ESC European Society of Cardiology

European Journal of Heart Failure (2020) **22**, 1598–1599
doi:10.1002/ejhf.1977

EDITORIAL COMMENT

Searching for diagnostic biomarkers of heart failure with preserved ejection fraction: methodological issues

Alberto Aimo[1,2]*, Georgios Georgiopoulos[3], Michele Senni[4], and Michele Emdin[1,5]

[1]Institute of Life Sciences, Scuola Superiore Sant'Anna, Pisa, Italy; [2]Cardiology Division, University Hospital of Pisa, Pisa, Italy; [3]School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK; [4]Cardiovascular Department and Cardiology Unit, ASST Papa Giovanni XXIII, Bergamo, Italy; and [5]Cardiology Division, Fondazione Toscana 'Gabriele Monasterio', Pisa, Italy

This article refers to 'Risk of bias in studies investigating novel diagnostic biomarkers for heart failure with preserved ejection fraction. A systematic review' by M.T.H.M. Henkens *et al.*, published in this issue on pages 1586–1597.

Over the past three decades, the prevalence of heart failure (HF) with preserved ejection fraction (HFpEF) has risen from 41% to 56% of all cases of HF, while the prevalence of HF with reduced ejection fraction (HFrEF) and HF with mid-range ejection fraction has fallen from 44% to 31% and from 15% to 13%, respectively.[1,2] Since the burden of HFpEF is growing, defining a standardized diagnostic approach to this condition becomes increasingly important.[3] Overt congestion in hospitalized patients can usually be readily detected from physical examination, chest X-ray, and measurement of B-type natriuretic peptides (NPs), whereas the diagnosis of HFpEF may be challenging in outpatients complaining of dyspnoea on effort, and relies largely on demonstration of elevated pulmonary pressures.[3,4] Invasive haemodynamic exercise testing (right heart catheterization with measurement of pulmonary capillary wedge pressure at rest or during exercise) has emerged as the gold standard to diagnose or exclude HFpEF in patients with exertional dyspnoea of unclear aetiology, but cost, risk, and the requirement for specialized training and equipment may limit its broad application in practice and in clinical trials, while exercise echocardiography cannot be proposed as a stand-alone diagnostic examination for HFpEF.[4] The search for non-invasive alternatives for the diagnosis of HFpEF has led to the introduction of the stepwise diagnostic algorithm by the European Society of Cardiology Heart Failure Association (HFA-PEFF),[5] and the H_2FPEF score derived from dichotomized variables or the HFpEF nomogram derived from continuous variables.[6] These systems are able to discriminate non-cardiac dyspnoea from HFpEF with a high diagnostic accuracy.[5–7] Briefly, the HFA-PEFF score adopts a stepwise approach that starts by establishing the pre-test likelihood of HFpEF through the assessment of risk factors and exercise intolerance. The score then incorporates three domains (functional, morphological, and biomarkers) to estimate the likelihood of HFpEF. A high-likelihood score is considered diagnostic for HFpEF, while a low-likelihood score allows to rule out HFpEF. For patients with an intermediate score, further evaluation by means of exercise echocardiography or invasive measurement of cardiac filling pressures is advised, together with additional diagnostic test to evaluate specific causes when appropriate.[5] A high HFA-PEFF score allows to diagnose HFpEF with 93% specificity, and a low HFA-PEFF score to rule out HFpEF with 99% sensitivity. Moreover, a similar pattern of HFA-PEFF score was found in two independent cohorts despite different patient characteristics, diagnosing >60% of HFpEF patients in the high-likelihood category, although a rather large group of patients with an intermediate likelihood requiring additional testing remained.[7] The H_2FPEF score includes obesity, atrial fibrillation, age > 60 years, treatment with ≥2 antihypertensives, E/e′ ratio > 9, and pulmonary artery systolic pressure > 35 mmHg, and ranges from 0 to 9. The odds of having HFpEF increased by a factor of two for every one-unit increase in the score, and the score allowed good discrimination of HFpEF from controls.[6] The same Authors proposed also the HFpEF nomogram derived from the same items, reported as continuous variables.[6]

Beyond NPs, several biomarkers have been evaluated as possible tools to diagnose HFpEF. Among them, there are several molecules reflecting the processes of inflammation and fibrosis (most notably galectin-3 and soluble suppression of tumorigenesis-2), extracellular matrix remodelling (such as matrix metalloproteinases 2 and 9, carboxy-terminal telopeptide of collagen type I and aminoterminal propeptide of type III procollagen), elevated ventricular wall

The opinions expressed in this article are not necessarily those of the Editors of the *European Journal of Heart Failure* or of the European Society of Cardiology. doi: 10.1002/ejhf.1944

*Corresponding author. Cardiology Division, University Hospital of Pisa, and Institute of Life Sciences, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, 56124 Pisa, Italy. Tel: +39 347 7084391, Email: a.aimo@santannapisa.it, albertoaimo@libero.it

tension (adrenomedullin), and a vast array of other biomarkers including growth differentiation factor-15, cystatin C, resistin, cancer antigen-125 and von Willebrand factor.[8] Many studies have been conducted with the challenging goals of establishing the diagnostic performance of these biomarkers. In this issue of the Journal, this body of literature is critically reappraised, stressing methodological issues that could affect the reliability of their conclusions.

Henkens *et al.*[9] performed a systematic review of studies evaluating the risk of bias (ROB) in 28 studies assessing the performance of circulating biomarkers for the diagnosis of HFpEF in the non-acute setting. The ROB was evaluated across the four domains of a dedicated tool for the quality assessment of diagnostic accuracy studies (the QUADAS-2 tool): patient selection, index test, reference standard, and flow and timing. The Authors report that all studies presented at least one domain with a high ROB, and 39% of studies had a high ROB within all four domains. The most common issues were the use of a case-control or two-gated design, the exclusion of difficult-to-diagnose patients, the absence of a pre-specified cut-off value for the index test with the lack of external validation, the use of inappropriate reference standards, and the unclear timing of the index test and/or reference standards. Because of these methodological issues, and, even more importantly, of the high degree of heterogeneity across trials, a comprehensive assessment of trial results was not performed.[9]

This article has been authored by leading experts in HFpEF, who tried to clarify the problems of studies investigating novel diagnostic biomarkers. It is interesting to notice that all studies had an intermediate or high ROB regarding the reference standard, given that HFpEF was diagnosed based on signs/symptoms of HF with left ventricular ejection fraction ≥40–50% and structural/functional abnormalities indicative of left ventricular diastolic dysfunction, or other reference standards.[9] This result confirms the limited use of right heart catheterization also in research settings, thus making the case for the use of the HFA-PEFF and H_2FPEF scores as reasonable alternatives to this gold standard to make the diagnosis of HFpEF, even while their validation against right heart catheterization is still pending.[5–7] Under this light, future studies on potential diagnostic biomarkers of HFpEF should assess if these biomarkers have a similar diagnostic accuracy than existing diagnostic scores, and if they improve discrimination (i.e. the area under the curve values) when added to these scores. A head-to-head comparison of the two scores would be important to clarify if either of them can be used, or one of them should be preferred because of its greater accuracy.

It is interesting to consider that the 28 studies included in the Henkens *et al.* evaluated around 40 single biomarkers as well as 'miscellaneous miRNAs', 'metabolites' and 'proteins'.[9] In other words, a wide array of biomarkers was evaluated in studies with small sample sizes (down to 32 subjects, with a median number of just 154 subjects), and all the other methodological issues highlighted in the paper.[9] Therefore, there is an urgent need to improve methodological quality of studies searching for diagnostic biomarkers of HFpEF. The Authors should be congratulated for pointing out some of the most crucial issues that must be considered when designing similar studies. They should have preferably a prospective design, should enrol consecutive patients referred for exertional dyspnoea of unclear aetiology (using clear but not too stringent inclusion and exclusion criteria to avoid a selection bias), and be large enough to capture the phenotypic heterogeneity and comorbidity burden of HFpEF and to enable meaningful subgroup analyses, possibly by means of international collaborative studies. As correctly pointed out by the Authors, biomarkers should be 'measured at the same moment as the HFpEF diagnosis is made and before any intervention occurs', to avoid the confounding effect of medications such as diuretics.[9] When right heart catheterization cannot be systematically performed, one of the two validated scores (the HFA-PEFF and H_2FPEF scores) can represent an acceptable alternative as the reference standard to diagnose HFpEF. NP levels were included in the HFA-PEFF but not in the H2FPEF score. The diagnostic role of NPs in HFpEF deserves further consideration, and NP measurement could provide a link between diagnosis and the following steps of characterization of patient phenotype and risk prediction, which are all crucial to define a tailored therapeutic approach to HFpEF.

Conflict of interest: none declared.

References

1. Vasan RS, Xanthakis V, Lyass A, Andersson C, Tsao C, Cheng S, Aragam J, Benjamin EJ, Larson MG. Epidemiology of left ventricular systolic dysfunction and heart failure in the Framingham Study: an echocardiographic study over 3 decades. *JACC Cardiovasc Imaging* 2018;**11**:1–11.
2. Owan TE, Hodge DO, Herges RM, Jacobsen SJ, Roger VL, Redfield MM. Trends in prevalence and outcome of heart failure with preserved ejection fraction. *N Engl J Med* 2006;**355**:251–259.
3. Paulus WJ. H_2FPEF score: at last, a properly validated diagnostic algorithm for heart failure with preserved ejection fraction. *Circulation* 2018;**138**:871–873.
4. Obokata M, Kane GC, Reddy YN, Olson TP, Melenovsky V, Borlaug BA. Role of diastolic stress testing in the evaluation for heart failure with preserved ejection fraction: a simultaneous invasive-echocardiographic study. *Circulation* 2017;**135**:825–838.
5. Pieske B, Tschöpe C, de Boer RA, Fraser AG, Anker SD, Donal E, Edelmann F, Fu M, Guazzi M, Lam CS, Lancellotti P, Melenovsky V, Morris DA, Nagel E, Pieske-Kraigher E, Ponikowski P, Solomon SD, Vasan RS, Rutten FH, Voors AA, Ruschitzka F, Paulus WJ, Seferovic P, Filippatos G. How to diagnose heart failure with preserved ejection fraction: the HFA-PEFF diagnostic algorithm: a consensus recommendation from the Heart Failure Association (HFA) of the European Society of Cardiology (ESC). *Eur J Heart Fail* 2020;**22**:391–412.
6. Reddy YN, Carter RE, Obokata M, Redfield MM, Borlaug BA. A simple, evidence-based approach to help guide diagnosis of heart failure with preserved ejection fraction. *Circulation* 2018;**138**:861–870.
7. Barandiarán Aizpurua A, Sanders-van Wijk S, Brunner-La Rocca HP, Henkens M, Heymans S, Beussink-Nelson L, Shah SJ, van Empel VP. Validation of the HFA-PEFF score for the diagnosis of heart failure with preserved ejection fraction. *Eur J Heart Fail* 2020;**22**:413–421.
8. Carnes J, Gordon G. Biomarkers in heart failure with preserved ejection fraction: an update on progress and future challenges. *Heart Lung Circ* 2020;**29**:62–68.
9. Henkens MT, Remmelzwaal S, Robinson EL, van Ballegooijen AJ, Barandiarán Aizpurua MA, Verdonschot JA, Raafs AG, Weerts J, Hazebroek MR, van Sanders-Wijk S, Handoko ML, den Ruijter HM, Lam CS, de Boer RA, Paulus WJ, van Empel VP, Vos R, Brunner-La Rocca HP, Beulens JW, Heymans SR. Risk of bias in studies investigating novel diagnostic biomarkers for heart failure with preserved ejection fraction. A systematic review. *Eur J Heart Fail* 2020;**22**:1586–1597.