Evaluation of diagnostic accuracy of current biomarkers in heart failure with preserved ejection fraction: a systematic review and meta-analysis

Hao Chen[a,b], Michael Chhor[a], Benjamin Rayner[b], Kristine McGrath[a] and Lana McClements[a*]

[a]School of Life Sciences, Faculty of Science, University of Technology Sydney, NSW, Australia.

[b]Inflammation Group, Heart Research Institute, University of Sydney, NSW, Australia.

***Corresponding author**: Dr Lana McClements, School of Life Sciences, Faculty of Science, University of Technology Sydney, PO Box 123 Broadway NSW 2007 Australia.

Email: lana.mcclements@uts.edu.au

Tel: +61 (0) 2 9514 4372

Declarations of interest: none

Summary

Background

A number of circulating biomarkers are currently utilised for diagnosis of chronic heart failure with preserved ejection fraction (HFpEF). However, due to HFpEF heterogeneity, the accuracy of these biomarkers remains unclear.

Aims

This study aimed to systematically determine the diagnostic accuracy of currently available biomarkers for chronic HFpEF.

Methods

PubMed, Web of Science, MEDLINE and SCOPUS databases were searched systematically to identify studies assessing the diagnostic potentials of biomarkers of chronic HFpEF with left ventricular ejection fraction (LVEF) ≥50%. All included studies were independently assessed for quality and relevant information extracted. Random-effects models were used to estimate pooled diagnostic accuracy of the biomarkers of HFpEF.

Results

The search identified 6,145 studies with 19 studies included. Four biomarkers were available for meta-analyses. The pooled sensitivity of BNP (0.787, 95% CI =0.719, 0.842) was higher than that of NT-proBNP (0.696, 95% CI =0.599, 0.779) in chronic HFpEF diagnosis. However, compared to BNP (0.796, 95% CI =0.672, 0.882), NT-proBNP showed improved specificity (0.882, 95% CI =0.778, 0.941). Gal-3 exhibited a reliable diagnostic adequacy for HFpEF

(sensitivity: 0.760, 95% CI =0.631, 0.855; specificity: 0.803, 95% CI =0.667, 0.893). However, ST2 displayed limited diagnostic performance for chronic HFpEF diagnosis (sensitivity: 0.636, 95% CI =0.465, 0.779; specificity: 0.595, 95% CI =0.427, 0.743).

Conclusion

NT-proBNP and BNP appear the most reliable biomarkers in chronic HFpEF with NT-proBNP showing higher specificity and BNP higher sensitivity. Although Gal-3 appears more reliable than ST2 in HFpEF diagnosis, the conclusions are limited as only three studies were included in this meta-analysis.

Résumé

Contexte

Un certain nombre de biomarqueurs circulants est actuellement utilisé pour le diagnostic de l'insuffisance cardiaque chronique avec fraction d'éjection préservée (HFpEF). Cependant, en raison de l'hétérogénéité de HFpEF, la précision de ces biomarqueurs demeure incertaine.

Objectif

Cette étude vise à déterminer de manière systématique la précision diagnostique des biomarqueurs actuellement disponibles pour HFpEF chronique.

Méthodes

Les bases de données PubMed, Web of Science, MEDLINE et SCOPUS ont été utilisées pour identifier les études évaluant les potentiels de diagnostique des biomarqueurs de HFpEF

chronique avec une fraction d'éjection ventriculaire gauche (FEVG) ≥50%. Toutes les études retenues ont chacune été évaluées pour la qualité et la pertinence des données obtenues. Des modèles à effets aléatoires ont été utilisés pour estimer l'exactitude diagnostique groupée des biomarqueurs de HFpEF.

Résultats

Cette étude a permis d'identifier 6,145 études. Les données de 19 d'entre elles ont été utilisées pour cette recherche. Quatre biomarqueurs ont été identifiés et évalués pour les méta-analyses. La sensibilité combinée du BNP (0,787, IC à 95% = 0,719, 0,842) était plus élevée que celle du NT-proBNP (0,696, IC à 95% = 0,599, 0,779) dans le diagnostic de HFpEF chronique. Cependant, par rapport au BNP (0,796, IC à 95% = 0,672, 0,882), le NT-proBNP a montré une meilleure spécificité (0,882, IC à 95% = 0,778, 0,941). Gal-3 a montré un potentiel diagnostique fiable pour HFpEF (sensibilité: 0,760, IC à 95% = 0,631, 0,855; spécificité: 0,803, IC à 95% = 0,667, 0,893). Cependant, ST2 a montré des performances diagnostiques limitées pour HFpEF chronique (sensibilité: 0,636, IC 95% = 0,465, 0,779; spécificité: 0,595, IC 95% = 0,427, 0,743).

Conclusions

La fiabilité diagnostique du NT-proBNP et du BNP semble être la plus prometteuse pour HFpEF chronique, avec une meilleure spécificité pour le NT-proBNP et une meilleure sensibilité pour le BNP. Bien que Gal-3 semble plus fiable que ST2 dans le diagnostic de HFpEF, les conclusions sont limitées car seules trois études ont été incluses dans cette méta-analyse.

Keywords

heart failure with preserved ejection fraction; HFpEF, biomarker; diagnosis; meta-analysis.

Mots clés

insuffisance cardiaque avec fraction d'éjection préservée; HFpEF; biomarqueur; diagnostic; méta-analyse.

Abbreviations

BNP, B-type natriuretic peptide; Gal-3, galectin-3; HFpEF, heart failure with preserved ejection fraction; LVEF, left ventricular ejection fraction; NT-proBNP, N-terminal pro-B-type natriuretic peptide; ST2, suppression of tumorigenesis-2.

Background

Heart failure (HF) is an increasingly prominent disease in developed countries, placing a significant burden on patients and healthcare systems. It currently affects ~64 million people worldwide, with rising prevalence [1]. HF is a complex syndrome characterised by abnormal cardiac structure and function of the heart with impaired ability to fill and/or eject blood at normal pressure. In line with this definition, the latest clinical guidelines commonly classify HF into two subtypes based on the left ventricular ejection fraction (LVEF) [2], [3]. A LVEF <50% is typically considered as HF with reduced LVEF (HFrEF), and LVEF ≥50% is defined as HF with preserved LVEF (HFpEF). Nevertheless, HF patients with LVEF ranging from 40% to 50% has recently been classified as HF with mid-range EF (HFmrEF) [2] or HFpEF borderline [3], an emerging grey area between HFrEF and HFpEF. HFpEF has increased in prevalence over the last number of years and is now associated with similar mortality rates as HFrEF [4]. However, this is controversial and HFrEF is still considered the more dangerous types of HF with the higher mortality rates [5], [6]. Although HFpEF is often associated with less severe manifestations, currently available treatments remain limited for symptomatic control and ineffective for HFpEF management [7].

Circulating biomarkers are employed regularly in the diagnosis and prognosis of HFpEF. They have additional potential to provide a better understanding of the underlying pathogenesis, which could lead to the development of effective therapies. Natriuretic peptides (NPs), including B-type (BNP) and N-terminal pro-BNP (NT-proBNP), are recommended for diagnosis of HFpEF [2], [3]. In addition, galectin-3 (Gal-3) and suppression of tumorigenesis-2 (ST2) are emerging as clinical markers for risk stratification of HFpEF [3].

Nevertheless, their diagnostic reliability remains controversial due to the heterogeneity of data reported. Meta-analyses were performed on the diagnostic accuracy of NT-proBNP and BNP for HFpEF with substantial heterogeneity observed [8], which may affect the interpretability of results. Another relevant meta-analysis reported biomarkers in female patients with HFpEF and pre-eclampsia, whereas there were insufficient included studies for meta-analyses in HFpEF solely [9]. In this study, we systematically performed meta-analyses to comprehensively assess the diagnostic potential of all current biomarkers in the context of HFpEF (defined as LVEF ≥50%).

Methods

Search strategy and selection criteria

The systematic search was conducted to assess the diagnostic accuracy of biomarkers in HFpEF using the following databases: PubMed, Web of Science, MEDLINE and SCOPUS (1900 to February 2021). The literature search was performed using the following terms "HFpEF AND biomarker" as well as other synonymous terms outlined in the Supplementary material online (Appendix S1). We included studies that defined HFpEF as per the latest clinical guidelines published by the American Heart Association or European Society of Cardiology including the presence of symptoms and signs of HF, and LVEF ≥50% as confirmed by echocardiography [2], [3]. The history of congestive HF and the aetiology of HFpEF were not restricted in the definition of HFpEF.

To determine the biomarkers' suitability for HFpEF diagnosis, published data from observational studies assessing the diagnostic accuracy of individual biomarkers to discriminate between cohorts with and without chronic HFpEF, was included. Studies were selected if diagnostic performance measures of individual biomarkers were reported. Studies were excluded if they were: non-English language publications, letters, editorials, conference abstracts, meta-analyses and reviews. The secondary or post-hoc studies in the excluded meta-analysis or review publications were considered only if the inclusion criteria were met.

Data extraction

Two independent investigators (HC, MC) extracted data from included studies. Disagreements were resolved by consensus with a third investigator (LM). The recommendations of PRISMA guidelines [10] and a relevant guideline specialising for biomarker meta-analysis [11] were followed for data extraction. A conventional 2 x 2 table consisting of true positive (TP), true negative (TN), false positive (FP) and false negative (FN), was extracted from each included study. Only published data were extracted.

Quality assessment

The included studies were assessed for quality independently by three co-authors (MC, BR, KM) using Quality Assessment for Diagnostic Accuracy Studies-2 (QUADAS-2) tool [12], which was composed of four domains including i) patient selection, ii) index test, iii) reference standard and iv) patient flow and timing. Low risk of bias in a domain referred to positive answers in all sub-questions. High risk of bias in a domain was defined as negative answers in two / three sub-questions. Unclear risk of bias was defined as one / two negative answers.

Results were compared between assessors and, in case of disagreement, individual studies were discussed to achieve consensus.

Statistical analysis

The analyses of diagnostic accuracy test (DTA) were performed in R (4.0.3) using 'mada' package, where a bivariate, random-effects meta-analysis model was applied. The analyses of diagnostic biomarkers were based on sensitivity and specificity discriminating between cohorts with and without HFpEF. The estimated sensitivity and specificity were calculated using the 2 x 2 tables extracted from included studies. The sensitivity and specificity were pooled and analysed to generate random-effects model forest plots and random-effects model hierarchical summary of receiver operating characteristic (HSROC) curves. Natural logarithm (ln)-transformed diagnostic odds ratio (DOR) was reported along with heterogeneity of Higgins' I^2 and Cochran's Q. Publication bias was assessed through visual inspection of funnel plots of ln(DOR). Analyses were generated only for those diagnostic markers, which were evaluated in three or more independent studies.

Results

Search results

The results for diagnostic markers of HFpEF yielded 6,145 articles, of which 19 [13-31] met the inclusion criteria with sufficient evidence to conduct meta-analyses on individual biomarkers (Table 1; Fig. 1a). The overall quality of these included studies was comparably high (Fig. 1b and 1c). The number of prospective and retrospective designs were equal across the included studies (n=10 and n=9, respectively; Table 1). In total, n=1,486 patients with

HFpEF and n=1,416 without HFpEF were included from all 19 studies. All patients were at the chronic stage of HFpEF and free of valvular diseases. Patients with HFpEF were reported at old age (>50 years old), with a non-diseased group appropriately matched for age and sex. Overall, selected studies yielded a total of four different diagnostic markers including NT-proBNP, BNP, Gal-3 and ST2. NPs were the most commonly reported diagnostic markers (n=17) studies, which is in line with the well-established role of NPs in current HFpEF management [2], [3]. We were unable to complete meta-analyses on emerging biomarkers such as matrix metalloproteinases (MMPs) and growth differentiation factor 15 (GDF15), due to a small number of studies identified in relation to their diagnostic potential in HFpEF (n<3). However, these biomarkers along with their supporting citations were recorded in Supplementary file (Table S1 and References S1).

N-terminal pro-B-type natriuretic peptide (NT-proBNP)

Studies using NT-proBNP as a diagnostic marker of chronic HFpEF (n=12 studies [13-24], n=975 patients), reported optimal sensitivity and specificity at NT-proBNP cut-off concentration ranging from 65 pg/mL to 477 pg/mL with the median at 227 pg/mL (Fig. 2a). Interestingly, the four studies using NT-proBNP cut-off at ~227 pg/mL [19-22], showed distinct values of sensitivity but consistent specificity. The pooled DOR was 2.97 (95% CI =2.19, 3.76) (Fig. 2b). Relatively low heterogeneity was observed (Higgins' I^2=26.32%, Cochran's Q =14.938, p =0.185). The random-effects HSROC curve revealed moderate sensitivity (0.696, 95% CI =0.599, 0.779) and reliable specificity (0.882, 95% CI =0.778, 0.941) in terms of the diagnostic performance of NT-proBNP in HFpEF, with area under the curve (AUC) estimated 0.836 (Fig. 2c). Fig. 2d showed the 95% CI region of each study using NT-proBNP as a diagnostic marker. Generally, the 95% CI region of specificity appeared larger than that of sensitivity for most

relevant studies. According to the funnel plot (Supplementary material online Figure S1), there was some evidence of publication bias in NT-proBNP. However, the high statistical significance (p <0.01) of all 12 relevant studies suggested that the publication bias is not the underlying cause of this funnel asymmetry.

B-type natriuretic peptide (BNP)

Seven studies [18], [24], [25-29] investigating the diagnostic performance of BNP in HFpEF were analysed, with data extracted from 367 patients with HFpEF (Fig. 3a). The cut-off levels of BNP were varied (40-353.6 pg/mL, median: 125 pg/mL). In the random-effects forest plot (Fig. 2b), the pooled DOR was 2.70 (95% CI =1.68, 3.72), with no heterogeneity observed (Higgins' I^2=0%, Cochran's Q =4.422, *p* =0.620). Pooled estimated sensitivity (0.787, 95% CI =0.719, 0.842) and specificity (0.796, 95% CI =0.672, 0.882) were well-balanced when using BNP to diagnose HFpEF (Fig. 3c). The pooled AUC was 0.842. The number of participants was relatively small in three studies [25-27], resulting in the largest variance shown in Fig. 3d. Similar to NT-proBNP, the funnel plot of BNP is asymmetrical (Supplementary material online Figure S2). However, the high statistical significance (p <0.01) of all relevant studies suggested that the publication bias is not the underlying cause of this funnel asymmetry.

Galectin-3 (Gal-3)

Analyses were performed on the diagnostic accuracy of Gal-3 using three studies [15], [19], [30]. The data were evaluated based on a total of 362 patients with HFpEF (Fig. 4a). Gal-3 was reported at 1.79 ng/mL to 10.68 ng/mL (median: 9.55 ng/mL) as cut-off levels. The pooled DOR was 2.94 (95% CI =1.61, 4.28), whereas substantial heterogeneity was observed (Higgins' I^2=48.598%, Cochran's Q =3.891, *p* =0.143) (Fig. 4b). Sensitivity was relatively high (0.760, 95%

CI =0.631, 0.855), so was specificity (0.803, 95% CI =0.667, 0.893) (Fig. 4c). The AUC was 0.851 for the diagnostic performance of Gal-3. Fig. 4d exhibited larger variance on specificity than sensitivity.

Suppression of tumorigenesis-2 (ST2)

Three studies [15], [20], [31] reported the diagnostic accuracy of ST2 in chronic HFpEF, with an adequate pooled number of patients with HFpEF (n=290), and the distribution of participants well-balanced across the studies (Fig. 5a). The cut-off levels of ST2 were substantially varied across the three studies, ranging from 68.6 pg/mL to 26.57 ng/mL. The pooled DOR of ST2 as an individual diagnostic marker in HFpEF was 1.00 (95% CI: -0.07, 2.07), with minimal heterogeneity (Higgins' I^2=3.959%, Cochran's Q =2.082, p =0.353) (Fig. 5b). In line with the poor DOR, sensitivity (0.636, 95% CI =0.465, 0.779) and specificity (0.595, 95% CI =0.427, 0.743) as well as AUC (0.647) were all unreliable (Fig. 5c). Although the number of participants was satisfactory in each study, the reported diagnostic accuracy highly varied, particularly in terms of specificity (Fig. 5d).

Discussion

HF may be categorised as either acute or chronic, where it is possible and common for patient to experience acute episodes (e.g. acute exacerbation or decompensation) of HF with underlying chronic symptoms in HF patients. Chronic underlying HFpEF accounts for a large proportion of its population and it must be noted that the biomarkers assessed in this study were performed in the context of patients with chronic HFpEF [32].

The diagnosis of chronic HFpEF is difficult as a multifactorial syndrome; it not only requires preserved LVEF, but additional symptoms of chronic HF [33]. However, the reliability of currently available biomarkers in diagnosis of HFpEF remains partially unclear. Our study is the first to systematically and comprehensively review the currently available circulating biomarkers (defined as proteins detected in blood-derived samples) in diagnosis of chronic HFpEF. The main findings of this study are: i) NT-proBNP (DOR =2.97) and BNP (DOR =2.70) remain the two most reliable individual diagnostic markers for HFpEF, whereas the diagnostic adequacy of both NPs remains only reasonable; ii) NT-proBNP shows higher specificity (0.882) than BNP (0.796) in diagnosis of chronic HFpEF, whereas the sensitivity and specificity of BNP (0.787 and 0.796, respectively) are more balanced compared to NT-proBNP (0.696 and 0.882, respectively); iii) Gal-3, an emerging biomarker for HFpEF management, displays promising diagnostic performance (DOR =2.94) for HFpEF; and iv) ST2 shows limited accuracy (DOR =1.00) as an individual biomarker for the diagnosis of chronic HFpEF.

Compared to another HFpEF biomarker meta-analysis [8], a lower degree of heterogeneity was reported in our study, as the heterogeneity statistics were only utilised for estimation of DOR rather than sensitivity and specificity. However, substantial heterogeneity remains on the diagnostic accuracy of Gal-3, which could be due to the retrospective design of all relevant selected studies [15], [19], [30]. In addition, an 100% specificity was introduced by one of the studies [19], which could be possibly due to a random chance. Another explanation for the heterogeneity could be caused by the wide difference of cut-off levels of Gal-3 (1.79 ng/mL, 9.55 ng/mL and 10.68 ng/mL). The heterogenous nature of HFpEF may also play a role in this heterogeneity. All these underlying causes of heterogeneity could limit the applicability of the

results of Gal-3. Therefore, it is important to note that the reliable diagnostic discriminative power of Gal-3 remains questionable.

A limited number of studies were included for evaluating the diagnostic accuracy of Gal-3 and ST2 in HFpEF, with only 362 and 290 patients with HFpEF, respectively. Trends of rising HF prevalence are shared amongst all countries, yet it must be noted that the studies included in this meta-analysis were conducted in Asia. With factors such as an ageing population and younger age range for HF patients, the generalisation of the findings in this paper to patients of different ethnicities may be limited [32].

Natriuretic peptides are currently the most widely utilised biomarkers in supporting HFpEF diagnosis. Frequently, laboratories and clinical guidelines recommend the use of NT-proBNP over BNP in HFpEF diagnosis as the first line option. This is likely due to the stability of NT-proBNP within blood samples for over 72 hours at room temperature without the need for using additives. On the other hand, BNP is stable within blood samples for only 24 hours at room temperature, and the blood collection tubes are required to be coated an ethylenediaminetetraacetic acid (EDTA) [34].

NT-proBNP and BNP are strongly recommended for HFpEF diagnosis by the current clinical guidelines [2], [3], generating a number of high-quality observational studies. As such, the diagnostic reliability of NT-proBNP and BNP is well-validated in our study. In this DTA meta-analysis, the pooled specificity of NT-proBNP in diagnosing HFpEF was higher than that of BNP, however the pooled sensitivity of BNP was more improved than NT-proBNP, consistent with another HFpEF biomarker meta-analysis [8]. Interestingly, both sensitivity and specificity of

BNP were well-balanced and reasonable. The AUC and DOR of NT-proBNP and BNP were satisfactory for diagnostic purposes. Therefore, the reliability of NT-proBNP and BNP is equal as diagnostic markers for chronic HFpEF, given that both NPs are in the same biological pathway [35]. However, differential sensitivity and specificity were reported for NT-proBNP and BNP in HFpEF diagnosis, suggesting different utility in clinical settings. Due to the high specificity in NT-proBNP in HFpEF diagnosis, we suggest that NT-proBNP possesses a prominent advantage in ruling out HFpEF. Higher sensitivity could be more preferable in secondary or tertiary care, whereas a reliable specificity could be more important in primary care settings.

Overall, fairly consistent cut-off levels of NT-proBNP were reported by relevant studies, with the most optimal specificity being observed at approximately 100 pg/mL [14], [16], [23], which is consistent with the cut-off (>125 pg/mL) suggested by 2016 European Society of Cardiology (ESC) clinical guidelines for HF [2] and the new HFA-PEFF diagnostic algorithms [36]. Three selected studies utilised significantly higher cut-off values of NT-proBNP (295.85 pg/mL [24], 424.31 pg/mL [17] and 477 pg/mL [18]), which led to the lowest specificity. This could be further supporting evidence for utilising the recommended cut-off values of NT-proBNP at approximately 100 pg/mL for diagnosis of HFpEF. Despite the fact that the recommended cut-off level of BNP is 35 pg/mL [2], [36], the cut-off values reported by included studies related to BNP was widely varied. In addition, a study reported that the cut-off value of ~35 pg/mL provided an unreliable diagnostic accuracy (sensitivity: 0.67; specificity: 0.73) for chronic HFpEF [37]. However, significantly higher cut-off levels of BNP were observed in most relevant studies. Further population-based comparable investigations of the diagnostic performance of BNP at different cut-off concentrations for HFpEF diagnosis are necessary.

ST2 is emerging as a new diagnostic marker for HFpEF and is recommended by the latest American Heart Association (AHA) guidelines [3]. Nevertheless, we observed a limited diagnostic accuracy of ST2 in chronic HFpEF diagnosis, supported by three studies [15], [20], [31]. The limited diagnostic value of ST2 in HFpEF is likely caused by the lack of association of ST2 with LV function and structure [38]. Despite the limited performance of ST2 in chronic HFpEF, ST2 is beneficial in acute settings [39]. Although ST2 was shown to be associated with HF diagnosis at 35 ng/mL, as recommended by the Food and Drug Administration (FDA), the diagnostic adequacy in HF subtypes including HFpEF and HFrEF, were modest in the elderly population [40]. Therefore, it is recommended that the optimal cut-off value of ST2 in HF subtypes should be re-evaluated in future observational studies.

Collectively, the specificity of NT-proBNP, BNP and Gal-3 are generally higher than their sensitivity, suggesting a more advanced ability of ruling out HFpEF, consistent with the proposals in current guidelines [2], [41]. Generally, these biomarkers play a critical role in discriminating acute HF from non-cardiac dyspnoea in acute settings, as their concentrations were significantly elevated [39], [42]. This is opposite in chronic settings where the levels of biomarkers could be closer to normal ranges. Therefore, diagnosis of chronic HFpEF is difficult, especially given the common comorbidities which further complicate the diagnosis. Overall, in line with the recommendations of HFA-PEFF diagnostic algorithms [36], biomarkers should be used on top of echocardiography for the early diagnosis of HFpEF. Future studies should therefore investigate the clinical utility of current biomarkers in combination with echocardiographic measurements.

Conclusions

HFpEF comprises approximately half of all patients with HF, and it is associated with similar mortality as HFrEF, yet it is ineffectively managed with pharmacotherapies. Due to poorly understood pathogenesis of HFpEF, there is a likely delay in the diagnosis and treatment, leading to worse outcomes for HFpEF patients. Accurate biomarkers are critical for early diagnosis of HFpEF, emphasising the urgent need for biomarker discovery and validation. Nevertheless, in this meta-analysis it was demonstrated that NT-proBNP and BNP remain the most reliable biomarker in HFpEF diagnosis. NT-proBNP is possibly more reliable for chronic HFpEF diagnosis given its more consistent and less varied cut-off diagnostic values and higher specificity than BNP. Gal-3 also displays a reliable diagnostic discriminative power, while the high heterogeneity limits the applicability of Gal-3's high diagnostic value. ST2 appears to have limited diagnostic potential for chronic HFpEF. Therefore, more robust and larger future studies are warranted.

Authors' contributions

HC conducted the search, identified the studies, performed the statistical analyses and wrote the first draft of this manuscript. MC conducted the search, screened, assessed and identified the studies, extracted the data and contributed to the writing. BR, KM and LM supervised HC. BR and KM reviewed the quality of the studies. LM conceptualised the study and edited the manuscript. All authors reviewed and approved the manuscript.

Sources of funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure of interest

The authors declare that they have no competing interest.

References

[1] Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. Eur J Heart Fail 2020;22:1342-1356.

[2] Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J 2016;37:2129-2200.

[3] Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Colvin MM, et al. 2017 ACC/AHA/HFSA Focused Update of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. Circulation 2017;136:e137-e161.

[4] 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 Eng J Med 2006;355:251-259.

[5] Lam CSP, Gamble GD, Ling LH, Sim D, Leong KTG, Yeo PSD, et al. Mortality associated with heart failure with preserved vs. reduced ejection fraction in a prospective international multi-ethnic cohort study. Eur Heart J 2018;39:1770-1780.

[6] Meta-Analysis Global Group In Chronic Heart Failure (MAGGIC). The survival of patients with heart failure with preserved or reduced left ventricular ejection fraction: an individual patient data meta-analysis. Eur Heart J 2018;33:1750-1757.

[7] Zheng SL, Chan FT, Nabeebaccus AA, Shah AM, McDonagh T, Okonko DO, et al. Drug treatment effects on outcomes in heart failure with preserved ejection fraction: a systematic review and meta-analysis. Heart 2018;104:407-415.

[8] Remmelzwaal S, van Ballegooijen AJ, Schoonmade LJ, Canto ED, Handoko ML, Henkens MTHM, et al. Natriuretic peptides for the detection of diastolic dysfunction and heart failure with preserved ejection fraction—a systematic review and meta-analysis. BMC Med 2020;18:290.

[9] Alma LJ, Bokslag A, Maas AHEM, Franx A, Paulus WJ, de Groot CJM. Shared biomarkers between female diastolic heart failure and pre-eclampsia: a systematic review and meta-analysis. ESC Heart Fail 2017;4:88-98.

[10] Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA GROUP. Preferred reporting items for systematic reviews and meta-analyses. BMJ 2009;339:b2535.

[11] Liu Z, Yao Z, Li C, Liu X, Chen H, Gao C. A step-by-step guide to the systematic review and meta-analysis of diagnostic and prognostic test accuracy evaluations. Br J Cancer 2013;108:2299-2303.

[12] Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-536.

[13] Barutcuoglu B, Parildar Z, Basol G, Gurgun C, Tekin Y, Bayindir O. The detection of left ventricular diastolic dysfunction in hypertensive patients: Performance of N-terminal probrain natriuretic peptide. Blood Press 2010;19:212-217.

[14] Celik A, Koc F, Kadi H, Ceyhan K, Erkorkmaz U, Burucu T, et al. Relationship between red cell distribution width and echocardiographic parameters in patients with diastolic heart failure. Kaohsiung J Med Sci 2012;28:165-172.

476 [15] Cui Y, Qi X, Huang A, Li J, Hou W, Liu K. Differential and predictive value of galectin-3 and
477 soluble suppression of tumorigenicity-2 (sST2) in heart failure with preserved ejection
478 fraction. Med Sci Monit 2018;24:5139-5146.
479 [16] Kasner M, Gaub R, Westermann D, Kaplan H, Akpulat S, Steendijk P, et al. Simultaneous
480 estimation of NT-proBNP on top to mitral flow Doppler echocardiography as an accurate
481 strategy to diagnose diastolic dysfunction in HFNEF. Int J Cardiol 2011;149:23-29.
482 [17] Liu S, Iskandar R, Chen W, Zhang Jie, Wang Y, Chen X, et al. Soluble glycoprotein 130 and
483 heat shock protein 27 as novel candidate biomarkers of chronic heart failure with preserved
484 ejection fraction. Heart Lung Circ 2016;25:1000-1006.
485 [18] Mason MJ, Hancock HC, Close H, Murphy JJ, Fuat A, de Belder M, et al. Utility of
486 biomarkers in the differential diagnosis of heart failure in older people: findings from the
487 heart failure in care homes (HFinCH) diagnostic accuracy study. PLoS One 2013;8:e53560.
488 [19] Polat V, Bozcali E, Uygun T, Opan S, Karakaya O. Diagnostic significance of serum galectin-
489 3 levels in heart failure with preserved ejection fraction. Acta Cardiol 2016;71:191-197.
490 [20] Santhanakrishnan R, Chong JPC, Ng TP, Ling LH, Sim D, Leong KTG, et al. Growth
491 differentiation factor 15, ST2, high-sensitivity troponin T, and N-terminal pro brain natriuretic
492 peptide in heart failure with preserved vs. reduced ejection fraction. Eur J Heart Fail
493 2012;14:1338-1348.
494 [21] Shuai XX, Chen YY, Lu YX, Su GH, Wang YH, Zhao HL. et al. Diagnosis of heart failure with
495 preserved ejection fraction: which parameters and diagnostic strategies are more valuable?
496 Eur J Heart Fail 2011;13:737-745.
497 [22] Stahrenberg R, Edelmann F, Mende M, Kockskamper A, Dungen HD, Luers C, et al. The
498 novel biomarker growth differentiation factor 15 in heart failure with normal ejection fraction.
499 Eur J Heart Fail 2010;12:1309-1316.

500 [23] Tschope C, Kasner M, Westermann D, Gaub R, Poller WC, Schultheiss HP. The role of NT-
501 proBNP in the diagnostics of isolated diastolic dysfunction: correlation with
502 echocardiographic and invasive measurements. Eur Heart J 2005;26:2277-2284.
503 [24] Zapata L, Betbese AJ, Roglan A, Ordonez-Llanos J. Use of B-type natriuretic peptides to
504 detect the existence and severity of diastolic dysfunction in non-cardiac critically ill patients:
505 a pilot study. Minerva Anestesiol 2014;80:194-203.
506 [25] Arques S, Roux E, Sbragia P, Pieri B, Gelisse R, Luccioni R, et al. Usefulness of bedside
507 tissue Doppler echocardiography and B-type natriuretic peptide (BNP) in differentiating
508 congestive heart failure from noncardiac cause of acute dyspnea in elderly patients with a
509 normal left ventricular ejection fraction and permanent, nonvalvular atrial fibrillation: insights
510 from a prospective, monocenter study. Echocardiography 2007;24:499-507.
511 [26] Dokainish H, Zoghbi WA, Lakkis NM, Quinones MA, Nagueh SF. Comparative accuracy of
512 B-type natriuretic peptide and tissue doppler echocardiography in the diagnosis of congestive
513 heart failure. Am J Cardiol 2004;93:1130-1135.
514 [27] Liu H, Zhang YZ, Gao M, Liu BC. Elevation of B-type natriuretic peptide is a sensitive
515 marker of left ventricular diastolic dysfunction in patients with maintenance haemodialysis.
516 Biomarker 2010;15:533-537.
517 [28] Lubien E, DeMaria A, Krishnaswamy P, Clopton P, Koon J, Kazanegra R, et al. Utility of B-
518 natriuretic peptide in detecting diastolic dysfunction: comparison with Doppler velocity
519 recordings. Circulation 2002;105:595-601.
520 [29] Wei T, Zeng C, Chen L, Chen Q, Zhao R, Lu G, et al. Bedside tests of B-type natriuretic
521 peptide in the diagnosis of left ventricular diastolic dysfunction in hypertensive patients. Eur
522 J Heart Fail 2005;7:75-79.

[30] Wu CK, Su MY, Lee JK, Chiang FT, Hwang JJ, Lin JL, et al. Galectin-3 level and the severity of cardiac diastolic dysfunction using cellular and animal models and clinical indices. Sci Rep 2015;5:17007.

[31] Wang YC, Yu CC, Chiu FC, Tsai CT, Lai LP, Hwang JJ, et al. Soluble ST2 as a biomarker for detecting stable heart failure with a normal ejection fraction in hypertensive patients. J Card Fail 2013;19:163-168.

[32] Chen YT, Wong LL, Liew OW, Richards AM. Heart Failure with Reduced Ejection Fraction (HFrEF) and Preserved Ejection Fraction (HFpEF): The Diagnostic Value of Circulating MicroRNAs. Cells 2019;8:1651.

[33] Gazewood JD, Turner PL. Heart Failure with Preserved Ejection Fraction: Diagnosis and Management. Am Fam Physician 2017;96:582-588.

[34] Weber M, Hamm C. Role of B‐type natriuretic peptide (BNP) and NT‐proBNP in clinical routine. Heart 2006;92:843-849.

[35] de Lemos JA, McGuire DK, Drazner MH. B-type natriuretic peptide in cardiovascular disease. Lancet 2003;362:316-322.

[36] Pieske B, Tschope C, de Boer RA, Fraser AG, Anker SD, Donal E, et al. 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 Heart J 2019;40:3297-3317.

[37] Arques S, Jaubert MP, Bonello L, Roux E, Armero S, Sbragia P, et al. Usefulness of basal B-type natriuretic peptide levels for the diagnosis of diastolic heart failure in young patients: an echocardiographic-catheterization study. Int J Cardiol 2009;14:551-552.

[38] AbouEzzeddine OF, McKie PM, Dunlay SM, Stevens SR, Felker GM, Borlaug BA, et al. Suppression of tumorigenicity 2 in heart failure with preserved ejection fraction. J Am Heart Assoc 2017;6:e004382.

[39] Mueller T, Gegenhuber A, Leitner I, Poelz W, Haltmayer M, Dieplinger B. Diagnostic and prognostic accuracy of galectin-3 and soluble ST2 for acute heart failure. Clin Chim Acta 2016;463:158-164.

[40] Parikh RH, Seliger SL, Christenson R, Gottdiener JS, Psaty BM, deFilippi CR. Soluble ST2 for prediction of heart failure and cardiovascular death in an elderly, community-dwelling population. J Am Heart Assoc 2016;5:e003188.

[41] Paulus WJ, Tschope C, Sanderson JE, Rusconi C, Flachskampf FA, Rademakers FE, et al. How to diagnose diastolic heart failure: a consensus statement on the diagnosis of heart failure with normal left ventricular ejection fraction by the Heart Failure and Echocardiography Associations of the European Society of Cardiology. Eur Heart J 2007;28:2539-2550.

[42] Maisel AS, McCord J, Nowak RM, Hollander JE, Wu AHB, Duc P, et al. Bedside B-Type natriuretic peptide in the emergency diagnosis of heart failure with reduced or preserved ejection fraction. Results from the Breathing Not Properly Multinational Study. J Am Coll Cardiol 2003;41:2010-2017.

569 **Table 1.** Study characteristics of included studies.

Study	Study Design	Location	Mean LVEF in HFpEF (%)	n (HFpEF)	Women in HFpEF (%)	Mean age in HFpEF (years)	Mean LVEF in control (%)	n (Control)	Women in control (%)	Mean age in control (years)
Liu et al. 2016	Retrospective	China	NA	50	46	64.28	NA	50	54	63.76
Cui et al. 2018	Retrospective	China	60	172	55.8	73	58.5	30	40	67
Tschope et al. 2005	Prospective	Germany	68	68	46	51	65	50	44	49
Santhanakrishnan et al. 2012	Prospective	Singapore	60	50	42	69	66	50	54	63
Stahrenberg et al. 2010	Retrospective	Germany	60	142	64	73	62	188	66	56
Kasner et al. 2011	Prospective	Germany	NA	107	40	53	NA	73	43	51
Dokainish et al. 2004	Prospective	USA	NA	19	NA	NA	NA	27	NA	NA
Liu et al. 2010	Prospective	China	65	39	50	52.2	67	20	46.2	46
Wei et al. 2005	Prospective	China	65	61	32	70	67	74	35	66
Lubien et al. 2002	Prospective	USA	NA	119	10.9	71	NA	175	9.1	60
Wang et al. 2013	Retrospective	China	68	68	54.4	68	68	39	33.3	60
Arques et al. 2007	Prospective	France	60	15	27	58	62	11	55	57
Mason et al. 2013	Retrospective	UK	NA	57	NA	NA	NA	308	NA	NA
Shuai et al. 2011	Prospective	China	66	101	52	67	67	48	50	62
Polat et al. 2016	Retrospective	Turkey	59	44	45.5	60	61	38	47.5	57
Celik et al. 2012	Retrospective	Turkey	72	71	63.4	57.09	68	50	38	56.16
Zapata et al. 2014	Prospective	Spain	60	35	51.4	68	59	36	19.4	57
Barutcuoglu et al. 2010	Retrospective	Turkey	NA	122	51.3	55	NA	119	54.9	53
Wu et al. 2015	Retrospective	China	68.2	146	61.6	70.06	NA	30	63.3	63.23

570

571 Control is defined as participants without evidence of HF.

572 *LVEF* left ventricular ejection fraction; *HFpEF* heart failure with preserved ejection fraction.

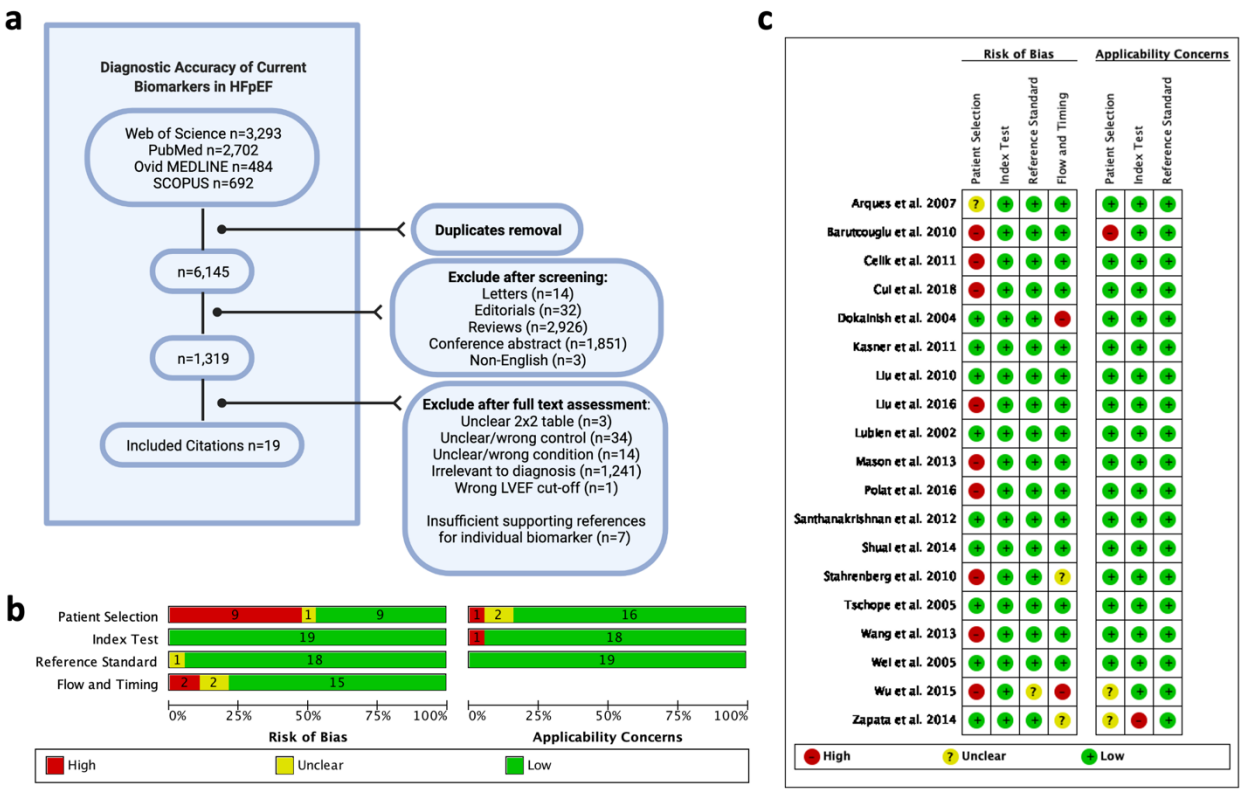


Figure 1. Summary of the study workflow and the number of included studies.

a Workflow of the systematic search adheres to PRISM guidelines. **b** Summary quality assessment of included studies independently evaluated using QUADAS-2 tool. **c** Outcomes of quality assessment of each individual included study.

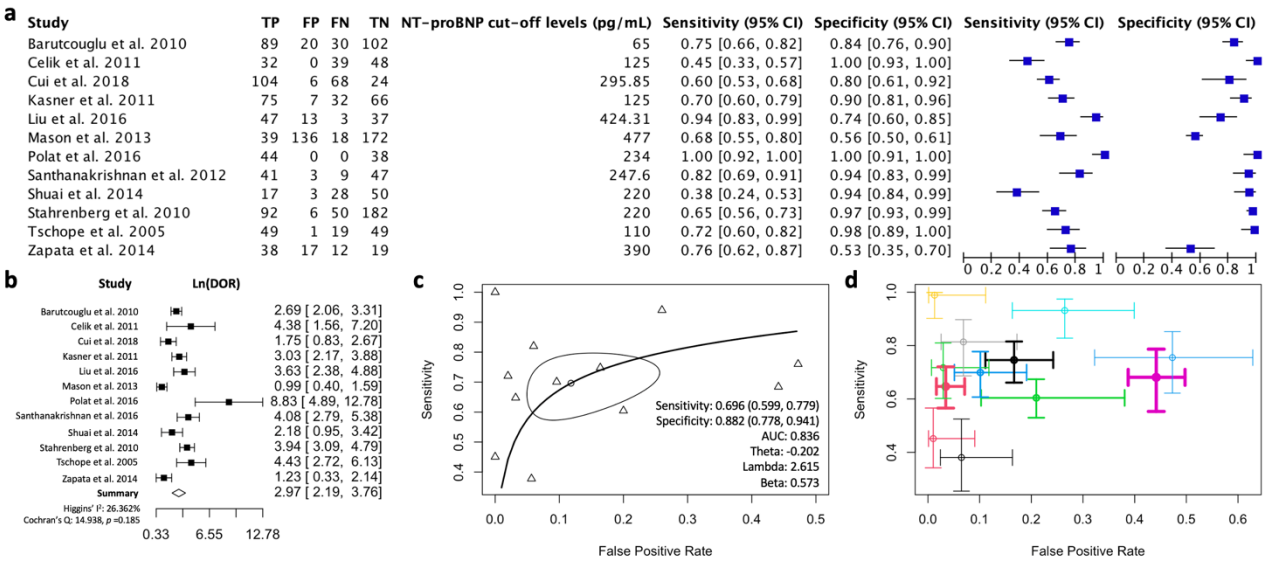


Figure 2. Diagnostic assessment of NT-proBNP in HFpEF using a bivariate, random-effects model.

a Forest plot of six studies investigating the diagnostic performance of NT-proBNP in HFpEF, with sensitivity and specificity reported. **b** Forest plot of ln(DOR) related to the diagnostic accuracy of NT-proBNP in HFpEF. **c** Plot of HSROC curve showing the estimated pooled diagnostic accuracy. **d** Plot of HSROC curve showing the 95% CI of each study evaluating the diagnostic accuracy of NT-proBNP in HFpEF.

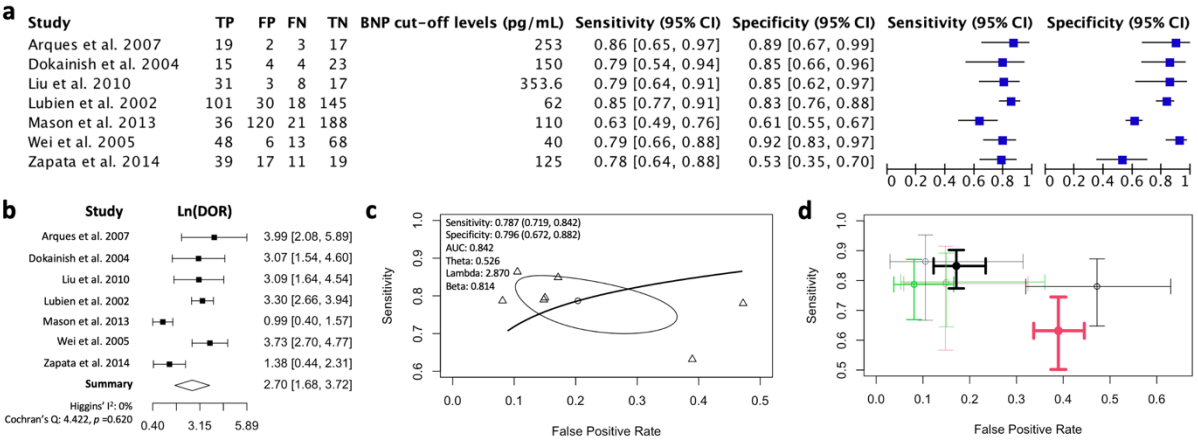
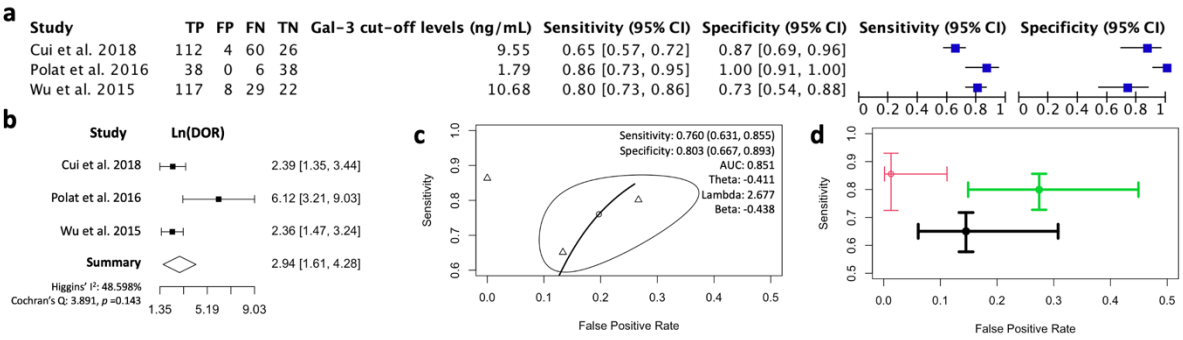


Figure 3. Diagnostic assessment of BNP in HFpEF using a bivariate, random-effects model. **a** Forest plot of five studies investigating the diagnostic performance of BNP in HFpEF, with sensitivity and specificity reported. **b** Forest plot of ln(DOR) related to the diagnostic accuracy of BNP in HFpEF. **c** Plot of HSROC curve showing the estimated pooled diagnostic accuracy. **d** Plot of HSROC curve showing the 95% CI of each study evaluating the diagnostic accuracy of BNP in HFpEF.



623

624 **Figure 4.** Diagnostic assessment of Gal-3 in HFpEF using a bivariate, random-effects model.

625 **a** Forest plot of three studies investigating the diagnostic performance of Gal-3 in HFpEF, with

626 sensitivity and specificity reported. **b** Forest plot of ln(DOR) regarding the diagnostic accuracy

627 of Gal-3 in HFpEF. **c** Plot of HSROC curve showing the estimated pooled diagnostic accuracy.

628 **d** Plot of HSROC curve showing the 95% CI of each study evaluating the diagnostic accuracy

629 of Gal-3 in HFpEF.

630

631

632

633

634

635

636

637

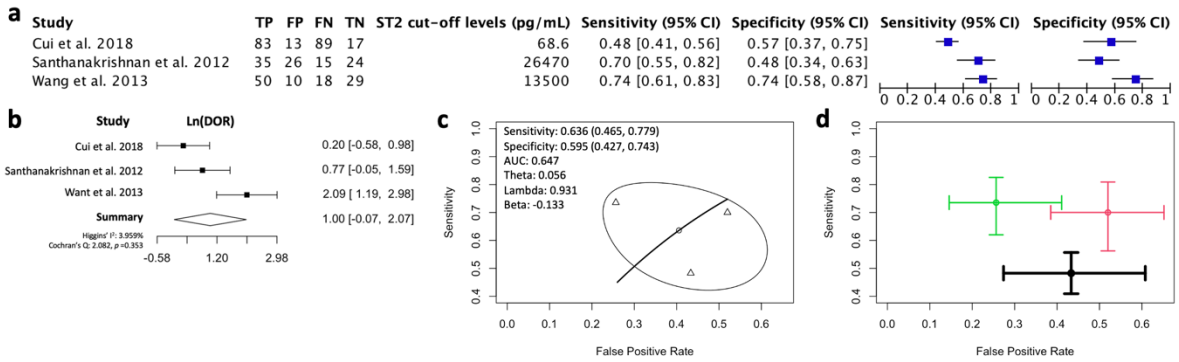
638

639

640

641

642



643

644 **Figure 5.** Diagnostic assessment of ST2 in HFpEF using a bivariate, random-effects model. **a**
645 Forest plot of three studies investigating the diagnostic performance of ST2 in HFpEF, with
646 sensitivity and specificity reported. **b** Forest plot of ln(DOR) regarding the diagnostic accuracy
647 of ST2 in HFpEF. **c** Plot of HSROC curve showing the estimated pooled diagnostic accuracy. **d**
648 Plot of HSROC curve showing the 95% CI of each study evaluating the diagnostic accuracy of
649 ST2 in HFpEF.