Open Peer Review

Any reports and responses or comments on the article can be found at the end of the article.

STUDY PROTOCOL

Diagnostic accuracy and clinical impact of natriuretic peptide screening for the detection of heart failure in the community: a protocol for systematic review and meta-analysis [version 2; peer review: awaiting peer review]

C. R. Goyder 1, A. K. Roalfe 1, N. R. Jones 1, K. S. Taylor 1, C. D. Plumptre 2, T. R. Fanshawe 1, F. D. R. Hobbs 1, C. J. Taylor 1

1Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, OX2 6GG, UK
2Nuffield Department of Medicine, University of Oxford, Oxford, OX3 9DU, UK

Abstract

Introduction: Patients diagnosed with heart failure in primary care have a better prognosis than those diagnosed in hospital. However, most cases are missed in the community. Recent attention has focussed on the potential of early detection through screening. Natriuretic peptides (NPs) are tested by GPs and used to rule out heart failure in patients presenting with symptoms. Evidence is now emerging that they may also have a role in screening but their accuracy in this context and the associated optimal thresholds, have not been established. The impact that NP screening would have on patients and health care systems also remains unclear.

Methods: We aim to undertake a systematic search of the following sources: Ovid Medline, Embase, Cochrane Database of Systematic Reviews and Cochrane Central Register of Controlled Trials. Screening, data extraction and critical appraisal will be carried out independently and in duplicate by two reviewers. We will include studies based in the community with >100 participants that recruited a screened population. We will not add a study design filter and there will be no language restriction. The primary outcome will be the sensitivity and the specificity of NP screening and optimal thresholds for screening will be explored. Outcomes of interest for the impact analysis will include mortality, hospital admissions and cost effectiveness. This protocol has been developed in accordance with guidelines from the preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P).

Discussion: This systematic review will identify how accurately NP screen for heart failure in the community and explore where NP screening thresholds should be set. It also aims to summarise the clinical impact of this strategy. Together, these results should inform future interventions that may provide an alternative pathway to facilitate improved detection of heart failure in the community.

Registration: PROSPERO CRD42018087498; registered on 11 May 2018.
Keywords
Heart Failure, Left ventricular systolic dysfunction, diastolic dysfunction, community screening, detection, natriuretic peptides, diagnostic accuracy
Introduction

Heart failure (HF) is a common and debilitating syndrome and presents a significant burden both to patients and society. Emerging evidence suggests that patients diagnosed with HF in primary care have a better prognosis than those diagnosed in acute settings\(^1\). However, detecting patients with HF in primary care is challenging\(^2\). Recent research highlights missed opportunities to diagnose HF in primary care\(^3\) and suggests that the diagnosis is most often made in hospitals despite many patients presenting earlier in primary care with suggestive symptoms. Some patients may also delay consulting their GP or community health care professional. For example, patients with breathlessness may modify their behaviour rather than consult a clinician, and older patients may normalise their symptoms, considering them to be secondary to ageing\(^7\). Patients who do present often describe vague and non-specific symptoms such as fatigue and breathlessness, which can have multiple causes. This can result in inappropriate alternative diagnoses, such as attributing HF to a respiratory condition resulting in ineffective treatment and management.

Given the difficulties in diagnosing HF in the community, recent research has focussed on early detection. Community screening is one potential strategy, especially in those individuals with cardiovascular risk factors\(^8\)\(^\text{–}\)\(^14\). In general, screening for HF meets many of the established criteria required of screening programmes\(^8\). It is clear that HF is an important health problem. Indeed, it is estimated that 40 million people worldwide are living with HF\(^10\). The estimated incidence rate rises from 1.4/1000 person-years in those aged 55–59 years to 47.4/1000 person-years in those aged 90 years or older\(^11\). HF is associated with a high morbidity and mortality. Patients with HF have a lower five-year survival compared to people with most common types of cancer\(^12\). Moreover, HF patients also have a ten-year survival of only 27%, compared to 75% for age-sex matched patients without HF\(^13\).

The natural history of disease progression in HF is well understood. An early, latent stage is recognised; this preclinical phase is termed left ventricular systolic dysfunction (LVSD) and can be asymptomatic\(^4\). Evidence suggests that identifying patients with LVSD early and commencing treatment with angiotensin-converting enzyme (ACE) inhibitors has been shown to reduce both incidence of HF and reduce rate of hospitalization\(^15\). Moreover, early treatment of high-risk individuals, even prior even to a diagnosis of LVSD, has been associated with reductions in the rate of HF associated hospitalization\(^16\). It is important to note that the prognosis and treatment of HF due to diastolic dysfunction (DD), more recently referred to as HF with preserved ejection fraction (HFP EF), is complex and currently there is less evidence to support early intervention in this group.

Natriuretic peptides (NPs) are tested by GPs and used to rule out HF in patients presenting with symptoms, so it is possible that NPs may perform well in the screening context. Indeed, a recent randomised control trial demonstrated that NP screening and associated collaborative care reduced the combined rates of systolic dysfunction, DD and HF as well as overall hospital admissions\(^17\). This study did not measure the impact of screening on patient’s quality of life and was also underpowered to measure any differences in mortality between control and intervention groups.

Screening with echocardiography alone is too expensive. However a screening strategy combining echocardiography with NP testing to stratify those at highest risk (who could then undergo echocardiography) would be more cost-effective\(^18\)\(^\text{–}\)\(^19\). Indeed, this recommendation has been included in the recent Canadian Cardiovascular Society Guidelines and the American College of Cardiology/American Heart Association/Heart Failure Society of America (ACCF/AHA/HFSA) guidelines although it does not yet feature in European guidelines\(^20\)\(^\text{–}\)\(^22\). Despite this, there have been no recent systematic reviews in this area. Previous reviews on population screening with NP combined patients from community settings with those recruited from secondary care\(^23\)\(^\text{–}\)\(^25\). Other systematic reviews have analysed the performance of NP to detect HF or LVSD in symptomatic, presenting patients\(^26\)\(^\text{–}\)\(^27\), rather than looking at asymptomatic patients or have combined screening with diagnostic studies\(^24\). There have been no systematic reviews to consider the overall impact of introducing NP screening for HF and to consider potential negative issues, such as overdiagnosis.

This systematic review will therefore provide an up-to-date summary of all the available evidence, focussing specifically on the diagnostic accuracy and clinical impact of NP screening for the detection of HF, LVSD and DD.

Research aims

The primary aim of this systematic review is to determine the diagnostic accuracy of NP screening for the detection of HF, LVSD and DD in the community. To clarify this question the following PIRT summarises the clinical question:

<table>
<thead>
<tr>
<th>Population: Patients in community-based NP screening studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index test: NP tests (point-of-care or laboratory-based)</td>
</tr>
<tr>
<td>Reference standard: Echocardiography or cardiac MRI</td>
</tr>
<tr>
<td>Target conditions: HF, LVSD, DD</td>
</tr>
</tbody>
</table>

This systematic review also aims to assess the impact that NP screening may have on both patients and health care systems.
and will therefore evaluate the clinical utility of this approach. The following PICO outlines the clinical question:

<table>
<thead>
<tr>
<th>Population: Patients in community-based NP screening studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention: NP test (point-of-care or laboratory-based)</td>
</tr>
<tr>
<td>Comparison: Usual care (no NP screening)</td>
</tr>
<tr>
<td>Outcomes: mortality, overall prognosis, quality of life, NYHA class at diagnosis, hospital admissions and cost effectiveness</td>
</tr>
</tbody>
</table>

**Methods**

**Protocol registration**

This systematic review protocol has been registered in the PROSPERO International Prospective Register of Systematic Reviews (registration number: CRD42018087498). This protocol has been developed according to recommendations from the Cochrane Collaboration. Checklists from the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement and the Joanna Briggs Institute have also been followed.

**Definitions**

In the literature, the term HF can have different meanings. For clarity, in this systematic review, the target condition HF refers to the European Society Cardiology (ESC) definition of HF that describes a clinical syndrome combining typical symptoms and signs plus evidence of structural or functional cardiac abnormalities, including elevated levels of NP.

The term LVSD has also been used to describe different scenarios. It can be used to define an early or precursor stage of HF when there is impairment of the left ventricle, but patients may not yet have developed the clinical syndrome of HF and are therefore asymptomatic. Clinical HF that is due to a reduction in left ventricular ejection fraction (LVEF) may also be called LVSD. Therefore, in more recent guidelines, the term LVSD is also used to describe HF with reduced ejection fraction (HFrEF) when the ejection fraction is below 40%. The ACCF/AHA classify HF in a spectrum from stage A through to D, where A refers to patients at risk of HF and stage D refers to patients with refractory clinical HF. Stage B HF refers to structural heart disease without signs or symptoms of HF and traditionally LVSD would have been included in this category. In order to capture these different scenarios, we are using HF, LVSD and DD as our target conditions, but for the purposes of the review we will use the definitions specified by the included studies and the data will be extracted so that test performance at each ejection fraction can be analysed.

**Eligibility criteria**

**Inclusion criteria**

**Study design.** We will not add a study design filter, and there will be no language restriction. To assess diagnostic accuracy, included diagnostic accuracy studies will follow cross-sectional and case-control designs. For impact analysis, we expect a broader range of study designs to be included, including randomised controlled trials and observational studies (retrospective or prospective) including cohort studies.

**Population.** We will include studies of adult patients (aged 18 years or older). Some screening studies may only recruit patients who do not have a prior diagnosis of LVSD or HF whereas other studies may adopt a more pragmatic approach to screening (noting that the prior diagnosis is often unreliable). We will include studies with both these recruitment approaches as long as they are based in community settings and recruit a screened population, in contrast to patients that present with HF symptoms. We will only include studies that combine these populations if they provide accuracy data independently for each group.

We will only include studies that included more than 100 participants to avoid introducing bias from studies with small sample sizes. Patients at high risk of HF, will be defined as those who had at least one cardiovascular risk factor such as hypertension or diabetes or conditions known to cause HF such as ischaemic heart disease. Selected non-general populations, such as cohorts of patients who all have chronic obstructive pulmonary disease (COPD), will also be included in the high risk group as the overlap between COPD and HF is known to be high.

**Intervention.** To assess diagnostic accuracy the included studies will compare NP measurement with either echocardiography or cardiac magnetic resonance imaging (MRI) for target conditions of HF and LVSD. For impact analysis, included studies will quantitatively measure the impact of introducing NP screening for HF, LVSD or DD. Studies that also compare NP screening with other strategies such as electrocardiogram (ECG) will be examined. Studies that employed multi-faceted interventions (such as NP screening plus collaborative cardiology care) will be included but will be analysed separately.

**Primary and secondary outcomes**

To evaluate diagnostic accuracy, the primary outcome will be the sensitivity and the specificity of NP screening. Where possible, the optimal threshold of NP to maximise screening performance will be explored, although this is not an individual patient data analysis so this assessment maybe will be limited by the available data. Outcomes of interest for the impact analysis will include NYHA class at diagnosis, mortality, overall prognosis, quality of life, hospital admissions and cost effectiveness.

**Exclusion criteria.**

The following studies will be excluded:

- All studies that recruit through secondary care.
- Studies containing duplicate datasets. We will select the papers that most closely align with inclusion criteria or were most recently published if they are otherwise equal.
- Conference abstracts, as these do not provide enough methodological detail to allow adequate analysis of risk of bias.
- Studies evaluating NP screening in restricted patient groups such as patients with rheumatoid arthritis, betathalassaemia, Marfan’s syndrome and Duchenne's Muscular dystrophy.
• Studies based on participants who consulted their GP or another community healthcare professional with symptoms of HF.

In evaluating diagnostic accuracy, specifically the following additional exclusion criteria will apply:
• Studies that assess NP as a prognostic marker.
• Studies that do not contain sufficient data to construct 2x2 tables, although we will contact all authors first to give the opportunity to provide missing data.

Search strategy
The search strategy was developed by CRG and the librarian/information specialist NR and subsequently refined to ensure that all appropriate papers were captured. The following databases will be searched: Ovid Medline (1946 to 2018), Embase (1974 to 2018), Cochrane Database of Systematic Reviews, Cochrane CENTRAL, DARE, Science Citation Index (1945–present). We will perform citation searches of all full-text papers included in final review. The full search strategy is available in Extended data, Appendix A.

Selection of studies
Each title/abstract will be screened by two reviewers independently and in duplicate. Any disagreements regarding inclusion will be resolved by a third reviewer. The screening process will be managed using Covidence software. Any potentially relevant articles will be selected and the full text obtained. Full-text papers will then be screened using the same process until the final studies are selected.

Data extraction
Data extraction will also be performed independently and in duplicate. A third author will compare extraction results to ensure that these are in agreement and highlight any areas of conflict so these can be resolved through discussion. If any ongoing studies are found, these will be described and the authors will be contacted to see if there are any relevant data that could be incorporated into review. Data extraction will be performed using a template. The minimum planned extraction fields are listed in Extended data, Appendix B. Each row of data will be coded so that data from specific populations (e.g. high risk compared to general), ages, outcomes and thresholds will be categorised and appropriate data combined at the data analysis stage.

Assessment of study quality
Two authors will independently assess risk of bias. The summary extraction fields in appendix B also incorporate a risk of bias template from the QUADAS-2 criteria, which will be used to assess methodological quality for diagnostic accuracy studies. Studies that assess clinical utility may be randomized trials, in this case the following domains will be assessed as per the Cochrane risk of bias tool: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, similarity of baseline measures, similarity of baseline characteristics, management of incomplete outcome data, selective outcome reporting, and other risks of bias. As per the recommendations of the GRADE working group, quality assessment will be used to interpret the quality of the evidence and the risk of bias present within the studies.

We will summarize the results of the risk of bias assessment across studies in a table.

Data synthesis
For the diagnostic accuracy analysis, bivariate meta-analysis will be used to calculate pooled performance estimates. Forest plots will be used to visually explore paired sensitivity and specificity of included studies and these will be generated in RevMan. Where studies measured sensitivity and specificity at multiple thresholds, the lowest threshold will be selected for inclusion in the forest plots. The most important role of community screening is to detect patients with the target condition and therefore to capture the highest sensitivity recorded by each study, it is for this reason that the lowest threshold was selected. If there are at least four studies with available data, then further analysis to create summary receiver operating characteristic (SROC) curves from pooled sensitivity and specificity at every specific threshold will be followed. This approach, outlined by Steinhauser et al., also enables the optimal threshold across studies to be determined. If there is insufficient data to generate SROCs with multiple thresholds, then hierarchical summary ROC curves will be drawn from the accuracy data provided on the lowest threshold per study data.

For the impact analysis, meta-analyses will be conducted separately for randomised controlled trials and cohort studies if the same outcome is reported from trials with similar contexts. RevMan software will be used to calculate and plot pooled effects. The differences between screening-related outcomes will be measured through comparison of relative risks (RRs) and mean differences (MDs). Random effects models will be used unless statistical and clinical heterogeneity are sufficiently low to warrant a fixed effects model. For these outcomes for which meta-analysis is not possible, we will construct evidence tables to report results descriptively. Statistical heterogeneity will be analysed using the I² statistic.

Data will also be reported on the search statistics including the number of references in the original search, and those included at full text stage and in the final review. The number of excluded papers will be recorded with the reasons for exclusion being noted. Included study characteristics will be summarised in Table 1 of the full systematic review.

Subgroup analysis
Where possible, subgroup analyses will be performed. These will include looking at different types of NP, including plasma brain natriuretic peptide (BNP) and N-terminal prohormone (NT-proBNP) as well as considering point-of-care (POCT) and laboratory tests. We will also consider factors which may affect NP measurement including participant:
• Demographics (age, sex)
• Body mass index (BMI)
• Renal function

Different populations of interest will also be included to compare screening in high risk populations (as defined above) and in general populations. Where possible, we also aim to investigate whether there is any difference in NP performance in asymptomatic participants compared to those in whom symptoms were described. We could also aim to explore whether the performance of NP changes with severity of LVSD/DD.

Sensitivity analysis
If studies with high or unclear risks of bias are found, we will use sensitivity analyses to explore the effect of excluding these studies from the data analysis.

For the analysis of diagnostic accuracy, sensitivity analysis will evaluate NP screening performance in studies that exclusively recruited participants who did not have a prior diagnosis of HF, LVSD or DD in comparison with those studies that pragmatically recruited a screened population.

Discussion
Identifying patients with HF in the community is crucial. Treatment improves patient survival and quality of life, and early diagnosis can prevent disease progression. A potential strategy to screen patients for HF by measuring NP presents an alternative diagnostic pathway, although evidence underpinning this is lacking. This systematic review aims to change this by providing an up to date summary of the literature on both the diagnostic accuracy and the impact of NP screening. Dissemination of results will be through publication in peer-reviewed journals and the presentation of this work at relevant conferences. We also plan to share our results with the public through organisations such as the British Heart Foundation.

Data availability
Underlying data
No underlying data are associated with this article.

Extended data
Figshare: Appendix A.docx. https://doi.org/10.6084/m9.figshare.10062191.v1

This file contains the search strategy for the systematic review.

Figshare: Appendix B.docx. https://doi.org/10.6084/m9.figshare.10062215.v1

This file contains an outline of the data extraction plan.

Reporting guidelines

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

References

16. Huisman M, Neuhold S, Resl M, et al.: PONTIAC (NT-proBNP selected...


