Cardiac neuronal imaging with ^{123}I -meta-iodobenzylguanidine in heart failure: implications of endpoint selection and quantitative analysis on clinical decisions

Mario Petretta · Teresa Pellegrino · Alberto Cuocolo

Received: 22 May 2014 / Accepted: 26 May 2014 / Published online: 11 June 2014
© Springer-Verlag Berlin Heidelberg 2014

There are a number of radiopharmaceuticals that can be used to investigate autonomic neuronal functions [1]. Among these, the norepinephrine analogue meta-iodobenzylguanidine (MIBG) labelled with ^{123}I has been widely used and validated as a marker of adrenergic neuron function [2–4]. The first study addressing the prognostic value of ^{123}I -MIBG imaging in heart failure (HF) was that of Merlet et al. [5] in 90 patients suffering from either ischaemic or idiopathic cardiomyopathy. After publication of this study, more recent studies have indicated that patients with HF and decreased late heart-to-mediastinum (H/M) ratio or increased myocardial MIBG washout have a worse prognosis than those with normal quantitative myocardial MIBG parameters [6]. However, MIBG scintigraphy has still to reach widespread clinical application mainly because of the value of other cheaper variables such as left ventricular (LV) ejection fraction and brain natriuretic peptide (BNP) plasma levels. The possibility that the detection of mechanical dyssynchrony by innervation imaging might identify patients who would benefit from resynchronization pacing is another area of research interest [7].

In 2010, the landmark AdreView Myocardial Imaging for Risk Evaluation in Heart Failure (ADMIRE-HF) study was published [8]. This trial consisted of two identical open-label phase III studies enrolling patients in 96 sites in North

America and Europe to provide prospective validation of the prognostic role of quantitation of sympathetic cardiac innervation using MIBG. The primary endpoint was the relationship between late H/M ratio and time-to-occurrence of the first event among a combination of HF progression, potentially life-threatening arrhythmic event, and cardiac death. The authors found that a H/M ratio <1.6 provided prognostic information beyond LV ejection fraction, BNP, and New York Heart Association (NYHA) functional class at the time of enrolment. In a recent article in this journal, Parker et al. [9] present the results of a secondary analysis of the ADMIRE-HF study exploring the association of abnormal MIBG imaging and hospitalization events. The results of this study indicate that the H/M ratio may risk-stratify HF patients for cardiac-related hospitalization, especially when used in conjunction with BNP.

Endpoint selection

It must be emphasized that the selection of the best response variable for assessment of the efficacy of a treatment in HF patients is still under debate [10, 11]. Also for HF prognostication, the endpoint selection may influence the results. Much of the burden of HF occurs around acute care for decompensation. HF is the leading cause of hospitalization in the US and Europe, resulting in over 1 million admissions as a primary diagnosis and representing 1 to 2 % of all hospitalizations; moreover, in the last decade, the early postdischarge mortality and readmission rates have remained largely unchanged and may even be worsening [12]. There are now over 1 million HF hospitalizations annually in the US alone and approximately 70 % of total direct costs are attributable to inpatient care [13]. Also if HF hospitalization is not to be considered a “surrogate endpoint” and certainly represents a “hard” event, it is associated with its own limitations [14]. In fact, hospitalization can

M. Petretta
Department of Translational Medicine, University Federico II,
Naples, Italy

T. Pellegrino
Institute of Biostructure and Bioimaging, National Council of
Research, Naples, Italy

A. Cuocolo (✉)
Department of Advanced Biomedical Sciences, University Federico
II, Naples, Italy
e-mail: cuocolo@unina.it

be distinguished in a variety of modes. The first hospitalization has been used in some studies as a single time-to-event measure. For HF patients who have already been hospitalized, readmission may be the endpoint of interest. Inconsistently, the endpoint of rehospitalization is related to the duration of stay of the index hospitalization, because patients with prolonged index hospitalizations have less time at risk of readmission from the time of the index admission. More inclusive measures of hospitalization may include total number of admissions over time or total days of hospitalization. The matter of cause-specific vs. any hospitalization also interacts with outcome. Regardless of the particular hospitalization endpoint selected, hospitalization should be considered with the perspective of overall mortality, since patients who do not survive are not at risk of hospitalization [15]. There are also other factors that may have an impact on hospitalization endpoints, such as social and regional differences. For example, duration of hospitalization for HF in European countries is about double that in the US [16]. Comorbidities often associated with HF such as chronic obstructive pulmonary disease, atrial fibrillation and renal dysfunction often make it difficult to assess whether an acute event requiring hospitalization is due to worsening HF or worsening of the comorbidity.

Despite the above limitation, it must be outlined that in most HF studies, the conventional analysis of composite endpoints treats all outcomes as of equal relevance, and only takes into account the first outcome. With this approach, many events occurring late are missing, the follow-up period is short and long-term questions might not be addressed. As a consequence, the true impact of hospital admissions due to worsening HF either for the individual or for the healthcare system is ignored [17]. Thus, these recurrent, nonfatal events are important to quantify, although it is uncertain how this should be done statistically [18]. Interestingly, in this retrospective substudy of prospective data, Parker et al. [9] explored the relationship between MIBG imaging findings and hospitalization and give potentially useful, also if not definitive, information. In fact the reported results can be confounded by the competing risk of death. Moreover, the conditional, sequential multiple-failure Cox proportional hazards models used by Parker et al. [9] is a generalization of the Cox proportional hazards model which considers each repeat event (hospitalization) as a separate term in the partial likelihood, assuming that the events are independent when in fact they may not be. Such assumptions can be relaxed by introducing into the model the number of prior recurrences as a time-dependent covariate, which may capture the dependence among the recurrence times. It must be also noted that, for post-hoc exploratory analyses, LV ejection fraction was dichotomized as severely depressed if $<25\%$. It appears questionable if HF patients with LV ejection fraction values below this threshold need further (invasive) prognostic stratification.

Quantitative analysis

A further aspect is that Parker et al. [9] report the “independent” prognostic value of low H/M ratio as a dichotomous variable (<1.6) that was associated with increased risk of hospitalization in the final multivariable model including elevated BNP and logarithm of time since HF diagnosis. Nakata et al. [19] recently performed a patient-level analysis of six prospective multicentre cohort studies of MIBG imaging of sympathetic innervation for assessment of long-term prognosis in HF. Multivariate Cox proportional hazards model analysis for all-cause mortality identified age, NYHA functional class, late H/M ratio and LV ejection fraction as significant independent predictors. Receiver operating characteristic analysis identified 1.68 as the optimal late H/M ratio threshold for dichotomizing the population into higher-risk and lower-risk patients for a lethal outcome. The late H/M ratio threshold <1.68 identified patients at significantly increased risk in any LV ejection fraction category. Survival rates decreased progressively with decreasing late H/M ratio, with 5-year all-cause mortality rates $>7\%$ annually for late H/M ratio <1.25 , and $<2\%$ annually for late H/M ratio ≥ 1.95 . Likewise, late H/M ratio differentiated the high-risk from the low-risk patients both for sudden cardiac death and for pump failure death at 5 years. Thus, the results of this pooled analysis of 1,328 HF patients from six Japanese medical centres suggest a threshold higher than that in the ADMIRE-HF study [8].

The study of Nakata et al. [19] also demonstrates that, although the late H/M ratio may have a threshold for dichotomizing patients into higher-risk and lower-risk categories for a fatal outcome, the patient survival rate decreases linearly depending on impairment in cardiac MIBG activity. More recently, we evaluated the classification certainty of observer reproducibility of planar late H/M ratio in patients with HF [20]. We found that within the range of H/M values from 1.54 to 1.66 the agreement between paired intraobserver measurements falls below 80 %, reaching a nadir of approximately 50 % around 1.60, the proposed clinical cut-off value. Therefore, when a single late H/M ratio measurement falls between 1.54 and 1.66, there is a substantial chance that the dichotomous classification will change if the measurement is repeated by the same observer. This area of uncertainty creates a “measurement gray zone” that for interobserver variability is somewhat larger (1.52 to 1.68). These results might have important implications both for interpretation of available MIBG studies and for application in patient’s clinical decision making. Our analysis outlines the potential limitations of a dichotomous interpretation of late H/M ratio results. In fact, similarly to all other measurements in medicine, late H/M ratio does not carry strict dichotomous implications as to which prognosis or treatment is best, and this is especially true close to the cut-off value. It must be outlined that for observer reproducibility we can consider the observations for

the same subject as a series of measurements of a quantity that does not vary over the period of observation. On the other hand, when the quantity being measured is physiologically unstable or may vary under different conditions, as in a test–retest study, the classification certainty could be lower and the probability of misclassification higher than we found in our observer reproducibility analysis of late H/M ratio.

Unfortunately, Parker et al. [9] did not evaluate the “incremental” prognostic value of H/M ratio over other available information. The performance of prediction models can be assessed using a variety of different methods and metrics, such as the concordance statistic for discriminative ability, and goodness-of-fit statistics for calibration. Moreover, several new measures have recently been proposed that can be seen as refinements of discrimination measures, including variants of the *c* statistic for survival, reclassification tables, net reclassification improvement, and integrated discrimination improvement [21]. The *c*-index, a measure of model discrimination that varies from 0.5 (no better than a coin flip) to 1.0 (perfect discrimination), is rarely above 0.8 for published HF survival models. It may be of interest to know if a model including MIBG information would perform substantially better than this value. All these challenges, included those noted by Parker et al. [9], have slowed the wider clinical use of cardiac MIBG imaging, and the great potential of adrenergic system imaging still needs to be investigated in larger prospective studies before cardiac ¹²³I-MIBG imaging can become a valuable addition to clinical guidelines.

References

1. Tamaki N, Kuge Y, Yoshinaga K. Molecular imaging in heart failure patients. *Clin Transl Imaging*. 2013;1:341–51.
2. Pellegrino T, Petretta M, De Luca S, Paolillo S, Boemio A, Carotenuto R, et al. Observer reproducibility of results from a low-dose 123I-metaiodobenzylguanidine cardiac imaging protocol in patients with heart failure. *Eur J Nucl Med Mol Imaging*. 2013;40:1549–57.
3. Paolillo S, Rengo G, Pagano G, Pellegrino T, Savarese G, Femminella GD, et al. Impact of diabetes on cardiac sympathetic innervation in patients with heart failure: a 123I metaiodobenzylguanidine (123I MIBG) scintigraphic study. *Diabetes Care*. 2013;36:2395–401.
4. De Rosa A, Pappatà S, Pellegrino T, De Leva MF, Maddaluno G, Fiumara G, et al. Reduced cardiac 123I-metaiodobenzylguanidine uptake in patients with spinocerebellar ataxia type 2: a comparative study with Parkinson's disease. *Eur J Nucl Med Mol Imaging*. 2013;40:1914–21.
5. Merlet P, Valette H, Dubois-Randé JL, Moysé D, Duboc D, Dove P, et al. Prognostic value of cardiac metaiodobenzylguanidine imaging in patients with heart failure. *J Nucl Med*. 1992;33:471–7.
6. Verberne HJ, Brewster LM, Somsen GA, van Eck-Smit BL. Prognostic value of myocardial 123I-metaiodobenzylguanidine (MIBG) parameters in patients with heart failure: a systematic review. *Eur Heart J*. 2008;29:1147–59.
7. Petretta M, Petretta A, Pellegrino T, Nappi C, Cantoni V, Cuocolo A. Role of nuclear cardiology for guiding device therapy in patients with heart failure. *World J Meta-Anal*. 2014;2:1–16.
8. Jacobson AF, Senior R, Cerqueira MD, Wong ND, Thomas GS, Lopez VA, et al. 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.
9. Parker MW, Sood N, Ahlberg AW, Jacobson AF, Heller GV, Lundbye JB. Relationship of quantitative cardiac neuronal imaging with iodine-123 meta-iodobenzylguanidine and hospitalization in patients with heart failure. *Eur J Nucl Med Mol Imaging*. 2014. doi:10.1007/s00259-014-2819-2.
10. Parker M. How should we judge the efficacy of drug therapy in patients with chronic congestive heart failure? The insights of six blind men. *J Am Coll Cardiol*. 1987;9:433–8.
11. Neaton JD, Gray G, Zuckerman BD, Konstam MA. Key issues in end point selection for heart failure trials: composite end points. *J Card Fail*. 2005;11:567–75.
12. Ambrosy AP, Fonarow GC, Butler J, Chioncel O, Greene SJ, Vaduganathan M, et al. The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries. *J Am Coll Cardiol*. 2014;63:1123–33.
13. Liao L, Allen LA, Whellan DJ. Economic burden of heart failure in the elderly. *Pharmacoeconomics*. 2008;26:447–62.
14. Zannad F, Stough WG, Pitt B, Cleland JG, Adams KF, Geller NL, et al. Heart failure as an endpoint in heart failure and non-heart failure cardiovascular clinical trials: the need for a consensus definition. *Eur Heart J*. 2008;29:413–21.
15. Allen LA, Spertus JA. End points for comparative effectiveness research in heart failure. *Heart Fail Clin*. 2013;9:15–28.
16. Blair JE, Zannad F, Konstam MA, Cook T, Traver B, Burnett Jr JC, et al. Continental differences in clinical characteristics, management, and outcomes in patients hospitalized with worsening heart failure results from the EVEREST (Efficacy of Vasopressin Antagonism in Heart Failure: Outcome Study with Tolvaptan) program. *J Am Coll Cardiol*. 2008;52:1640–8.
17. Cook C, Cole G, Asaria P, Jabbour R, Francis DP. The annual global economic burden of heart failure. *Int J Cardiol*. 2014;171:368–76.
18. Rogers JK, Pocock SJ, McMurray JJ, Granger CB, Michelson EL, Östergren J, et al. Analysing recurrent hospitalizations in heart failure: a review of statistical methodology, with application to CHARM-Preserved. *Eur J Heart Fail*. 2014;16:33–40.
19. Nakata T, Nakajima K, Yamashina S, Yamada T, Momose M, Kasama S, et al. A pooled analysis of multicenter cohort studies of (123)I-MIBG imaging of sympathetic innervation for assessment of long-term prognosis in heart failure. *JACC Cardiovasc Imaging*. 2013;6:772–84.
20. Petretta M, Pellegrino T, Cuocolo A. The “gray zone” for the heart to mediastinum MIBG uptake ratio. *J Nucl Cardiol*. 2014. doi:10.1007/s12350-014-9894-4.
21. Petretta M, Cuocolo A. Prediction models for risk classification in cardiovascular disease. *Eur J Nucl Med Mol Imaging*. 2012;39:1959–69.