increase, and with it so will the need for more creative and specific strategies.

Riyad Karmy-Jones, MD, FCCP
Gabriel Aldea, MD, FCCP
Edward M. Boyle Jr., MD
Seattle, WA

Drs. Karmy-Jones, Aldea, and Boyle are from the Division of Cardiothoracic Surgery, University of Washington.

Correspondence to: Riyad Karmy-Jones, MD, FCCP, Department of Surgery, Harborview Medical Center, Box 359796, 325 Ninth Ave, Seattle, WA 98104; e-mail: riyad.karmy-jones@sos.washington.edu

REFERENCES

Killip and Forrester Classifications

Should They Be Abandoned, Kept, Reevaluated, or Modified?

To question the past is human; to do so is even imperative in the domain of science. Classifications have a certain “life span”: they come and go as their utility is surpassed by increasing scientific insight and experience. New classifications replace the old ones, just declared outmoded. At best, they can be retained in a modified form.

The science of classification was invented by Linnaeus, the Swedish naturalist-physician. His taxonomy has been at the heart of all attempts to bring some order to science in general, and the biological realm in particular, by emphasizing similarities in members of a class, while ignoring their individual attributes. Classification is a prerequisite for language and abstract thought. Scholars maintain that classifying in “pairs of opposites” is inherent to our brain’s basic function. However, implicit to classification is its artificiality and arbitrariness. There is nothing inherent to a body of information that calls for a single and unique way of classification. An infinite number of classifications can be imposed onto a set of data. Thus, it is a corollary of the above that results of various “outcomes” from the analysis of a given database will be different, depending on the classification system implemented.

The medical literature is replete of classifications, scores, types, screens, instruments, stratifications, predictors, assessment equations, determinants, methods, measures, categories, indicators, tools, markers, indices, systems, profiles, and prognosticators, used for diagnosis, risk assessment, and management. In this age of the computer, there has been an acceleration of this trend, as everyone tries to graduate from qualitative to semiquantitative, and even to quantitative assessment of problems. This treatment of data is amenable to statistical analysis, and this is not a small matter. Although many of these classification tools have remained unaltered, others have undergone modification; for example, we have experienced APACHE III (acute physiology and chronic health evaluation III), through APACHE I and APACHE II.

In the field of myocardial infarction (MI), different scores and classifications have been used for many years as prognostic instruments of short- or long-term outcome. The end points have evolved to become uniform and include, with rare variations in the theme, mortality (overall, cardiovascular, arrhythmic, or nonarrhythmic), recurrence of MI, angina or hospitalization, and need for revascularization procedures, used either singly or in various composite forms. On the other hand, the classifying variables have comprised a large array of clinically derived signs, comorbidity, risk factors, or information from different noninvasive and invasive testing. There are large differences among classification systems in the time of their design, the number, kind, and clarity of definition of the classifiers used, and the inclusion criteria of the study cohorts on which they were based. Some of these predictive tools were based on retrospective analysis, whereas others employed prospective analysis of data. The clinical parameters comprising these indices were either used to determine “membership”
in a class, or numerical weightings or “points” (reflecting gradations of risk from a particular parameter, ie, age, or location, or extent of MI) were applied to them prior to their employment in the predictive instrument. Both of these two approaches were based on clinically derived information (physical examination and chest radiography), with the Killip classification representing the most popular example. A third approach, primarily based on invasive hemodynamic assessment, has employed the cardiac index (CI) and the pulmonary capillary wedge pressure for categorization of patients with a MI, and is known as the Forrester classification. This system also has enjoyed wide application, and it is currently engulfed in the controversy whether right heart catheterization (RHC) provides any benefit in the management of patients with critical illness, and (what is even more disquieting) whether it is downright harmful. In addition to the hemodynamic information, RHC has led to the employment of mixed venous oxygen saturation as a parameter used in the follow-up of critically ill patients. This variable, along with the oxygen extraction ratio, reflects oxygen supply to and uptake by the body. Also, information on the saturation of arterial and mixed venous blood, coupled with data on cardiac output, has led to a host of parameters (oxygen transport, oxygen consumption), extending the hemodynamic to an “oxymhemodynamic” assessment of the critically ill patient.

The Killip classification was based on the evaluation of 250 patients admitted with an MI in a coronary care unit. The “pigeonholing” of patients to classes I to IV (no heart failure, some evidence of such a complication, pulmonary edema, and cardiogenic shock, respectively) was based on routine physical examination-derived parameters. The proportion of patients in class IV was 19%, higher than the rates encountered today. This system carries the imprimatur of simplicity, and intuitively makes sense, which explains its popularity and relative longevity. However, there is some fuzziness in the definition of classes II and III, resulting in ambiguity in their demarcation. Also, patients with or without lung congestion can be placed in class IV if they are in cardiogenic shock; however, one wonders whether these two subgroups within class IV have the same prognosis.

The Forrester classification assigns patients to four categories, and it was based on the clinical and hemodynamic evaluation of 200 patients. The demarcation of the classes here is clear. Thus, patients in class I have no pulmonary congestion (PC) or peripheral hypoperfusion (PH), whereas those in class IV have both. Class II patients have PC but no PH. These three classes correspond to classes I, IV, and II of the Killip system. Patients in Forrester class III, on the other hand, have PH but no PC. Thus, Forrester class III does not correspond to Killip class III. Assignment of patients to these four classes was based on RHC-derived cut-off values of CI < 2.2 L/min/m², reflecting PH, and pulmonary wedge pressure ≥ 18 mm Hg, defining PC. Forrester et al found that the classification of their patients according to the presence of PC and PH, when carried out using physical examination, differed from the one based on the information obtained from RHC in 32% of their cohort; also of interest is that reclassification of their patients after management again showed that 30% of them were not assigned to the same class by both the clinical and hemodynamic assessment.

In contemplating the Killip and Forrester classifications, one cannot help but detect an aberration based on the nature of Forrester’s class III. The classes from I to IV in both systems represent subsets of increasing severity by design, and this is reflected in the correspondingly increasing mortality. However, patients in Forrester class III with isolated PH (entirely different from Killip class III) often improve after careful volume loading to find themselves in Forrester class I. This is clearly exemplified by the patient with inferior and right ventricular MI who is admitted with the provisional diagnosis of “cardiogenic shock” and shows a gratifying clinical response after administration of a large fluid load over a relatively short time period.

Comparisons of various scoring systems have been carried out in the past. A recent trend has emerged: the incorporation of the Killip classification in a larger scoring system, which includes other predictive variables, or of the reassessment of the utility of Killip classification by applying it to current databases. The study by Siniorakis et al in this issue of CHEST (see page 1286) belongs to this genre. These authors evaluated their nonthrombolysed patients with acute MI using the Killip classification; also, hemodynamic characterization of one third of their patients according to the Forrester classification was employed, along with use of parameters reflecting body oxygen extraction. The particular parameter they focused on was the “normalized” cardiac index (NCI), derived by dividing the CI by the oxygen extraction ratio. They found a significant decline in mortality in both their entire cohort and the subgroup that underwent RHC, in comparison with what was expected 25 to 30 years ago when the two classifications were proposed. NCI in their study had the best sensitivity/specificity/predictive value from all other predictors of early mortality. These
workers did not detect any complications attributable to the RHC and concluded that NCI outperformed the two old classifications.

The current hotly contested issue of the usefulness or even harmfulness of RHC is not the main focus of this commentary. Suffice to say that this procedure has come under attack by some, while others maintain that increased mortality noted with its use can be traced to the severity of the illness of the patients who have undergone the procedure. Although some have called for a moratorium in the use of RHC, or have recommended that a randomized study be undertaken, neither of these are likely to happen. Meanwhile, we all should use some restraint in using RHC, and resort to its employment after a trial of diuresis, nitrates, or “fluid challenge” have taken place. The performance of the procedure by less-experienced physicians should be carried out under supervision, and the catheters should be removed as soon as they are deemed not to be essential for the management of the patients. The indications for RHC for patients with an acute MI are spelled out in the recently released update of guidelines of the American College of Cardiology/American Heart Association.

If one considers the major progress in diagnostics and therapeutics realized in the past 3 decades, it should not be surprising that mortality has declined so dramatically. Consequently, when patients with an MI are evaluated presently by the Killip and Forrester classifications, observed mortality rates are expected to be much lower than those originally reported. This improvement applies to patients who have not received thrombolysis or revascularization procedures (as observed by Siniorakis et al), but it is even more pronounced in patients with such therapies. This is not a limitation of the classifications, but the result of the medical progress made in the interim. Thus, the issue at stake here is whether the two classifications applied today predict mortality and other outcomes with lower accuracy than the more-recently designed prognostic instruments.

In tackling this matter, one should make a differentiation between the two classifications, since the one (Killip) is designed for general “consumption” and the other (Forrester) by necessity depends on RHC. Accordingly, and although it is important that both be reappraised, it is the Killip classification that, if modified, will have the greater impact, since it is available at the bedside to all physicians, without the need for RHC. In contrasting the two classifications, it should also be pointed out that in the study of Siniorakis et al (Table 1), the actual mortality in the four Killip classes (as applied by the authors) was much lower than that reported 30 years ago, whereas the same comparison using Forrester classification revealed numerically almost the same actual and predicted mortality for classes I, II, and IV (Table 2). Patients in Forrester class I are always (then and now) expected to do well, and patients in class IV are destined to do poorly. Perhaps early revascularization in such patients < 75 years old may change their plight. Also, it is long held that patients in class II respond to nitrates and diuresis. On the other hand, the discrepancy noted for predicted (23%) and actual (12%) mortality for the Forrester class III patients could be due to the more voluminous fluid administration used nowadays, particularly in patients with right ventricular MI, present or suspected. Accordingly, it would have been revealing to contrast the proportion of patients with inferior MI in the cohorts of Forrester et al and Siniorakis et al. The above comparisons indicate that the Forrester classification, although less accessible and more technically demanding, is more accurate and has shown greater stability over time than the Killip classification.

In general, all predicting tools are employed on admission of the patients to the hospital, and thus they may not reflect outcomes due to subsequent complications (extension of MI, mitral regurgitation due to evolving papillary muscle dysfunction, or rupture, ventricular septal defect, systemic or pulmonary embolism, cardiac rupture, atioventricular blocks, or arrhythmias). As a corollary of the above, this author has used “peak Killip class” (reached by patients during their entire hospitalization) as a separate prognostic variable, in addition to the Killip class on admission. However, this approach has its drawbacks, since the peak Killip class may be reached late in the clinical course, and thus it may have marginal utility as a prognosticator. Accordingly, a compromise may be considered here: a reclassification of patients with an MI, a few hours following their initial on admission assessment, and after they have received therapy (fluids, nitrates, diuresis, or inotropic drugs). Thus, a modification in the application of both Killip and Forrester classifications may render them more robust in predicting outcome, on a par with other more complex and time- and labor-demanding prognostic tools. Nevertheless, in this last quoted study, the Killip classification was the third most powerful predictor of 30-day mortality, after age and low systolic BP; also, the composite variable of age/Killip class revealed a highly significant relation to mortality.

Thus, the Killip and Forrester classifications should not be abandoned, but kept. Their value can be enhanced and their endurance secured by being reevaluated in the current medical milieu. Also,
some modification of their application will make them go a longer way. This is a challenge to all of us.

John E. Madias, MD
Elmhurst, NY

References

Wet Nebulization in Acute Asthma

The Last Refrain?

“Music, when soft voices die, vibrates in the memory”

Percy Bysshe Shelley

For 6 years, my office was located next to our clinic treatment rooms. Almost every day from my position on that hallway, I listened to the high-pitched descant song of the nebulizer performing the ballad of asthma treatment. The song is a familiar one to those working in emergency departments, pediatric wards, and ICUs, and it resonates through all my memories of caring for children with asthma. In coming years, however, it is a song we will hear less often. And quite possibly, it is a melody our academic grandchildren will not know at all.

In this issue of CHEST (see page 1309), Avigdor Mandelberg and his group from Tel Aviv report on a study comparing nebulizers with metered-dose inhaler (MDI) and spacer for the treatment of acute wheezing in infants and children. It is a well-designed and executed randomized, double-blind, placebo controlled trial comparing 2.5 mg of nebulized salbutamol with four puffs (400 μg) of salbutamol MDI delivered in a metal, nonelectrostatic spacer with face mask. The study had sufficient power to detect a 5% change in clinical scores between the two groups, and no significant difference between

Downloaded From: http://journal.publications.chestnet.org/pdaccess.ashx?url=/data/journals/chest/21945/ on 06/27/2017