Curr Cardiovasc Imaging Rep (2016) 9:3
DOI 10.1007/s12410-015-9365-6

CrossMark

CARDIAC NUCLEAR IMAGING (A CUOCOLO, SECTION EDITOR)

Prediction Models for Cardiac Risk Classification with Nuclear Cardiology Techniques

Mario Petretta[1] · **Alberto Cuocolo**[2]

Abstract Regression modeling strategies are increasingly used for the management of subjects with cardiovascular diseases as well as for decision-making of subjects without known disease but who are at risk of disease in the short- or long-term or during life span. Accurate individual risk assessment, taking in account clinical, laboratory, and imaging data is useful for choosing among prevention strategies and/or treatments. The value of nuclear cardiology techniques for risk stratification has been well documented. Many models have been proposed and are available for diagnostic and prognostic purposes and several statistical techniques are available for risk stratification. However, current approaches for prognostic modeling are not perfect and present limitations. This review analyzes some specific aspects related to prediction model development and validation.

Keywords Cardiovascular disease · Risk stratification · Nuclear cardiology · Myocardial perfusion imaging · Algorithms for risk prediction

This article is part of the Topical Collection on *Cardiac Nuclear Imaging*

✉ Mario Petretta
petretta@unina.it

[1] Department of Translational Medical Sciences, University Federico II, Naples, Italy

[2] Department of Advanced Biomedical Sciences, University Federico II, Naples, Italy

"*Forecasting is the art of saying what will happen, and then explaining why it didn't!*"
Anonymous (communicated by Balaji Rajagopalan)

Introduction

Historically, test accuracy has been evaluated considering the sensitivity and specificity for the detection of a disease. Nevertheless, as stated by Dr. Charles Edgar Lea [1] in a review article that appeared in the British Medical Journal in 1915, "It is probably true that in no branch of medicine is an accurate prognosis more widely demanded and, withal, more difficult to offer than in diseases of the heart."

Prognostic modeling is a vivid area of research. Many articles are published monthly on prognostic factors and new models are proposed to the clinicians. This area of interest is also the topic of many recent books. Considering an individual's absolute risk in relation to the potential harm from therapy can target therapy to the subjects who most likely will benefit without unwarranted risk for harm. This approach gives also the chance for clinicians and patients in discussing preventive as well therapeutic choices when absolute risk is small. As working definitions, predictive modeling is based on the prospective (or concomitant) application of person level risk measures and statistical analytic technique to recognize subjects with high medical need who are expected to benefit from care management interventions. On the other hand, risk adjustment takes into account the health status of a population when setting budgets, capitation rates or premiums, evaluating provider performance, or outcomes of care. The aim of regression modeling strategies is to build an accurate and discriminating prediction model from several variables. In certain

Published online: 06 January 2016

settings, models may be extremely complex but, in the clinical realm, they should be practical, simple, and interpretable.

Stress Myocardial Perfusion Imaging and Risk Stratification

In the second half of the 1970s it was demonstrated that cardiovascular nuclear medicine provides insights into the functional state of the myocardium not obtainable from other procedures, and was applied in coronary care units and in exercise testing of ambulatory patients [2, 3]. Iskandrian et al [4] outlined the advantages and limits of thallium-201 (Tl-201) myocardial scintigraphy. On the other hand, Gibson et al [5•] demonstrated that resting anterior Tl-201 defect is able to recognize those patients with inferior infarction at great risk for subsequent coronary events, probably due to stenosis of the left anterior descending coronary artery. Brown et al [6••] in 1983 first demonstrated that the presence or extent of jeopardized viable myocardium, as assessed by Tl-201 imaging, is directly related to the risk of subsequent cardiac events. From this time, several studies reported the independent and incremental prognostic significance of Tl-201 scintigraphy [7–12].

In the 1990s technetium-99 m (Tc-99 m) perfusion agents were developed and introduced for assessing myocardial perfusion in the clinical setting [13]. On the other hand, risk assessments become increasingly important and there was a change toward comprehensive patient approach in which both the extent and severity of inducible ischemia are considered for therapeutic decision-making. As a consequence, at this time nuclear cardiology plays a pivotal role in patient management choices [14]. In the field of nuclear cardiology, the type of stressor, choice of radionuclide tracers, and detection technology are continuously evolving: from exercise stress alone to a variety of pharmacologic agents [15–17], from Tl-201 to Tc-99 m tracers [18, 19] and to positron emission tomography tracers [20–22], from planar to single-photon emission computed tomography imaging, from single head to dual-head cameras, and from Na-I crystals to cadmium zinc telluride solid state cameras [23]. Moreover, nuclear cardiology is in competition with other emerging noninvasive techniques (stress echocardiography, cardiac computed tomography coronary angiography, cardiac magnetic resonance) and has been also integrated in hybrid imaging techniques [24, 25]. Documented changes in the clinical manifestations and treatment of coronary artery disease are challenging the role of cardiac radionuclide imaging for prognostication. Nevertheless, recent reviews [26, 27] and meta-analysis [28•] have well described the abundance of evidence published on the prognostic accuracy of stress myocardial perfusion imaging and "the prognosis for prognosis remains excellent" [29]. This review, rather than focus on individual studies, will analyze some specific aspects related to prediction model building and validation.

Prediction Model Building

Prognostic studies search for predictors associated with specific endpoints. An objective may be to investigate the relative effects of covariates in a model. For instance, we may be concerned in the independent or the incremental prognostic significance of summed stress score for the clinical progression and consequence of known or suspected coronary artery disease. Additionally, clinical prediction models may be used to stratify groups of patients and guide treatment choices. Finally, prediction for individual patients is often preferred to risk grouping, providing absolute risk estimates for individual patients [30].

Adequate prediction requires the selection of adequate endpoints, generally involves multiple prognostic factors, and requires a correct statistical approach. The most frequently used statistical tool for the forecast of binary events, including short-term mortality, is logistic regression analysis, whereas Cox regression is commonly used for time-to-event data, such as long-standing mortality. Despite its popularity, the Cox model has some limitations, such as the restraining assumption of proportional hazards for covariate effects, and the lack of evaluation of the baseline hazard function due to conditioning on event times. The Cox model is semi-parametric, as it makes no assumption about the parametric form of the survival time distribution, since the only parameters to estimate in the model are those describing how the predictors affect the hazard. However, as the hazard function is a central aspect of the time course of the illness, it is also a limitation of the Cox model. Therefore, several parametric models have been developed, including exponential and Weibull models, Gompertz model, log-normal and log-logistic models, and generalized gamma model [31]. The Akaike information criterion [32] or Bayesian information criterion [33] may be used to choose the parametric distribution that gives the best fit with the fewest parameters. If h(t) is the hazard at time t, for a Weibull model with k covariates $h(t)=\lambda pt^{p-1}$, where $\lambda=\exp(\beta_0+\beta_1x_1+\beta_2x_2+\ldots.\beta_kx_k)$, β_0 is the constant (intercept) term of the model and β_1, β_2,... β_k the coefficient of the x_k covariates estimated through maximum likelihood, and p is the shape parameter. If $p>1$ the hazard rises with time, if $p=1$ the hazard is constant and the Weibull model shrinks to an exponential model, if $p<1$ the hazard declines. By means of direct modeling of the baseline hazard function, it is possible to better understand how the patient's risk profile changes during the course of the time, obtaining absolute measures of risk [34–37]. It should be noted that the Weibull model makes restraining suppositions of the baseline hazard function, such

as monotonicity, which is often violated in the clinical setting. Thus, more flexible parametric survival models have been developed [38, 39].

Key Steps to Construct Useful Prognostic Models

Seven key steps to construct useful prognostic models have been recommended [40•]: consideration of the research question and initial data inspection; coding of predictors; model specification; model estimation; evaluation of model performance; internal validation; and model presentation. The validity of a prediction model should be theoretically assessed in entirely independent data, utilizing measures to estimate model performance such as calibration-in-the-large (ie, the model intercept; calibration slope; discrimination, with a concordance statistic; and clinical usefulness, with decision-curve analysis) [40•]. In prognostic studies, it also is essential to assess and identify biases sufficiently large to distort study results. For this purpose, the Quality in Prognosis Studies tool recognizes 6 key areas to be considered when evaluating validity and bias in studies of prognostic variables: participation, attrition, prognostic factor measurement, confounding measurement and account, outcome measurement, and analysis and reporting [41]. The quality of the validation analysis may be positively influenced by guidelines such as the recently proposed list of 22 items considered crucial for worthy reporting of a research building or validating multivariable prediction model [42].

Endpoint Selection

An imperative initial step is to judiciously select the endpoint of interest [43]. As in other fields, also in nuclear cardiology several types of endpoints have been used. Among hard endpoints, all-cause death is the most valid and unbiased as it does not necessitate pronouncement about the reason of death. In addition, all-cause mortality is an outcome that can capture unpredicted fatal side effects of medical treatment. Nevertheless, this endpoint has limitations, in particular in older patients with high prevalence of comorbidity. In patients with coronary artery disease, death is usually assumed to be cardiac, thus cardiac death is another endpoint commonly used in this patients. However, cardiac death is susceptible to information bias (in particular misclassification bias) [44]. Within cardiovascular imaging research, the study population often comes from only one or few centers [45]. Thus, due to the small sample size and the limited number of events, the choice of endpoints tends to be a combination of lethal and morbid complications, including biased outcomes, such as unstable angina or revascularization or heart failure hospitalization [46–48]. Moreover, revascularization procedures are treated differently among imaging prognostic studies. In fact, in some studies patients who underwent revascularization in the first 60 days after the imaging study are omitted from the analysis. This choice is based on available data showing that referral to coronary revascularization early after nuclear testing tends is biased by the results of the scan while referral to revascularization >60 days after testing is more frequently due to deterioration of the patient's clinical status [49]. However, the cut-off between "early" and "late" revascularizations is subjective and not standardized. In fact, in other studies, the cut-off is 90 days [50] or only 30 days [51]. In further investigations, revascularizations are considered "soft events" and analyzed separately, not combined with hard events [52, 53]. Other prognostic studies use as endpoint a composite of major adverse cardiovascular events; however, heterogeneous among studies [54]. The use of composite endpoints increases the number of events and lowers accrual time but has numerous limitations [55], such as the assumption of risk homogeneity among the component events. Moreover, soft endpoints are more influenced by physicians' or patients' treatment preferences than harder endpoints; also reporting study results when one of the composite endpoints is the principal event occurring may be challenging [45, 56].

In most studies, hard events are precluded from consideration as the primary endpoint for analysis because of the low cumulative incidence. Researches focusing on "surrogate" endpoints have shown some benefits and revealed that results based on validated intermediate markers are often useful and should be taken more in account for decision-making guidelines [57]. Some prognostic imaging studies included in the composite endpoint components that seem to have particular biological and/or clinical meaning. As an example, the AdreView myocardial imaging for risk evaluation in heart failure [58] provides prospective validation of the prognostic value of quantitation of sympathetic cardiac innervation by means of iodine-123-meta-iodobenzylguanidine. In this report, the main outcome of interest was the relation between late heart to mediastinum ratio and the time for the manifestation of the first event among a combination of heart failure progression, life threatening arrhythmic event, and death for cardiac cause. However, this approach can distract the interpretation of results despite the formal incorporation of multiple analysis procedures, such as recurrent event [59] or competing risk [60•] analyses. In fact, the conventional analysis of composite endpoints treats all outcomes as of equal relevance, and frequently only take into account the first occurrence. With this approach, many events occurring late are missing, the follow-up period is short, and long-term questions might not be addressed [61]. In the presence of competing risks, each contending event has a related hazard function, the cause-specific hazard that measures the hazard of suffering an event from a precise cause. Linked with this cause-specific hazard is the cumulative incidence function, also known as sub-

distribution function, which quantifies the likelihood of undergoing an event on a given point in time or earlier due to a specific reason [62]. Thus, although most of the prognostic cardiovascular imaging studies utilize composite endpoints, it has been suggested that these latter should be at best considered a temporary analytic approach awaiting larger data sets [43]. Of note, myocardial perfusion imaging is increasingly used by itself as an endpoint in single- as well multi-center studies [63–67]. These studies differ in scope and complexity and may involve issues beyond reproducibility and repeatability. Thus, at this time it remains unclear if myocardial perfusion imaging provides a meaningful endpoint based upon global and regional assessments of perfusion and function [68, 69].

Standardized Definition of Cardiovascular Endpoints

The definitions of cardiovascular endpoints is fated to vary over time, as new biomarkers or other diagnostic tests become available, or as standards evolve and awareness of clinical relevance become modified. Thus, the terminology used to define specific cardiovascular endpoint is continuously evolving. As an example, consider the case of myocardial infarction. In 2000, the First Global Myocardial Infarction Task Force offered a novel definition of myocardial infarction, which states that any necrosis in the presence of myocardial ischemia should be categorized as myocardial infarction [70]. These codes were advanced by a more recent article in 2007, which highlighted the diverse situations that might lead to myocardial infarction [71]. The progress of even more sensitive assays for markers of myocardial necrosis dictates additional amendment, chiefly when such necrosis follows in the setting of the critically ill, after percutaneous coronary procedures or after cardiac surgery [72]. For example, the introduction of cardiac troponin assays radically augmented the amount of diagnosable myocardial infarction [73]. Endpoint definitions are necessary in clinical trials as well as in prognostic studies so that events are clearly characterized by objective criteria and reported uniformly. The American College of Cardiology/American Heart Association Task Force on Clinical Data Standards recently chose cardiovascular disorders and procedures that will benefit from design of a standard dataset [74••]. The aim was to recognize and harmonize the common data features involved in crucial cardiovascular endpoint classifications. In fact, progress in database expertise and statistical tools provided the opportunity to aggregate large trial datasets, but only if uniformly defined events among different studies may be easily analyzed, compared, and trends identified. Obviously, also an incorrect classification of disease code of major adverse cardiovascular events by software may impact the results observed [75].

Population-Based vs Patient-Centered Risk Prediction

Obtaining accurate risk information for a single patient from a population-based risk model is inherently uncertain. Making personalized treatment decisions based solely on risk-estimates that have limited applicability to the patient may result in poor preventive care. Thus, decisions based on risk may prove too risky. In contrast, emerging data suggest that screening for illness reduces uncertainty for the patient and may improve tailored decision-making. We must be willing to revise our models as new information emerges, to admit when our knowledge is incomplete, and to consider new strategies to optimize the care of our patients. It may be time to forgo the flip of a coin that often dictates care in preventive cardiology [76]. At the population level, it seems to be correct to recommend a treatment on the basis of some level of risk. At the individual level, such a broad range of vagueness indicates that numerous subjects whose predicted risks are near the selected cutoff will be incorrectly classified. Thus, individual-level predictions have a higher level of uncertainty than population means and may be inappropriate for guidelines and quality measures to use with particular numerical cutoffs. Multilevel models are defined by the presence of an inherent hierarchical or clustered structure. Two modeling approaches are generally used to estimate the associations between individual variables and health endings in multilevel studies. The former is the random effects or mixed model, which utilizes maximum likelihood estimation, and the latter is the population average model, which classically utilizes a generalized estimating equations. The assumptions of each method and how these assumptions affect the inferences from the analysis govern the best approach to exploring the data [77].

Novel Metrics for Risk Stratification

Several summary statistics are commonly used for quantifying the improvement in prediction performance by added covariates (Table 1). However, which measure is most appropriate is a controversial issue [78–80]. The net reclassification improvement (NRI) is widely used to evaluate the incremental value of markers added to a risk prediction model, beyond the change in the area under the receiver operating characteristics curve (AUC) [81]. Reclassification tables are commonly used to show how many subjects are reclassified by adding a marker to a model. The NRI is computed as the sum of differences in proportions of individuals in whom the estimated risk increases (“moving up”) minus the proportion in whom the estimated risk decreases (“moving down”) for those with the

Table 1 Performance of novel metrics for risk stratification

	Description	Advantages	Limitations
NRI with categories	Net proportion of patients with events reassigned to a higher risk category and of patients without events reassigned to a lower risk category	Reclassification tables constructed separately for participants with and without events. Quantification of the correct movement in categories. Related to clinical practice	Only individuals that cross a category threshold contribute to the NRI. Dependent on choice and number of categories. Ranges of meaningful improvement not established
NRI without categories	Net proportion of events with increased model-based probability plus net proportion of non-events with decreased model-based probability	Counts the direction of change for every individual rather than the crossing of a threshold. Not affected by calibration	Ranges of meaningful improvement not established
Absolute IDI	Actual change in calculated risk for each individual separately for those with and those without events	Compares models and not individual variables. Summary measure	Sensitive to differences in event rate. Ranges of meaningful improvement not established
Relative IDI	Ratio of differences between means of model-based probability for events and non-events with and without new predictor	Independent of category. The use of a relative scale may improve interpretability	Ranges of meaningful improvement not established

IDI integrated discrimination improvement, *NRI* net reclassification improvement.

outcome of interest, and the proportion of individuals moving down minus the proportion moving up for those without the outcome: P(up|diseased) − P(down|diseased) + P(down|non-diseased) − P(up|non-diseased).

The NRI was introduced with an example in cardiovascular disease prevention, where three 10-year risk categories are commonly considered (0–6 %, 6–20 %, >20 %) [82••]. For binary classification in low or high risk, NRI reduces to the sum of the improvements (delta,Δ) in sensitivity and specificity. It should be noted that using a single cut-off, the AUC is (sensitivity + specificity)/2. Thus, the NRI in the 2-category case is $\Delta_{sensitivity} + \Delta_{specificity}$, or $2 \times \Delta_{AUC}$ and also equal to Δ in Youden index, ie, Δ [(sensitivity + specificity) − 1] [83]. Categorical NRI is highly sensitive to both the number of risk categories and the thresholds between categories [84]. When no specific cut points for risk categories are established, a sensitivity analysis may be performed evaluating different sets of cut points [85]. In this setting, a category-free version, named NRI(>0) or continuous NRI, has been also recommended, in particular when the purpose is to make comparisons between different studies [86•]. The formula is not different for the category-free NRI(>0), but upward or downward movement are more simply defined, indicating any increase or decrease in probabilities of the endpoint of interest. Independently from the number of risk sets, it should be remarked that NRI is not a proportion; rather, it is a mixture of 4 proportions [87]. In survival analysis, NRI_{event} is the difference in the proportion of patients with events moving to a higher-risk category minus the proportion with events moving to a lower-risk category. The $NRI_{noevent}$ is the difference in the proportion of patients without events moving to a lower-risk category minus the proportion without events moving to a higher-risk category. If all "event" and "noevent" subjects are correctly reclassified, the NRI is 2 (the highest possible value), and thus it is better reported in absolute units than in percentage, to circumvent erroneous interpretation. As the interpretation of NRI_{event} and $NRI_{noevent}$ is not straightforward, it has been proposed to weight by the prevalence of events in alternative to take a naïve sum (or unweighted mean) to produce the NRI. This weighting expands the interpretations of NRI_{event} and $NRI_{noevent}$ to the whole population. The "population-weighted NRI" is $p\ NRI_{event} + (1-p)\ NRI_{novent}$, where p is the prevalence of the disease or outcome of interest. In this manner, the population-weighted NRI can be read as the net change in the proportion of subjects allocated to a more appropriate risk or risk category under the new model [88]. However, the NRI suffers a serious problem, being "too sensitive" even to nonexistent improvements in prediction [89]. Moreover, based on simulation studies, it appears that the NRI statistics can be misleading, and some authors recommend avoiding use of the NRI in practice [90].

A further criticism of NRI is that when there are 3 or more risk categories, this index does not effectively account for clinically relevant dissimilarities in shifts among risk groups [88]. To overcome this limit, it has been proposed an adjustment to the original definition of the NRI, which ponders each reclassification by the number of categories by which a subject is reclassified [91]. With this approach, the event and nonevent NRI may be simply considered as sums of changes in sensitivities and specificities measured at the risk thresholds.

While NRI aims to calculate the number of subjects who are correctly reclassified when an additional marker is included into a model, the integrated discrimination improvement (IDI) quantifies how far away from each other the probabilities of events and non-events become after consideration of

the additional marker [86•]. IDI quantifies the new model's improvement in sensitivity (true positive rate) without sacrificing its specificity (true negative rate). In comparing 2 models, an old model (model 1) and a new model (model 2), IDI measures the increase in the predicted probabilities for the subjects experiencing an event and the decrement for the subjects not experiencing an event, giving a measure of how far apart on average they are. IDI can also be defined as the difference in the discrimination slopes in the binary context (ie, as differences in the mean predicted probabilities of events and non-events). Two types of IDI may be calculated: absolute IDI and relative IDI. The formula to calculate the absolute IDI is $(\overline{p}\_\text{event}\_2 - \overline{p}\_\text{event}\_1) + (\overline{p}\_\text{noevent}\_1 - \overline{p}\_\text{noevent}\_2)$ where $\overline{p}$_event and $\overline{p}$_noevent are the average predicted probability of event and non-event for model 1 and model 2, in that order. The absolute value of IDI depends on incidence of the outcome of interest; thus, if the predicted event rate is small, it may be of value to calculate the relative IDI defined as the increase in discrimination slopes divided by the slope of the old model: $(\overline{p}\_\text{event}\_2 - \overline{p}\_\textit{noevent}\_2) / (\overline{p}\_\text{event}\_1 - \overline{p}\_\text{noevent}\_1)$.

Conclusions

Prognostic modeling is an intense area of research and is becoming more and more important in cardiovascular medicine. The goal of a useful prediction model is to recognize subjects at increased risk and to improve medical practice. For clinical use, it is recommendable to consider easily available predictors, to evaluate performance in a broad selection of models, and to select model minimizing generalized measure of error. The model should be internally and externally validated. Bootstrap is the preferred technique for internal validation. For external validation, predictions calculated from the previously developed model are tested in new data that are different from the development population. Growing emphasis focuses on techniques in which models can be enhanced using new markers, and there is a call for separating prediction, where established measures of performance, for instance discrimination and calibration, are helpful, from classification, where various other available statistic techniques have been proposed. Finally, methods for evaluating treatment selection markers are an active field of research.

Compliance with Ethical Standards

Conflict of Interest Mario Petretta and Alberto Cuocolo declare that they have no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:

- • Of importance
- •• Of major importance

1. Lea CE. Prognosis in heart disease. Br Med J. 1915;1:1036–8.
2. Wagner Jr HN. Cardiovascular nuclear medicine: a progress report. Hosp Pract. 1976;11:77–83.
3. Zaret BL. Myocardial imaging with radioactive potassium and its analogs. Prog Cardiovasc Dis. 1977;20:81–94.
4. Iskandrian AS, Wasserman L, Segal BL. Thallium 201 myocardial scintigraphy. Advantages and limitations. Arch Intern Med. 1980;140:320–7.
5.• Gibson RS, Taylor GJ, Watson DD, et al. Prognostic significance of resting anterior thallium-201 defects in patients with inferior myocardial infarction. J Nucl Med. 1980;21:1015–21. **This paper demonstrates that resting Tl-201 scintigraphy is a sensitive method to detect myocardial infarction and is able to identify patients at high risk for subsequent coronary events.**
6.•• Brown KA, Boucher CA, Okada RD, et al. Prognostic value of exercise thallium-201 imaging in patients presenting for evaluation of chest pain. J Am Coll Cardiol. 1983;1:994–1001. **This study suggests an approach to evaluate the risk of future cardiac events in patients with possible ischemic heart disease.**
7. Ladenheim ML, Pollock BH, Rozanski A, et al. Extent and severity of myocardial hypoperfusion as predictors of prognosis in patients with suspected coronary artery disease. J Am Coll Cardiol. 1986;7:464–71.
8. Brown KA. Prognostic value of thallium-201 myocardial perfusion imaging. A diagnostic tool comes of age. Circulation. 1991;83:363–81.
9. Pollock SG, Abbott RD, Boucher CA, et al. Independent and incremental prognostic value of tests performed in hierarchical order to evaluate patients with suspected coronary artery disease. Validation of models based on these tests. Circulation. 1992;85:237–48.
10. Petretta M, Bonaduce D, Cuocolo A, et al. Incremental prognostic value of thallium imaging and coronary angiography in patients with a symptom-limited ECG stress test. Coron Artery Dis. 1993;4:637–44.
11. Petretta M, Cuocolo A, Carpinelli A, et al. Prognostic value of myocardial hypoperfusion indexes in patients with suspected or known coronary artery disease. J Nucl Cardiol. 1994;1:325–37.
12. Petretta M, Cuocolo A, Bonaduce D, et al. Incremental prognostic value of thallium reinjection after stress-redistribution imaging in patients with previous myocardial infarction and left ventricular dysfunction. J Nucl Med. 1997;38:195–200.
13. Beller GA. New directions in myocardial perfusion imaging. Clin Cardiol. 1993;16:86–94.
14. Shaw LJ, Iskandrian AE. Prognostic value of gated myocardial perfusion SPECT. J Nucl Cardiol. 2004;11:171–85.
15. Travain MI, Wexler JP. Pharmacological stress testing. Semin Nucl Med. 1999;29:298–318.
16. Currie GM, Wheat JM, Wang L, et al. Pharmacology in nuclear cardiology. Nucl Med Commun. 2011;32:617–27.
17. Ghimire G, Hage FG, Heo J, et al. Regadenoson: a focused update. J Nucl Cardiol. 2013;20:284–8.
18. Sorrentino AR, Acampa W, Petretta M, et al. Comparison of the prognostic value of SPECT after nitrate administration and metabolic imaging by PET in patients with ischaemic left ventricular dysfunction. Eur J Nucl Med Mol Imaging. 2007;34:558–62.

19. Acampa W, Cuocolo A, Petretta M, et al. Tetrofosmin imaging in the detection of myocardial viability in patients with previous myocardial infarction: comparison with sestamibi and Tl-201 scintigraphy. J Nucl Cardiol. 2002;9:33–40.
20. Dilsizian V, Taillefer R. Journey in evolution of nuclear cardiology: will there be another quantum leap with the F-18-labeled myocardial perfusion tracers? JACC Cardiovasc Imaging. 2012;5:1269–84.
21. Murthy VL, Lee BC, Sitek A, et al. Comparison and prognostic validation of multiple methods of quantification of myocardial blood flow with 82Rb PET. J Nucl Med. 2014;55:1952–8.
22. Schindler TH, Quercioli A, Valenta I, et al. Quantitative assessment of myocardial blood flow–clinical and research applications. Semin Nucl Med. 2014;44:274–93.
23. Slomka PJ, Pan T, Berman DS, et al. Advances in SPECT and PET Hardware. Prog Cardiovasc Dis. 2015;57:566–78.
24. Gaemperli O, Kaufmann PA, Alkadhi H. Cardiac hybrid imaging. Eur J Nucl Med Mol Imaging. 2014;41 Suppl 1:S91–103.
25. Nappi C, Acampa W, Pellegrino T, et al. Beyond ultrasound: advances in multimodality cardiac imaging. Intern Emerg Med. 2015;10:9–20.
26. Bourque JM, Beller GA. Stress myocardial perfusion imaging for assessing prognosis: an update. JACC Cardiovasc Imaging. 2011;4:1305–19.
27. Shaw LJ, Hage FG, Berman DS, et al. Prognosis in the era of comparative effectiveness research: where is nuclear cardiology now and where should it be? J Nucl Cardiol. 2012;19:1026–43.
28.• Acampa W, Cantoni V, Green R, et al. Prognostic value of normal stress myocardial perfusion imaging in diabetic patients: a meta-analysis. J Nucl Cardiol. 2014;21:893–902. **A large meta-analysis showing that stress myocardial perfusion single-photon emission computed tomography has a high negative predictive value for adverse cardiac events in diabetic patients leading to define a "relatively low-risk" patients category.**
29. Hansen CL. The prognosis for prognosis remains excellent. J Nucl Cardiol. 2013;20:501–3.
30. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338:b375.
31. George B, Seals S, Aban I. Survival analysis and regression models. J Nucl Cardiol. 2014;21:686–94.
32. Akaike H. A new look at the statistical model identification. IEEE Trans Autom Control. 1974;19:716–23.
33. Schwarz G. Estimating the dimension of a model. Ann Stat. 1978;6:461–4.
34. Acampa W, Evangelista L, Petretta M, et al. Usefulness of stress cardiac single-photon emission computed tomographic imaging late after percutaneous coronary intervention for assessing cardiac events and time to such events. Am J Cardiol. 2007;100:436–41.
35. Petretta M, Acampa W, Evangelista L, et al. Impact of inducible ischemia by stress SPECT in cardiac risk assessment in diabetic patients: rationale and design of a prospective multicenter trial. J Nucl Cardiol. 2008;15:100–4.
36. Daniele S, Nappi C, Acampa W, et al. Incremental prognostic value of coronary flow reserve assessed with single-photon emission computed tomography. J Nucl Cardiol. 2011;18:612–9.
37. Acampa W, Petretta M, Cuocolo R, et al. Warranty period of normal stress myocardial perfusion imaging in diabetic patients: a propensity score analysis. J Nucl Cardiol. 2014;21:50–6.
38. Crowther MJ, Lambert PC. A general framework for parametric survival analysis. Stat Med. 2014;33:5280–97.
39. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med. 2002;21:2175–97.
40.• Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 2014;35:1925–31. **This article illustrates the usefulness of a framework to strengthen the methodological rigour and quality for prediction models in cardiovascular research.**
41. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158:280–6.
42. Collins GS, Reitsma JB, Altman DG, et al. TRIPOD Group. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. The TRIPOD Group. Circulation. 2015;131:211–9.
43. Hachamovitch R. Assessing the prognostic value of cardiovascular imaging: a statistical exercise or a guide to clinical value and application? Circulation. 2009;120:1342–4.
44. Delgado-Rodríguez M, Llorca J. Bias. J Epidemiol Community Health. 2004;58:635–41.
45. Shaw LJ, Min JK, Hachamovitch R, et al. Cardiovascular imaging research at the crossroads. JACC Cardiovasc Imaging. 2010;3:316–24.
46. Metz LD, Beattie M, Hom R, et al. The prognostic value of normal exercise myocardial perfusion imaging and exercise echocardiography: a meta-analysis. J Am Coll Cardiol. 2007;49:227–37.
47. Shaw LJ, Narula J. Cardiovascular imaging quality—more than a pretty picture! JACC Cardiovasc Imaging. 2008;1:266–9.
48. Hadamitzky M, Freissmuth B, Meyer T, et al. Prognostic value of coronary computed tomographic angiography for prediction of cardiac events in patients with suspected coronary artery disease. JACC Cardiovasc Imaging. 2009;2:404–11.
49. Hachamovitch R, Berman DS, Kiat H, et al. Exercise myocardial perfusion SPECT in patients without known coronary artery disease: incremental prognostic value and use in risk stratification. Circulation. 1996;93:905–14.
50. Kim HL, Kim YJ, Lee SP, et al. Incremental prognostic value of sequential imaging of single-photon emission computed tomography and coronary computed tomography angiography in patients with suspected coronary artery disease. Eur Heart J Cardiovasc Imaging. 2014;15:878–85.
51. Steel K, Broderick R, Gandla V, et al. Complementary prognostic values of stress myocardial perfusion and late gadolinium enhancement imaging by cardiac magnetic resonance in patients with known or suspected coronary artery disease. Circulation. 2009;120:1390–400.
52. Georgoulias P, Demakopoulos N, Tzavara C, et al. Long-term prognostic value of Tc-99m tetrofosmin myocardial gated-SPECT imaging in asymptomatic patients after percutaneous coronary intervention. Clin Nucl Med. 2008;33:743–7.
53. Galassi AR, Azzarelli S, Tomaselli A, et al. Incremental prognostic value of technetium-99m-tetrofosmin exercise myocardial perfusion imaging for predicting outcomes in patients with suspected or known coronary artery disease. Am J Cardiol. 2001;88:101–6.
54. Kip KE, Hollabaugh K, Marroquin OC, et al. The problem with composite end points in cardiovascular studies: the story of major adverse cardiac events and percutaneous coronary intervention. J Am Coll Cardiol. 2008;51:701–7.
55. Hachamovitch R, Di Carli MF. Methods and limitations of assessing new noninvasive tests: part II: outcomes-based validation and reliability assessment of noninvasive testing. Circulation. 2008;117:2793–801.
56. Goldberg R, Gore JM, Barton B, et al. Individual and composite study endpoints: separating the wheat from the chaff. Am J Med. 2014;127:379–84.
57. Asmar R, Hosseini H. Endpoints in clinical trials: does evidence only originate from 'hard' or mortality endpoints? J Hypertens Suppl. 2009;27:S45–50.
58. Jacobson AF, Senior R, Cerqueira MD, et al. ADMIRE-HF Investigators. Myocardial iodine-123 meta-iodobenzylguanidine imaging and cardiac events in heart failure. Results of the prospective ADMIRE-HF (AdreView Myocardial Imaging for Risk

Evaluation in Heart Failure) study. J Am Coll Cardiol. 2010;55: 2212–21.
59. Diamond GA, Kaul S. Forbidden fruit: on the analysis of recurrent events in randomized clinical trials. Am J Cardiol. 2013;111:1530–6.
60.• Wolbers M, Koller MT, Stel VS, et al. Competing risks analyses: objectives and approaches. Eur Heart J. 2014;35:2936–41. **This paper stresses the importance of choosing statistical methods that are appropriate if competing risks are present and also clarifies the role of competing risks for the analysis of composite endpoints.**
61. Petretta M, Pellegrino T, Cuocolo A. Cardiac neuronal imaging with ^{123}I-meta-iodobenzylguanidine in heart failure: implications of endpoint selection and quantitative analysis on clinical decisions. Eur J Nucl Med Mol Imaging. 2014;41:1663–5.
62. Aban I. Time to event analysis in the presence of competing risks. J Nucl Cardiol. 2015;22:466–7.
63. Perin EC, Willerson JT, Pepine CJ, et al. Cardiovascular Cell Therapy Research Network (CCTRN). Effect of transendocardial delivery of autologous bone marrow mononuclear cells on functional capacity, left ventricular function, and perfusion in chronic heart failure: the FOCUS-CCTRN trial. JAMA. 2012;307:1717–26.
64. Phillips LM, Hachamovitch R, Berman DS, et al. Lessons learned from MPI and physiologic testing in randomized trials of stable ischemic heart disease: COURAGE, BARI 2D, FAME, and ISCHEMIA. J Nucl Cardiol. 2013;20:969–75.
65. Sadat K, Ather S, Aljaroudi W, et al. The effect of bone marrow mononuclear stem cell therapy on left ventricular function and myocardial perfusion. J Nucl Cardiol. 2014;21:351–67.
66. Iskandrian AE, Hage FG, Shaw LJ, et al. Serial myocardial perfusion imaging: defining a significant change and targeting management decisions. JACC Cardiovasc Imaging. 2014;7:79–96.
67. El-Hajj S, AlJaroudi WA, Farag A, Bleich S, Manaoragada P, Iskandrian AE, et al. Effect of changes in perfusion defect size during serial regadenoson myocardial perfusion imaging on cardiovascular outcomes in high-risk patients. J Nucl Cardiol. 2016. doi: 10.1007/s12350-015-0174-8.
68. Sherwood M, Hage FG, Heo J, et al. SPECT myocardial perfusion imaging as an endpoint. J Nucl Cardiol. 2012;19:891–4.
69. Petretta M, Salvatore M, Cuocolo A. Immortality time and serial myocardial perfusion imaging: only those who do not die may repeat the exam. J Nucl Cardiol. 2016. doi:10.1007/s12350-015-0171-y.
70. Alpert JS, Thygesen K, Antman E, et al. Myocardial infarction redefined—a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. J Am Coll Cardiol. 2000;36: 959–69.
71. Thygesen K, Alpert JS, White HD. Joint ESC/ACCF/AHA/WHF Task Force for the Redefinition of Myocardial Infarction. Universal definition of myocardial infarction. J Am Coll Cardiol. 2007;50: 2173–95.
72. Thygesen K, Alpert JS, Jaffe AS, et al. Joint ESC/ACCF/AHA/WHF Task Force for the Universal Definition of Myocardial Infarction. Third universal definition of myocardial infarction. Circulation. 2012;126:2020–35.
73. Luepker RV, Duval S, Jacobs Jr DR, et al. The effect of changing diagnostic algorithms on acute myocardial infarction rates. Ann Epidemiol. 2011;21:824–9.
74.•• Hicks KA, Tcheng JE, Bozkurt B, et al. ACC/AHA key data elements and definitions for cardiovascular endpoint events in clinical trials: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Cardiovascular Endpoints Data Standards). Circulation. 2015;132:302–61. **This paper aims to identify and harmonize the common data elements involved in key cardiovascular endpoint definitions.**
75. Andersson C. Incorrect ICD-10 Code and MACE Endpoint. JAMA Intern Med. 2016. doi:10.1001/jamainternmed.2015.3219.
76. McEvoy JW, Diamond GA, Detrano RC, et al. Risk and the physics of clinical prediction. Am J Cardiol. 2014;113:1429–35.
77. Hubbard AE, Ahern J, Fleischer NL, et al. To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. Epidemiology. 2010;21:467–74.
78. Gu W, Pepe M. Measures to summarize and compare the predictive capacity of markers. Int J Biostat. 2009;5:27.
79. Gerds TA, Cai T, Schumacher M. The performance of risk prediction models. Biom J. 2008;50:457–79.
80. Pepe MS, Kerr KF, Longton G, et al. Testing for improvement in prediction model performance. Stat Med. 2013;32: 1467–82.
81. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21:128–38.
82.•• Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, et al. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27:157–72. **The authors discuss the properties of new predictive measures and develop simple asymptotic tests of significance.**
83. Steyerberg EW, Van Calster B, Pencina MJ. Performance measures for prediction models and markers: evaluation of predictions and classifications. Rev Esp Cardiol. 2011;64:788–94.
84. Pickering JW, Endre ZH. New metrics for assessing diagnostic potential of candidate biomarkers. Clin J Am Soc Nephrol. 2012;7:1355–64.
85. Shaw LJ, Giambrone AE, Blaha MJ, et al. Long-term prognosis after coronary artery calcification testing in asymptomatic patients: a cohort study. Ann Intern Med. 2015;163:14–21.
86.• Pencina MJ, D'Agostino Sr RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med. 2011;30:11–21. **In this paper the authors develop a general form for the net reclassification improvement that presents it as a prospective measure, which quantifies the correctness of upward and downward reclassification or movement of predicted probabilities as a result of adding a new marker.**
87. Petretta M, Cuocolo A. Prognosis in the era of comparative effectiveness research. J Nucl Cardiol. 2013;20:313.
88. Kerr KF, Wang Z, Janes H, et al. Net reclassification indices for evaluating risk prediction instruments: a critical review. Epidemiology. 2014;25:114–21.
89. Hilden J, Gerds TA. A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. Stat Med. 2014;33:3405–14.
90. Pepe MS, Janes H, Li CI. Net risk reclassification p values: valid or misleading? J Natl Cancer Inst. 2014;106:dju041.
91. Pencina KM, Pencina MJ, D'Agostino Sr RB. What to expect from net reclassification improvement with three categories. Stat Med. 2014;33:4975–87.