + Angina pectoris code · 5 + diabetes · 2 + Hypertension + Smoking now + Hypercholestrolemia + Family history of coronary artery disease (CAD) + Obesity.

Where age is less than 40 = 3, age between 40 to 55 = 6, and age more than 55 = 9. For estrogen status, 3 points were subtracted for positive and 3 points were added for estrogen-negative status. Typical chest pain = 5, atypical chest pain = 3, nonanginal chest pain = 1, and no chest pain = 0. For diabetes mellitus, 2 points were added and 1 point was added for each of the other 5 risk factors (hypertension, current smoking, hypercholesterolemia, family history of CAD, and obesity).

Duke treadmill score. Duration of exercise in minutes – (5 · Maximal ST-segment deviation during or after exercise in millimeters) – (4 · Treadmill angina index). Where angina index has a value of 0 if the patient had no angina during exercise, 1 if the patient had nonlimiting angina, and 2 if angina was the reason the patient stopped exercising.

Glossary

- Equation—Mathematical representation of the result of a multivariable statistical technique that attempts to discriminate those with and without disease
- Code—A numeric value for the variables included in an equation or score
- Score—A simplified version of an equation that only requires adding or subtracting of coded points
- Multiple logistic model—A multivariable statistical technique that attempts to discriminate those with and without disease and provides a probability of being in the diseased group from 0 to 1 calculated by a log equation
- ROC—Receiver operating characteristic curve is a graphic representation of the relationship between sensitivity and specificity for a diagnostic test
- AUC—Area under the ROC curve is a measure of how well the model separates patients with and without a given outcome (CAD). The AUC ranges from 0 to 1, with 0.5 corresponding to no discrimination (ie, random performance), 1.0 to perfect discrimination, and values less than 0.5 to worse-than-random performance.
- Portability—Ability of a score or equation to discriminate in other than the population in which it is derived
- Calibration—How well the cut points of a score or equation correlate with actual disease probabilities in different populations