Risk of bias in studies investigating novel diagnostic biomarkers for heart failure with preserved ejection fraction. A systematic review

Henkens, Michiel T. H. M.; Remmelzwaal, Sharon; Robinson, Emma L.; van Ballegooijen, Adriana J.; Aizpurua, Arantxa Barandiaran; Verdonschot, Job A. J.; Raafs, Anne G.; Weerts, Jerremy; Hazebroek, Mark R.; Wijk, Sandra Sanders-van

Published in: European Journal of Heart Failure

DOI: 10.1002/ejhf.1944

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date: 2020

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Henkens, M. T. H. M., Remmelzwaal, S., Robinson, E. L., van Ballegooijen, A. J., Aizpurua, A. B., Verdonschot, J. A. J., Raafs, A. G., Weerts, J., Hazebroek, M. R., Wijk, S. S., Handoko, M. L., den Ruijter, H. M., Lam, C. S. P., de Boer, R. A., Paulus, W. J., van Empel, V. P. M., Vos, R., Rocca, H-P. B-L., Beulens, J. W. J., & Heymans, S. R. B. (2020). Risk of bias in studies investigating novel diagnostic biomarkers for heart failure with preserved ejection fraction. A systematic review. European Journal of Heart Failure, (9), 1586-1597. https://doi.org/10.1002/ejhf.1944

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Risk of bias in studies investigating novel diagnostic biomarkers for heart failure with preserved ejection fraction. A systematic review

Michiel T.H.M. Henkens1*, Sharon Remmelzwaal2, Emma L. Robinson1, Adriana J. van Ballegooijen2, Arantxa Barandiarán Aizpurua1, Job A.J. Verdonschot1,3, Anne G. Raafs1, Jeremey Weerts1, Mark R. Hazebroek1, Sandra Sanders-van Wijk1, M. Louis Handoko4, Hester M. den Ruijter5, Carolyn S.P. Lam6,7,8, Rudolf A. de Boer8, Walter J. Paulus9,10, Vanessa P.M. van Empel1, Rein Vos11, Hans-Peter Brunner-La Rocca1, Joline W.J. Beulens2,12, and Stephane R.B. Heymans1,10,13

1Department of Cardiology, Maastricht University Medical Centre, Maastricht University, Maastricht, The Netherlands; 2Department of Epidemiology and Biostatistics, Amsterdam Cardiovascular Sciences Research Institute, Amsterdam UMC, Amsterdam, The Netherlands; 3Department of Clinical Genetics, Maastricht University Medical Centre, Maastricht, The Netherlands; 4Department of Cardiology, Amsterdam Cardiovascular Sciences Research Institute, Amsterdam UMC, Amsterdam, The Netherlands; 5Laboratory of Experimental Cardiology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; 6National Heart Centre Singapore, Singapore, Singapore; 7Duke-National University of Singapore, Singapore, Singapore; 8Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; 9Department of Physiological Sciences Research Institute, Amsterdam UMC, Amsterdam, The Netherlands; 10Netherlands Heart Institute (CIN), Utrecht, The Netherlands; 11Department of Methodology and Statistics, Maastricht University, Maastricht, The Netherlands; 12Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands; and 13Department of Cardiovascular Research, University of Leuven, Leuven, Belgium

Received 20 March 2020; revised 19 June 2020; accepted 20 June 2020

Aim

Diagnosing heart failure with preserved ejection fraction (HFP EF) in the non-acute setting remains challenging. Natriuretic peptides have limited value for this purpose, and a multitude of studies investigating novel diagnostic circulating biomarkers have not resulted in their implementation. This review aims to provide an overview of studies investigating novel circulating biomarkers for the diagnosis of HFP EF and determine their risk of bias (ROB).

Methods and results

A systematic literature search for studies investigating novel diagnostic HFP EF circulating biomarkers in humans was performed up until 21 April 2020. Those without diagnostic performance measures reported, or performed in an acute heart failure population were excluded, leading to a total of 28 studies. For each study, four reviewers determined the ROB within the QUADAS-2 domains: patient selection, index test, reference standard, and flow and timing. At least one domain with a high ROB was present in all studies. Use of case-control/two-gated designs, exclusion of difficult-to-diagnose patients, absence of a pre-specified cut-off value for the index test without the performance of external validation, the use of inappropriate reference standards and unclear timing of the index test and/or reference standard were the main bias determinants. Due to the high ROB and different patient populations, no meta-analysis was performed.
Introduction

Heart failure with preserved ejection fraction (HFpEF) is a clinical syndrome that is associated with high mortality rates, poor quality of life and significant healthcare resource utilization.\(^1,2\) Currently, more than 5% of the elderly (>65 years of age) suffer from this debilitating syndrome.\(^3,4\) The prevalence is expected to rise even further in the upcoming years, due to the ageing population and the growing occurrence of other HFpEF risk factors.\(^5\)

Unfortunately, diagnosing HFpEF in the non-acute setting remains challenging. Natriuretic peptides (NPs) have limited diagnostic value for this purpose, which is mainly due to the high prevalence of conditions within this syndrome that can lead to higher [e.g. atrial fibrillation (AF), hypertension, pulmonary diseases, renal function disorders] and lower (e.g. obesity) circulating NP levels.\(^3–11\) Moreover, 18% to 30% of patients with haemodynamically proven HFpEF have NP levels below ‘diagnostic’ threshold.\(^12–14\)

The limited diagnostic accuracy of NPs, and the concept that other circulating biomarkers could help to diagnose this complex syndrome on a molecular level, has resulted in a multitude of studies investigating novel diagnostic HFpEF biomarkers.\(^3,11\) Remarkably, none of the suggested circulating biomarkers have been implemented in the HFpEF clinics. The heterogeneous and systemic nature of the syndrome could contribute to their lack of success,\(^11\) but a comprehensive overview of the literature on this topic is absent. We therefore aimed to provide an overview of studies investigating the diagnostic value of novel biomarkers for non-acute HFpEF and determine their risk of bias (ROB).

Methods

A systematic literature search—based on the PRISMA-DTA statement\(^15\)—of PubMed and EMBASE was performed to find diagnostic papers within the field of HFpEF from its inception until 21 April 2020. A broad search (online supplementary Appendix S1) was used for a set of systematic reviews and a meta-analysis for the (early) detection of left ventricular diastolic dysfunction (LVDD) and/or HFpEF. The search strategy and the protocol can be found on PROSPERO (CRD42018065018). Studies that reported the diagnostic value within a rare patient population (e.g. beta thalassemia); or (ii) were written in English. Studies were excluded if they: (i) studied the diagnostic value of a biomarker in acute heart failure; (ii) only studied the diagnostic value of NPs; (iii) studied the diagnostic value within a rare patient population (e.g. beta thalassemia); or (iv) were a (systematic) review, meta-analysis, editorial, or conference abstract.

Data extraction

The following data were extracted for each study: publication details (first author, year of publication), study characteristics (patient population description, exclusion criteria), used reference standard, and the biomarker(s) studied (index test).

Risk of bias assessment

The methodological quality of the full-text articles was independently evaluated by four reviewers (SR, ER, RV, MH) by utilising the QUADAS-2 tool.\(^17\) This tool was used to determine the ROB within four domains: patient selection, index test, reference standard, and flow and timing. Based on the information provided in the included studies, the ROB was rated low, intermediate, or high for these domains separately.

For the reference standard domain the ROB was rated low if (exercise) right-sided heart catheterisation was used for the diagnosis of HFpEF, intermediate if signs/symptoms of heart failure with left ventricular ejection fraction ≥40–50% and structural/functional abnormalities indicative of LVDD was used,\(^16,18–21\) and high for all other reference standards. Within the remaining domains the ROB was rated low, intermediate or high when respectively all, two, and one or none of the supporting questions (online supplementary Table S1; three pre-defined questions per domain) were answered in a positive manner. However, certain study characteristics—no avoidance of case-control/two-gated designs, or unclear/inappropriate timing for the index test and/or reference standard—would immediately lead to a high ROB for the respective domain. Inconsistencies in quality assessment between the four reviewers were resolved by discussion until consensus was reached.

Results

Search results

A total of 20,757 studies were derived from the extensive literature search. A total of 28 studies were deemed eligible for this review (online supplementary Figure S1). The 28 selected studies included a wide range of potential novel diagnostic HFpEF circulating biomarkers (Table 1).\(^22–49\)
### Table 1 Overview of the diagnostic heart failure with preserved ejection fraction circulating biomarker studies

| Study/country | Biomarkers | Cases (reference standard) | Controls | Cases/controls descriptives |
|---------------|------------|----------------------------|----------|-----------------------------|
| Martos, 2009[2] | CITP; MMP-1,-2,-9; PICP; PNP; RINP; TIMP | HFpEF (n = 32) | No HFpEF (n = 53) | **Age (years)** | **Sex (% female)** | **NT-proBNP a (pg/mL)** | **LVEF (%)** |
| Ireland       |            | 72 ± 11/66 ± 9 | 47/75 | 265 ± 182/98 ± 132 | BNP | 63 ± 14/87 ± 10[3] | **E/e’** | **LAVI[4]** | **LVM[5]** |
|               |            |                |       |                |     |   |   |   |
| Stahrenberg, 2010[23] | GDF-15 | HFnE fresc (n = 142) | Healthy controls (n = 188) | 73 [66–78]/56 [52–63] | 64/66 | 326 [133–634]/64 [39–112] | 60 [56–65]/61 [56–66[6] | 12 [9–15]/7 [6–9] |
| Germany       |            |                |       |                |     |   |   |   |
|               |            |                |       |                |     |   |   |   |
| Zile, 2011[24] | CITP; CTP; MMP-1,-2,-3,-7,-8,-9; osteopontin; PNP; PIIINP; sRAGE; TIMP-1,-2,-3,-4 | LVH with DHF (n = 61) | LHV, no DHF (n = 144) | 66 ± 1/60 ± 1 | 59/55 | 214 ± 34/109 ± 12 | 69 ± 1/69 ± 1[7] | **E/e’** | **LAVI[4]** | **LVM[5]** |
| America       |            |                |       |                |     |   |   |   |
|               |            |                |       |                |     |   |   |   |
| Celic, 2012[25] | RDW | DHF (n = 71) | No signs/symptoms of HF (n = 50) | 57 ± 7/56 ± 7 | 63/58 | 97 [57–26]/65 [26–94] | 72 [63–75]/68 [63–73[8] | 9 ± 3/6 ± 2[9] | 103 ± 24/91 ± 20[10] |
| Turkey        |            |                |       |                |     |   |   |   |
|               |            |                |       |                |     |   |   |   |
| Santhasankrithan, 2012[26] | GDF-15; sST2; hsTnT | HFpEF (n = 50) | No history of CAD/HF (n = 50) | 69 ± 12/63 ± 8 | 42/54 | 942 [309–276]/69 [41–102] | 60 ± 7/66 ± 3[11] | 18 ± 9/9 ± 2[12] | **E/e’** | **LAVI[4]** | **LVM[5]** |
| Singapore     |            |                |       |                |     |   |   |   |
|               |            |                |       |                |     |   |   |   |

© 2020 The Authors. European Journal of Heart Failure published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.
| Study/country     | Biomarkers                           | Cases (reference standard) | Controls | Cases/controls descriptives                        |
|------------------|--------------------------------------|-----------------------------|----------|---------------------------------------------------|
|                  |                                      |                             | Age (years) | Sex (% female) | NT-proBNP* (pg/mL) | LVEF (%)<sup>5</sup> |
|                  |                                      |                             |           |                       | E/e<sup>6</sup> | LAVI<sup>7</sup> | LVM<sup>8</sup> |
| Baessler, 2012<sup>27</sup> Germany | GDF-15 | LVDD with possible HF (n = 88) | No LVDD (n = 119) | 50 ± 7/41 ± 12 | 55/73 | 52 [29–96]/42 [25–66] | 64 ± 9/64 ± 7<sup>4</sup> | 8 ± 3/5 ± 1<sup>4</sup> | 136 ± 32/102 ± 20 |
| Mason, 2013<sup>28</sup> England | Copeptin; hsCRP; MR-proANP; MR-proADM | HFpEF (n = 57) | No HF (n = 308) | 87 ± 6/84 ± 7 | 83/73 | 1300 ± 160/764 ± 1280 |                     |                     |                     |
| Wang, 2013<sup>29</sup> China | sST2 | HFpEF (n = 68) | No symptoms/signs HF (n = 39) | 68 ± 10/60 ± 12 | 54/33 | 262 ± 470/71 ± 53 | 68 ± 7/68 ± 7<sup>4</sup> | 12 ± 4/6 ± 1<sup>4</sup> |                     |
| Jang, 2014<sup>30</sup> China | Angiogenin | HFpEF (n = 16) | Healthy controls (n = 16) | 76 ± 4/68 ± 8 | 62/38 | 3377 [2178–3995]/55 [27–93] | 55 ± 12/70 ± 4<sup>4</sup> |                     |                     |
| Wong, 2015<sup>31</sup> Singapore | Miscellaneous miRNAs | HFpEF (n = 30) | No history of CAD/HF (n = 30) | 64 ± 9/66 ± 7 | – | 1712 (± 263)/86 (± 83) | 59 ± 5/64 ± 4<sup>4</sup> |                     |                     |
|                   |                                      | HFpEF ≤ 40% (n = 30) | 64 ± 9/65 ± 7 | – | 1712 (± 263)/6727 (± 6290) | 59 ± 5/25 ± 7<sup>4</sup> |                     |                     |
| Zordoky, 2015<sup>32</sup> Canada | Miscellaneous metabolites | HFpEF (n = 24) | Healthy controls and patients at risk (n = 38) | 68 [58–75]/64 [54–69] | 25/30 | 110 ± 140/238 ± 294 |                     |                     |                     |
|                   |                                      | HFpEF <45% (n = 20) | 68 [58–75]/64 [56–69] | 25/30 | 110 ± 140/238 ± 294 |                     |                     |                     |

* NT-proBNP: N-terminal pro-B-type natriuretic peptide, LVEF: left ventricular ejection fraction, E/e: ratio of early (E) to late (e) diastolic mitral inflow velocities, LAVI: left atrial volumetric index, LVM: left ventricular mass index.

<sup>1</sup> Data from Baessler et al., 2012.
<sup>2</sup> Data from Mason et al., 2013.
<sup>3</sup> Data from Wang et al., 2013.
<sup>4</sup> Data from Jang et al., 2014.
<sup>5</sup> Data from Wong et al., 2015.
<sup>6</sup> Data from Zordoky et al., 2015.
| Study/country      | Biomarkers                          | Cases (reference standard) | Controls                                | Cases/controls descriptives |
|--------------------|-------------------------------------|----------------------------|----------------------------------------|----------------------------|
|                     |                                     | Age (years) | Sex (% female) | NT-proBNP (pg/mL) | LVEF (%) |
| Watson, 2015 & Ireland | Miscellaneous miRNAs               | HFrEF (n = 75) | 75 ± 7/10 ± 11 | 39/27 | 215 [126–353]/139 [71–256] |
|                     |                                     | HFrEF <50% (n = 75) | 52 ± 13/40 ± 58 | 52 ± 13/46 ± 14 | 114 ± 36/1 26 ± 38 |
| Sanders-van Wijk, 2015 & Switzerland and Germany | Cys-C; Hb; hsCRP; hsTnT; sST2 | HFrEF (n = 112) | 80 ± 7/16 ± 7 | 64/33 | 2142 [1473–429]/4202 |
|                     |                                     | HFrEF ≤40% (n = 458) | 57 ± 6/29 ± 7 |
| Barroso, 2016 & Germany | IGFBP-7; IGF-1            | HFrEF (n = 77) | 73 [68–77]/54 | 60/47 | 344 [152–703]/90 |
|                     |                                     | No LVDD, LVEF >50% (n = 55) | 57 ± 6/29 ± 7 |
| Liu, 2016 & China | sgp130; hsP27; CTSS; DPP4         | HFrEF (n = 50) | 64 ± 6/64 ± 6 | 46/54 | 982 ± 46/33 ± 227 |
|                     |                                     | No history of heart disease(s) (n = 50) | 57 ± 6/29 ± 7 |
| Polat, 2016 & Turkey | Gal-3                              | HFrEF (n = 44) | 60 ± 7/15 ± 9 | 46/47 | 618 ± 271/66 ± 54 |
|                     |                                     | No systolic/diastolic dysfunction (n = 38) | 57 ± 6/29 ± 7 |
| Li, 2016 & China | AqJ-Ca                              | HFrEF (n = 104) | 76 ± 9/68 ± 12 | 54/41 | 645 ± 264/192 ± 70 |
|                     |                                     | No HFrEF (n = 701) | 57 ± 6/29 ± 7 |
| Berezin, 2016 & Ukraine | CD31+ annexin V+ EMPs to CD14+CD309+ cell ratio | HFrEF (N = 79) | 55 ± 7/58 ± 7 | 53/42 | 2131 [935–305]/2774 [1520–3870] |
|                     |                                     | HFrEF ≤45% (n = 85) | 55 ± 7/58 ± 7 |

[^1]: LVEF (%)
[^2]: E/e'<E/e
[^3]: LAVI '<LVM
[^4]: LVEF ≥50%
[^5]: LVEF ≥50%
[^6]: NT-proBNP ≥2x ULN
[^7]: NT-proBNP ≥2x ULN
[^8]: NT-proBNP ≥2x ULN
[^9]: NT-proBNP ≥2x ULN
[^10]: NT-proBNP ≥2x ULN
[^11]: NT-proBNP ≥2x ULN
[^12]: NT-proBNP ≥2x ULN
[^13]: NT-proBNP ≥2x ULN
| Study/country    | Biomarkers                                                | Cases (reference standard) |
|-----------------|-----------------------------------------------------------|----------------------------|
| Toma, 2017<sup>20</sup> Canada | Misc proteins and transcripts | HFpEF (n=21) ♦ Symptoms consistent with HF ♦ LVEF ≥50% |
|                 |                                                            | Controls                   |
|                 |                                                            | Age (years)                | Sex (% female) | NT-proBNP<sup>+</sup> (pg/mL) | LVEF (%) ▲ E/e ▼ LAVI ▽ LVMI ▽ |
| Toma, 2017<sup>20</sup> Canada |                                                            | 70 [63–79]/66              | 52/27          | 295 [143–150]/1174            | 60 [56–62]/30 [23–36]<sup>▼</sup> |
| Sinning, 2017<sup>21</sup> Germany | GDF-15; sST2; CRP | HFpEF (n=70) ♦ NYHA II–IV or treatment for HF ♦ LVEF ≥50% ♦ LVDD |
|                 |                                                            | Controls                   |
| Sinning, 2017<sup>21</sup> Germany |                                                            | 67 [62–72]/64              | 50/21          | 146 [76–294]/956              | 64 [59–70]/43 [36–48]<sup>▼</sup> |
| Cui, 2018<sup>12</sup> China | Gal-3; sST2 | HFpEF (n=172) ♦ HFpEF ESC, 2016<sup>13</sup> |
| Cui, 2018<sup>12</sup> China |                                                            | Controls                   |
| Cui, 2018<sup>12</sup> China |                                                            | 73 ± 9.7 ± 9               | 56/39          | 614 [243–1479]/4330           | 60 [56–62]/31 [28–35]<sup>▼</sup> |
| Nikolova, 2018<sup>23</sup> America | cBIN1 | HFpEF (n=52) ♦ History of fluid overload, prior HFH, or invasive evidence of elevated cardiac filling pressures ♦ LVEF ≥50% |
| Nikolova, 2018<sup>23</sup> America |                                                            | Controls at risk           |
| Nikolova, 2018<sup>23</sup> America |                                                            | 57 ± 15.5 ± 6              | 37/37          | 277 [99–1264]/36              | 58 ± 7<sup>▼</sup> |

© 2020 The Authors. European Journal of Heart Failure published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.
| Study/country | Biomarkers | Cases (reference standard) | Controls | Cases/controls descriptives |
|---------------|------------|-----------------------------|----------|-----------------------------|
|               |            |                             |          | Age (years) | Sex (% female) | NT-proBNP* (pg/mL) | LVEF (%)³ | E/e⁶ | LAVI⁸ | LVMI⁹ |
| Farinacci, 2019 | CECs       | HFpEF (n = 27)              | Healthy Controls (n = 10) | 69 ± 8/56 ± 3 | 44/55 | – | – | – | – |
| Germany       |            | • NYHA I–III                |          |             |                |                |                |                |                |                |
|               |            | • HFH during last year     |          |             |                |                |                |                |                |                |
|               |            | • Cardiac functional/structural abnormalities suggestive for HFpEF or elevated NP levels | | | | | | | | |
| Wong, 2019 | Miscellaneous miRNAs | HFpEF (n = 179) | HFpEF ≤40% (n = 145) | 77 ± 9/70 ± 14 | 46/17 | 2557 ± 2490/4898 ± 7887 | 62 ± 7/29 ± 7⁵ | – | – | – | – |
| Singapore and New Zealand | | | | | | | | | | |
| Chi, 2019 | CTGF, TGF-β1 | DHF (n = 114) | No HF (n = 72) | 71 ± 1/169 ± 11 | 53/43 | 1224 [499–2473]/70 [25–126] | 62 ± 9/67 ± 6⁵ | 13 ± 6/10 | – | – | – |
| China         | | | | | | | | | | |
| Berezin, 2019 | CD31+/annexin V+ MVs; GDF-15 | HFpEF (n = 178) | HFmrEF/HFrEF (n = 210) | 55 ± 7/57 ± 7 | 57/40 | 2131 [955–3056]/HFmrEF 2701 [1590–3541]; HFrEF 2375 [1520–3870] | 55 [51–58]/HFmrEF 44 [41–48]; HFrEF 37 [31–39]⁵ | – | – | – | – |
| Ukraine       | | | | | | | | | | |
| Fang, 2019 | RDW | HFpEF (n = 62) | I. No substantial cardiac dysfunction (n = 107) | 74 ± 9/67 ± 12 | 45/48 | 1095 [575–2027] | 58 ± 7/60 ± 6⁷ | 14 ± 5/13 ± 4⁶ | – | – | – |
| China         | | | II. Possible HFpEF | | | | | | | | |

© 2020 The Authors. European Journal of Heart Failure published by John Wiley & Sons Ltd on behalf of European Society of Cardiology.
### Quality assessment

All papers had at least one domain with a high ROB, and 11 papers (39%) showed a high ROB within all four domains (online supplementary Table S2). Main reasons for bias within the QUADAS-2 domains of each individual article are shown in online supplementary Table S3.

The ROB within the patient selection domain was high in 24 out of 28 studies (86%; Figure 7). This was mainly driven by the use of case-control/two-gated designs. Additionally, in 13 studies inappropriate exclusion criteria were not avoided (online supplementary Tables S3 and S4). This was often the result of excluding difficult to diagnose patients—e.g. patients with AF, obesity, and/or pulmonary diseases—or by excluding patient conditions which could possibly influence the outcome of the index test (e.g. kidney function disorders). Only nine studies (32%) did not use a case-control design, in only two of these studies inappropriate exclusion criteria were avoided (online supplementary Table S3).

Even though the index tests of all studies were classified as objective, the ROB for the index test domain was rated high in 26 out of 28 studies (93%; Figure 7). This was caused by the fact that most studies did not use pre-specified cut-off values and did not perform any external validation. Only one article provided information about the sensitivity and specificity of a pre-specified cut-off value for the index test studied, and one article performed validation of their findings in an external cohort.

All studies suffered from an intermediate or high ROB within the reference standard domain, being rated as intermediate/high in 14 out of 28 studies (50%; Figure 7). Different reference standards and definitions of LVDD were used, and none of the studies performed (exercise) right-sided heart catheterisation in all study subjects (Table 1).

A total of 27 out of 28 studies (96%; Figure 7) scored a high ROB within the flow and timing domain. In all these studies this was caused by the fact that the exact timing of the index test and/or reference standard was unclear (online supplementary Table S3).

Given the high ROB, combined with limited overlap in investigated biomarkers and different statistical methods used, no areas under the receiver operating curve were reported and no meta-analysis was performed.

### Discussion

This is the first study that provides a comprehensive overview of studies that included diagnostic evaluation of novel circulating biomarkers for the detection of HfPEF. All included studies in this review contributed to our current level of knowledge of this complex syndrome. However, this systematic review exposes multiple study limitations that together limit our ability to evaluate the true diagnostic value of circulating biomarkers. The main limitations that we found were: (i) use of case-control/two-gated designs; (ii) exclusion of a relevant/representative subset of the true HfPEF population; (iii) use of optimal rather than pre-specified cut-off points for the index test without the performance of external validation; (iv) inadequate and highly variable reference standards, none including the true gold standard; and (v) unknown...
Risk of bias in studies investigating novel diagnostic HFpEF biomarkers

Figure 1 Percentage of studies with low, intermediate or high risk of bias within the four QUADAS-2 domains (patient selection, index test, reference standard, flow and timing) and the main reasons for a high risk of bias within these domains.

timing of the index and/or reference standard. The overall high ROB might play an important role in the limited uptake of these biomarkers in the HFpEF clinics and calls for methodologically well-designed studies.50–51

Patient selection
Most studies determined the diagnostic value of the biomarkers in cases with known HFpEF compared to (healthy) controls. During the early stages of novel biomarker discovery, these designs with contrasting populations can be useful to screen whether novel biomarkers might be of any interest for future analysis.52 Such studies may also reveal mechanistic insights into the syndrome. However, for diagnostic utility these designs induce spectrum bias, which overestimates the diagnostic value of the investigated biomarker(s).52–55

Additionally, extensive exclusion criteria including AF, pulmonary diseases, or even chronic kidney function disorders were often used, which are all highly prevalent comorbid conditions in HFpEF.56–58 For example, over 50% of HFpEF patients have AF.59–61 Excluding these patients introduces selection bias that could result in a serious misinterpretation of the diagnostic value and reduce external validity of these biomarkers in unselected HFpEF populations.52,54,62

Index test
The use of optimal cut-off values for the index test without performing external validation within the majority of previous studies will have resulted in an overestimation of the diagnostic performance of the biomarkers examined.63 Moreover, a biomarker should have incremental value on top of easy to determine characteristics—e.g. age, sex and body mass index—to really yield potential for clinical use. While this was not part of the ROB assessment within this study, it will partially explain the lack of the implementation of novel diagnostic HFpEF biomarkers.

Reference standard
Test accuracy of a novel biomarker is based on the concept that every inconsistency between the index test and reference standard is due to an incorrect index test.17,51 Since different reference standards will significantly alter the prevalence of cases within the cohort of interest—as already shown within the field of LVDD64—this will significantly affect the diagnostic value of the biomarker(s) studied. None of the included studies used (exercise) right-sided heart catheterisation—the real gold standard for HFpEF—as uniform reference standard. Studies validating the biomarker value against this gold standard are urgently needed.

Recognising the challenges of widespread implementation of gold standard invasive haemodynamic testing, we also examined the use of guideline-recommended reference standards that were published at the moment of publication for the diagnosis of heart failure with normal ejection fraction since 200718 or HFpEF since 2016,19 and found that most studies did not apply these. Also, these reference standards were not in line with the recently published H2FPEF59 or HFA-PEFF scores.10 Nonetheless, even...
the recommended reference standards and risk scores differ significantly in included diagnostic criteria, used cut-off values and the role comorbidities play within these standards, highlighting the uncertainty of diagnosing HFpEF.

**Flow and timing**
Most studies did not provide (detailed) information regarding the timing of the index test and the reference standard. This lack of information is regrettable given that biomarker levels will likely change over time. Moreover, it is highly likely that diuretics are prescribed and/or dosage were changed in patients with signs of congestion. Diuretics will reduce filling pressure and very likely influence the concentration of the circulating biomarker measured. It has already been shown that diuretics affect the urinary proteome in rats, and the pleural protein concentration in patients with congestive heart failure. In the latter also an increase in total serum protein content after the administration of diuretics was observed. Therefore, it is highly desirable that the circulating biomarkers are measured at the same moment as the HFpEF diagnosis is made and before any intervention occurs.

**Phenotype specific biomarkers**
The question remains to which extent the absence of novel diagnostic HFpEF biomarkers is due to the real lack of diagnostic value of these biomarkers, vs. the heterogeneity of the syndrome itself. In contrast to HFrEF, heart failure with reduced ejection fraction, characterised by cardiomyocyte loss and ventricular dilatation, is diagnostically well-captured by natriuretic peptides that increase in response to wall stress and by troponins indicating cardiomyocyte injury. In the more heterogeneous HFpEF syndrome, biomarkers likely reflect less well the complex, mainly non-cardiac multi-organ nature of the syndrome. Therefore, biomarkers reflecting more general pathophysiological processes like inflammation (growth differentiation factor-15), fibrosis (soluble ST2, galectin-3), and metabolic dysfunction (insulin-like growth factor binding protein-7) could have potential; moreover, the search for one single biomarker may not be sufficient. An approach with multiple biomarkers in methodologically well-designed studies may be more appropriate and successful. One may postulate if it will ever be possible to find a single diagnostic test or panel of biomarkers with adequate diagnostic value for the entire syndrome, and perhaps the optimal approach may be to use specific biomarkers to diagnose distinct subtypes of HFpEF, which could eventually also lead to a more tailored therapy.

**Future perspectives**
There is an urgent need for prospective studies to validate the diagnostic value of the HFA-PEFF score against gold standard invasive exercise haemodynamic testing in unselected symptomatic patients with suspected HFpEF. The inclusion of blood biomarker testing in such a study will enable the evaluation of the possible role of novel biomarkers in the HFA-PEFF algorithm on top of NPs and echocardiographic biomarkers. Possibilities that warrant investigation include implementation of biomarker testing in step 1 (pre-test assessment) or step 2 (diagnostic work-up) of the HFA-PEFF algorithm. Furthermore, promising novel biomarkers may be assessed as potential alternatives to NPs. NP levels should not be used as a selection criterium in these studies since 18% to 30% of patients with haemodynamically proven HFpEF have NP levels below ‘diagnostic’ threshold. Such studies will require close collaboration between basic scientists, clinicians, epidemiologists, industry, and (federal) sponsors.

**Study limitations**
Although all papers were reviewed and discussed by our interdisciplinary team until consensus was reached, the ROB classifications are based on the information provided in the studies, the pre-defined risk of bias criteria, as well as on the interpretation of the reviewers themselves. Therefore, it is possible that analysis of the studies by another group of reviewers results in another level of bias within certain domains of studies. However, we defined clear roles and results are rather uniform and unambiguous, making it highly unlikely that the main conclusion would differ significantly. Our review did not aim for a head to head comparison between these studies, and therefore should not be used for this purpose.

To the best of our knowledge, this review includes all current novel diagnostic circulating biomarker studies to detect chronic HFpEF. However, given the extent of the search performed, it cannot be completely excluded that studies were missed if diagnostic performance measures were not mentioned in the abstract. Additionally, the main aim of some studies was not to study the diagnostic value of circulating biomarkers to detect HFpEF, though since they studied the diagnostic value in sub-analysis, they were still included in this review to provide a complete overview of current circulating diagnostic HFpEF biomarker analysis.

Finally, since some studies included (previous) hospitalised patients and timing of the reference standard and the drawing of blood was often unclear, we may have unintentionally included acute HFpEF populations. Since this does not affect the main conclusion of this review, we decided not to exclude these studies.

**Conclusion**
The majority of current diagnostic HFpEF biomarker studies have a high ROB, reducing the reproducibility and the potential for clinical care. Methodological well-designed studies with a uniform reference diagnosis are urgently needed to determine the incremental value of circulating biomarkers for the diagnosis of HFpEF.

**Supplementary Information**
Additional supporting information may be found online in the Supporting Information section at the end of the article. Appendix S1. Search string for PubMed and EMBASE. Figure S1. PRISMA flow diagram of study selection.
Table S1. Predefined questions that were used for the risk of bias assessment.
Table S2. Overview risk of bias within the QUADAS-2 domains.
Table S3. Main determinants of level of bias within the QUADAS-2 domains for the articles included in this review.
Table S4. Overview of the patient population and exclusion criteria of the diagnostic HFpEF circulating biomarker studies included in this review.

Funding
This work was supported by the European Union Commission’s Horizon 2020, and IM2-CARDIATEAM [N°821508]. We acknowledge the support from the Netherlands Cardiovascular Research Initiative: an initiative with support of the Dutch Heart Foundation, CVON2015-RECONNECT, CVON2016-Early HFpEF and CVON 2017-ShePREDICTS. Additionally, J.W.v.B. is supported by a ZonMw VIDI grant.

Conflict of interest: none declared.

References
1. Dunlay SM, Roger VL, Redfield MM. Epidemiology of heart failure with preserved ejection fraction. Nat Rev Cardiol 2017;14:391–602.
2. Ziaei B, Fonarow GC. Epidemiology and aetiology of heart failure. Card Fail Rev 2002;1:368–378.
3. Lam CS, Rienstra M, Tay WT, Liu LC, Hummel YM, van der Meer P, de Boer RA, Bakker SJ, Heymans S, van Empel V, Schroen B, van der Harst P, van Veldhuisen DJ, Pezzali N, Pieske B, Dickstein K, Fraser AG, Brutsaert DL. 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.
4. Ibrahim NE, Januzzi JL Jr. Established and emerging roles of biomarkers in heart failure. Circ Res 2018;123:614–629.
5. McInnes MD, Mohler D, Thombs BD, McGrath TA, Bossuyt PM. PRISMA-DTA Group Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA 2018;319:388–396.
6. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–536.
7. Meijers WC, Hoekstra T, Jaarsma T, van Veldhuisen DJ, de Boer RA. Patients with heart failure with preserved ejection fraction and low levels of natriuretic peptides. Neth Heart J 2016;24:287–295.
8. Brunner-La Rocca HP, Sanders-van Wijk S. Natriuretic peptides in chronic heart failure. Curr Cardiol Rev 2012;8:265–278.
9. Anjan VY, Loftus TM, Burke MA, Akhter N, Fonarow GC, Gheorghiade M, Shah AB, Trott KL, Paulus WJ, Tschope C, Flachskampf FA, Ramakers FE, Marino P, Smiseth OA, De Keulenaer G, Leite-Moreira AF, Borbely A, Eides I, Handoko ML, Heymans S, Pezzali N, Pieske B, Dickstein K, Fraser AG, Brutsaert DL. 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.
fraction: a novel potential biomarker for diagnosis and prognosis. J Cell Mol Med 2014;18:2189–2197.
31. Weng LL, Armugam A, Sepramanian S, Karolina DS, Lim KY, Lim JY, Chong JP, Ng JY, Chen YT, Chan MM, Chen Z, Yeo PS, Ng TP, Ling LH, Sim D, Leong KT, Ong HY, Jafferthier F, Wong R, Chai P, Low AF, Lam CS, Jeyaseelan K, Rahbari AH. Circulating miRNAs in heart failure with preserved and reduced left ventricular ejection fraction. Eur J Heart Fail 2015;17:393–404.
32. Zordory BN, Sung MM, Erezkowitz J, Mandal R, Han B, Bjornsdal TC, Bouatra S, Anderson T, Oudit GY, Washir DS, Dyck JR, Alberti H. Metabolomic fingerprint of heart failure with preserved ejection fraction. PLoS One 2015; 10:e0124466.
33. Watson CJ, Gupta SK, O’Connell E, Thum S, Glezeva N, Fendrich J, Gallagher J, Ledwidge M, Grote-Levi L, McDonald K, Thum T. MicroRNA signatures differentiate preserved from reduced ejection fraction heart failure. Eur J Heart Fail 2015;17:405–415.
34. Sanders-van Wijk S, van Empel V, Davarzani M, Maeder MT, Handschin R, Pflister ME. Brunner-La Rocca HP. TIME-CHF Investigators. Circulating biomarkers of distinct pathophysiologic pathways in heart failure with preserved vs. reduced left ventricular ejection fraction. Eur J Heart Fail 2015;17:1006–1014.
35. Barroso MC, Kramer F, Greene SJ, Scheyer D, Kahler T, Karoff M, Seyfart M, Gheorgheida M, Dinh W. Serum insulin-like growth factor-1 and its binding protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. BMC Cardiovasc Disord 2016;16:199.
36. Liu S, Iskandar R, Chen W, Zhang J, Wang Y, Chen X, Xiang F. Soluble protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. Eur J Heart Fail 2020;
37. Nikolova AP, Hitzeman TC, Baum R, Caldaruse AM, Agvanian S, Xie Y, Geft DR, Wu AH, Kobashigawa JA, Hamilton M, Hong T, Shaw RM. Association of a novel diagnostic biomarker, the plasma cardiac bridging integrator 1, with preserved ejection fraction. Eur J Heart Fail 2016;18:1039–1046.
38. Chang DH, Moriguchi JD, Hage A, Azarbal B, Czer LS, Kittleson MM, Patel JK, Butt SJ, Wilson-McManus JE, Ignaszewski A, Anderson T, Dyck JR, Howlett J, Ezekowitz J, McManus BM, Oudit GY. Differentiating heart failure phenotypes - distinction of chronic heart failure with reduced from preserved ejection fraction. JACC Heart Fail 2017;5:200–206.
39. Streng KW, Hillegel HK, Anker SD, Cieland JG, Dickstein K, Filippatos G, Lam CS, Santema BT, Voors AA. Atrial fibrillation in heart failure: a common and deadly combination. JACC Heart Fail 2017;5:575–577.
40. Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, Grobbee DE, Moons KG, Reitsma JB, Zwinderman AH, Zannad F, Damman K, van der Meulen JH, van Helvoorden DJ, Zwinderman AH, Zannad F, Damman K, van der Meer P, Voors AA. Non-cardiac comorbidities in heart failure with reduced, mid-range and preserved ejection fraction. Int J Cardiol 2018;271:132–139.
41. Sinning C, Kempf T, Schwarzl M, Lanfermann S, Ojeda F, Schnabel RB, Zengin E, Wild PS, Lackner KJ, Munzel T, Blankenberg S, Wollert KC, Zeller T, Westermann D. Biomarkers for characterization of heart failure – distinction of heart failure with preserved and reduced left ventricular ejection fraction. EBM Medicine 2016;6:86–94.
42. Cui Y, Qi X, Huang A, Li J, Hou W, Liu K. Differential and predictive value of differential microRNAs in heart failure with non-reduced ejection fraction: diagnostic application and new cut-off points. Heart Lung 2020;49:388–392.
43. Barroso MC, Kramer F, Greene SJ, Scheyer D, Kahler T, Karoff M, Seyfart M, Gheorgheida M, Dinh W. Serum insulin-like growth factor-1 and its binding protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. BMC Cardiovasc Disord 2016;16:199.
44. Liu S, Iskandar R, Chen W, Zhang J, Wang Y, Chen X, Xiang F. Soluble protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. Eur J Heart Fail 2015;17:1006–1014.
45. Barroso MC, Kramer F, Greene SJ, Scheyer D, Kahler T, Karoff M, Seyfart M, Gheorgheida M, Dinh W. Serum insulin-like growth factor-1 and its binding protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. BMC Cardiovasc Disord 2016;16:199.
46. Liu S, Iskandar R, Chen W, Zhang J, Wang Y, Chen X, Xiang F. Soluble protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. Eur J Heart Fail 2015;17:1006–1014.
47. Barroso MC, Kramer F, Greene SJ, Scheyer D, Kahler T, Karoff M, Seyfart M, Gheorgheida M, Dinh W. Serum insulin-like growth factor-1 and its binding protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. BMC Cardiovasc Disord 2016;16:199.
48. Liu S, Iskandar R, Chen W, Zhang J, Wang Y, Chen X, Xiang F. Soluble protein-7: potential novel biomarkers for heart failure with preserved ejection fraction. Eur J Heart Fail 2015;17:1006–1014.
49. Merino-Merino A, Saez-Maleta R, Salgado-Aranda R, AlKassam-Martinez D, Pascual-Terjeira V, Martin-Gonzalez J, Garcia-Fernandez J, Perez-Rivera JA. Biomarkers in atrial fibrillation and heart failure with non-reduced ejection fraction: diagnostic application and new cut-off points. Heart Lung 2020;49:388–392.
50. Vasan RS. Biomarkers of cardiovascular disease: molecular basis and practical considerations. Circulation 2006;113:2335–2362.
51. Ahmad T, Fiuzaat M, Pencina MJ, Geller NL, Zannad F, Cleland JG, Archer J. Factors that limit the utility of biomarkers in heart failure with preserved ejection fraction. Eur J Heart Fail 2011;13:569–573.
52. Whiting P, Rutjes AW, Reitsma J, Glas AS, Bossuyt PM. Case-control and two-gate designs in diagnostic accuracy studies. Clin Chem 2005;51:1333–1341.
53. Whiting P, Rutjes AW, Reitsma J, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. Ann Intern Med 2004;140:199–202.
54. Whiting P, Rutjes AW, Westwood ME, Mallett S, QUADAS-2 Steering Group. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. J Clin Epidemiol 2013;66:1093–1104.
55. Lijmer JG, Moi BV, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, Bossuyt PM. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999;282:100;2016.
56. Paulus WJ, Tiohope C. A novel paradigm for heart failure with preserved ejection fraction: comorbidities drive myocardial dysfunction and remodeling through coronary microvascular endothelial inflammation. J Am Coll Cardiol 2013;62:263–271.
57. Streng KW, Nauta JF, Hillegel HK, Anker SD, Cieland JG, Dickstein K, Filippatos G, Lang CC, Metra M, Ng LL, Ponikowski P, Samani NJ, van Veldhusen DJ, Zwinderman AH, Zannad F, Damman K, van der Meer P, Voors AA. Non-cardiac comorbidities in heart failure with reduced, mid-range and preserved ejection fraction. Int J Cardiol 2018;271:132–139.
58. Ergoutsou C, Schaafelder M, Andorssons B, Pivodic A, Dahlstrom U, Fu M. Non-cardiac comorbidities and mortality in patients with heart failure with reduced vs. preserved ejection fraction: a study using the Swedish Heart Failure Registry. Clin Res Cardiol 2019;108:1025–1033.
59. 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.
60. Reddy YN, Obokata M, Gersh BJ, Borlaug BA. High prevalence of occult heart failure with preserved ejection fraction among patients with atrial fibrillation and dyspnea. Circulation 2018;137:534–535.
61. Lam CS, Santema BT, Voors AA. Atrial fibrillation in heart failure: a common and deadly combination. JACC Heart Fail 2017;5:575–577.
62. Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, Groobbe DE, Moons KG. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol 2003;56:826–832.
63. Leelathong MM, Moons KG, Reitsma JB, Zwinderman AH. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitudes, and solutions. Clin Genom 2008;5:729–737.
64. Hutton O, Fraser AG, Coozo S, Bozec E, Selton-Suty C, Lamiral Z, Frikha Z, Rossignon P, Zannad F, Girerd N. Impact of changes in consensus diagnostic recommendations on the echocardiographic prevalence of dystrophic failure. J Am Coll Cardiol 2017;69:3119–3121.
65. Li X, Zhao M, Li M, Jia L, Gao Y. Effects of three commonly-used diuretics on the urinary proteome. Genomics Proteomics Bioinformatics 2014;12:106–126.
66. Romero-Candeira S, Fernandez C, Martin C, Sanchez-Paya J, Hernandez L, Genomics Proteomics Bioinformatics 2014;12:106–126.