

The REFER (REFer for Echocardiogram) study: a prospective validation and health economic analysis of a clinical decision rule, NT-proBNP or their combination in the diagnosis of heart failure in primary care

*Clare J Taylor, Mark Monahan, Andrea K Roalfe, Pelham Barton, Rachel Iles
and FD Richard Hobbs on behalf of the other REFER investigators*



***National Institute for
Health Research***

The REFER (REFer for EchocaRdiogram) study: a prospective validation and health economic analysis of a clinical decision rule, NT-proBNP or their combination in the diagnosis of heart failure in primary care

Clare J Taylor,¹ Mark Monahan,² Andrea K Roalfe,² Pelham Barton,² Rachel Iles² and FD Richard Hobbs^{1*} on behalf of the other REFER investigators

¹Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

²Institute of Applied Health Research, University of Birmingham, Birmingham, UK

*Corresponding author

Declared competing interests of authors: FD Richard Hobbs has received grants from Roche Diagnostics outside the submitted work. The work was supported by funding from the National Institute for Health Research Efficacy and Mechanism Evaluation programme. Roche Diagnostics provided the N-terminal pro-B-type natriuretic peptide testing equipment but did not have any influence on study design, conduct or reporting.

Published April 2017

DOI: 10.3310/eme04030

This report should be referenced as follows:

Taylor CJ, Monahan M, Roalfe AK, Barton P, Iles R, Hobbs FDR on behalf of the other REFER investigators. The REFER (REFer for EchocaRdiogram) study: a prospective validation and health economic analysis of a clinical decision rule, NT-proBNP or their combination in the diagnosis of heart failure in primary care. *Efficacy Mech Eval* 2017;**4**(3).

Efficacy and Mechanism Evaluation

ISSN 2050-4365 (Print)

ISSN 2050-4373 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full EME archive is freely available to view online at www.journalslibrary.nihr.ac.uk/eme. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Efficacy and Mechanism Evaluation* journal

Reports are published in *Efficacy and Mechanism Evaluation* (EME) if (1) they have resulted from work for the EME programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

EME programme

The Efficacy and Mechanism Evaluation (EME) programme was set up in 2008 as part of the National Institute for Health Research (NIHR) and the Medical Research Council (MRC) coordinated strategy for clinical trials. The EME programme is broadly aimed at supporting 'science driven' studies with an expectation of substantial health gain and aims to support excellent clinical science with an ultimate view to improving health or patient care.

Its remit includes evaluations of new treatments, including therapeutics (small molecule and biologic), psychological interventions, public health, diagnostics and medical devices. Treatments or interventions intended to prevent disease are also included.

The EME programme supports laboratory based or similar studies that are embedded within the main study if relevant to the remit of the EME programme. Studies that use validated surrogate markers as indicators of health outcome are also considered.

For more information about the EME programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/eme>

This report

The research reported in this issue of the journal was funded by the EME programme as project number 09/160/13. The contractual start date was in July 2011. The final report began editorial review in August 2015 and was accepted for publication in July 2016. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The EME editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, NETSCC, the EME programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the EME programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2017. This work was produced by Taylor *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Efficacy and Mechanism Evaluation Editor-in-Chief

Professor David Crossman Bute Professor of Medicine and Dean and Head of Faculty of Medicine, University of St Andrews, and Honorary Consultant Cardiologist, NHS Fife Health Board, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the EME Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research Group, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Director, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

The REFER (REFer for EchocaRdiogram) study: a prospective validation and health economic analysis of a clinical decision rule, NT-proBNP or their combination in the diagnosis of heart failure in primary care

Clare J Taylor,¹ Mark Monahan,² Andrea K Roalfe,²
Pelham Barton,² Rachel Iles² and FD Richard Hobbs^{1*}
on behalf of the other REFER investigators

¹Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

²Institute of Applied Health Research, University of Birmingham, Birmingham, UK

*Corresponding author richard.hobbs@phc.ox.ac.uk

Background: Heart failure is a treatable condition but making a diagnosis can be challenging.

Objective: To evaluate the performance of a clinical decision rule (CDR) with or without a natriuretic peptide assay for identifying heart failure in symptomatic patients presenting to primary care.

Design: Prospective, observational, diagnostic validation study and economic evaluation.

Setting: Twenty-eight general practices in central England, UK.

Participants: Primary care patients aged ≥ 55 years presenting with recent new-onset shortness of breath, lethargy or peripheral ankle oedema of > 48 hours' duration.

Instrument: The CDR included a clinical element (male, history of myocardial infarction, crepitations at the lung bases and oedema) and N-terminal pro-B-type natriuretic peptide (NT-proBNP) blood test. The reference standard was an expert consensus panel of three cardiology specialists.

Main outcome measure: The main outcome measure was test performance of the CDR and the natriuretic peptide test alone, and in combination, in estimating sensitivity and specificity, positive predictive value (PPV) and negative predictive value (NPV) for a diagnosis of heart failure. Economic evaluation of a decision tree with a NHS/Personal Social Services perspective determined the cost per quality-adjusted life-year (QALY) gained.

Results: In total, 304 participants were recruited to the validation cohort. The mean age was 73.9 years (standard deviation 8.8 years) and 124 (40.8%) participants were male. In total, 104 [34.2%, 95% confidence interval (CI) 28.9% to 39.8%] had a confirmed diagnosis of heart failure. The CDR had a sensitivity of 90% (95% CI 83% to 95%), specificity of 46% (95% CI 39% to 53%), PPV of 46% (95% CI 39% to 53%) and NPV of 90% (95% CI 83% to 95%). NT-proBNP level alone with a cut-off point of < 400 pg/ml had a sensitivity of 77% (95% CI 68% to 85%) and specificity of 92% (95% CI 87% to 95%). At the lower cut-off point of 125 pg/ml, sensitivity was 94% (95% CI 88% to 98%) and specificity was 49% (95% CI 42% to 56%). The economic model results suggest that referring a patient for a confirmatory diagnosis if they have had a previous myocardial infarction or have a NT-proBNP level that is greater than a 400 pg/ml threshold (current practice in England) is the most cost-effective option, with a cost of £4400 per QALY gained compared with a do nothing strategy. The base-case results were robust to deterministic and probabilistic sensitivity analyses.

Conclusions: Natriuretic peptide testing alone performed as well as the validated CDR in determining which patients presenting with symptoms went on to have a diagnosis of heart failure. The current NT-proBNP cut-off point of 400 pg/ml used in the UK is too high and means that one in five patients with heart failure may not be appropriately referred for further investigation and diagnosis, but this threshold was cost-effective in the REFer for Echocardiogram (REFER) trial. The study found only three patients with heart failure with reduced ejection fraction (HFREF), which might limit the benefits of early detection. The other diagnostic strategies with lower NT-proBNP referral levels become more cost-effective as the proportion of HFREF patients increases. International consensus on the optimal cut-off point for natriuretic peptide testing in patients with symptoms suggestive of heart failure should be sought.

Trial registration: Current Controlled Trials ISRCTN17635379.

Funding: This project was funded by the Efficacy and Mechanism Evaluation (EME) programme, a MRC and NIHR partnership.

Contents

List of tables	ix
List of figures	xi
List of boxes	xiii
List of abbreviations	xv
Plain English summary	xvii
Scientific summary	xix
Chapter 1 The REFER study: background	1
Heart failure diagnosis	1
Clinical decision rules	1
Natriuretic peptide testing	1
Aim of the study	2
Chapter 2 The REFER study: methods	3
Participants	3
Test methods	3
Statistical methods	4
Chapter 3 The REFER study: results	5
Participants	5
Comparison with the clinical decision rule derivation data set	7
Rates of confirmed heart failure	7
Diagnostic accuracy estimates	8
Assessing incorporation bias	10
Chapter 4 The REFER study: discussion	13
Summary of findings	13
Strengths and limitations	13
Comparison with existing literature	13
Implications for practice and research recommendations	14
Conclusions	15
Chapter 5 The REFER study health economic evaluation: background	17
The REFER study data set	17
Economic evaluation	17
Chapter 6 The REFER study health economic evaluation: methods	19
Economic evaluation diagnostic pathways	19
<i>MICE clinical decision rule</i>	19
<i>National Institute for Health and Care Excellence-recommended strategy</i>	19
<i>Echo all strategy</i>	19
<i>NT-proBNP 125 strategy</i>	19
<i>Do nothing strategy</i>	19

Decision tree structure	19
Effects of early heart failure diagnosis	21
Effects of delayed heart failure diagnosis	22
Treatment effects on mortality	22
Early heart failure treatment effect on hospital admissions	25
Costs associated with the diagnostic strategies	26
Deterministic sensitivity analysis	27
Probabilistic sensitivity analysis	27
Chapter 7 The REFER study health economic evaluation: results	29
Deterministic sensitivity analysis results	30
Probabilistic sensitivity analysis results	30
Chapter 8 The REFER study health economic evaluation: discussion	33
Summary of findings	33
Strengths and limitations	33
Comparison with existing literature	34
Implications for practice and research recommendations	34
Conclusions	34
Acknowledgements	35
References	37
Appendix 1 The REFER study clinical record form	41
Appendix 2 Patient and public involvement	55

List of tables

TABLE 1 Characteristics of the REFER study participants	6
TABLE 2 Objective abnormalities found on ECG and echocardiogram in participants with and without heart failure	7
TABLE 3 Characteristics of participants with and without heart failure	8
TABLE 4 Performance characteristics of the CDR and NT-proBNP level	9
TABLE 5 Diagnostic accuracy measures for participants with and without a NT-proBNP measurement	10
TABLE 6 Performance characteristics of CDR and NT-proBNP at steps 1–3	11
TABLE 7 Patient clinical characteristics	21
TABLE 8 Estimated undiscounted survival probabilities for male patients with heart failure on and off therapy	23
TABLE 9 Estimated undiscounted survival probabilities for female patients with heart failure on and off therapy	23
TABLE 10 Health utility states for the REFER patient population	25
TABLE 11 Cost parameters	26
TABLE 12 Parameter distributions for the probabilistic sensitivity analysis	27
TABLE 13 Base-case results	29
TABLE 14 Sensitivity analysis scenario results	30

List of figures

FIGURE 1	Flow diagram showing the number of participants in the REFER study	5
FIGURE 2	Venn diagram showing the presenting symptoms of participants	6
FIGURE 3	Receiver operating characteristic curves for CDR and NT-proBNP for predicting heart failure	10
FIGURE 4	Decision tree of the different diagnostic strategies for patients presenting with heart failure symptoms in primary care	20
FIGURE 5	Discounted survival curves for male patients on and off treatment	24
FIGURE 6	Discounted survival curves for female patients on and off treatment	24
FIGURE 7	Cost-effectiveness plane	29
FIGURE 8	Cost-effectiveness acceptability frontier showing the optimal diagnostic strategy across a range of willingness-to-pay thresholds	31

List of boxes

BOX 1 The MICE CDR

2

List of abbreviations

ACE	angiotensin-converting enzyme	HFREF	heart failure with reduced ejection fraction
ARB	angiotensin receptor blocker		
AUROC	area under the receiver operating characteristic curve	ICER	incremental cost-effectiveness ratio
BNF	<i>British National Formulary</i>	MICE	Male, history of myocardial Infarction, Crepitations at the lung bases and oEdema
BNP	B-type natriuretic peptide	NICE	National Institute for Health and Care Excellence
CDR	clinical decision rule		
CEAF	cost-effectiveness acceptability frontier	NIHR	National Institute for Health Research
CI	confidence interval	NPV	negative predictive value
COPD	chronic obstructive pulmonary disease	NT-proBNP	N-terminal pro-B-type natriuretic peptide
ECG	electrocardiogram	PPV	positive predictive value
EQ-5D	European Quality of Life-5 Dimensions	QALY	quality-adjusted life-year
ESC	European Society of Cardiology	REFER	REFer for Echocardiogram
GP	general practitioner	ROC	receiver operating characteristic
HFPEF	heart failure with preserved ejection fraction	SD	standard deviation
		SF-12	Short Form questionnaire-12 items

Plain English summary

Heat failure is a common problem in which the heart muscle functions less well and this results in unpleasant symptoms for patients (e.g. shortness of breath) and can shorten life. Fortunately, there are treatments available that can improve symptoms and delay the condition becoming worse, but symptoms can be vague, making diagnosis difficult. Accurate diagnosis requires referral for further investigation, which usually includes an ultrasound scan of the heart.

The aim of the REFER study was to investigate whether or not a clinical decision aid (called the MICE rule), with or without a simple blood test (called NT-proBNP), could improve the accuracy of general practitioner (GP) diagnosis of heart failure in people over the age of 55 years. The results showed that the MICE rule was not helpful in improving diagnosis but the blood test was helpful (and cost-effective) in deciding which people needed referral for a heart scan. The REFER study also indicated that the cut-off point for this blood test set by the National Institute for Health and Care Excellence (NICE) may be too high.

Patients and their carers should be reassured that this study showed that, if they go to see their GP with heart failure symptoms, the best thing for the GP to do is to carry out a blood test for NT-proBNP, which is what currently happens in the NHS. We recommend, however, that NICE considers lowering the blood test threshold level for further testing to ensure that patients with a possible diagnosis of heart failure are not missed.

Scientific summary

Background

Heart failure is a chronic disease associated with significant mortality and poor quality of life for patients. There are several evidence-based therapies that will delay heart failure progression, improve quality of life, reduce cardiovascular events and help avoid hospital admissions. A reliable and early diagnosis is, therefore, essential to guide the most appropriate management strategies. However, making an accurate and timely diagnosis requires referral for objective testing, but deciding who to refer can be challenging. The symptoms of heart failure are often non-specific and include gradual-onset breathlessness, fatigue and ankle swelling, symptoms not unique to heart failure and often associated with other conditions, or patients may have several coexisting diseases. Help with deciding who to refer and what tests to use is, therefore, crucial.

Clinical decision rules (CDRs) may help clinicians to assess the probability that a patient has a particular condition. The 'MICE' CDR was developed from an individual patient database meta-analysis of all ($n = 11$) prospective epidemiological studies of heart failure screening in primary care, which was commissioned as a health technology assessment by the National Institute for Health and Care Excellence (NICE). The CDR comprised four clinical elements – Male, history of myocardial Infarction, Crepitations at the lung bases and oEdema – and was combined with natriuretic peptide levels to identify those likely to have heart failure and who should be referred for further diagnostic testing.

Objectives

The REFER (REFer for EchocaRdiogram) trial aimed to assess the performance of the CDR, CDR and N-terminal pro-B-type natriuretic peptide (NT-proBNP) testing or NT-proBNP testing alone in identifying patients with heart failure presenting to primary care. The REFER trial was a prospective, observational, diagnostic validation study of the MICE CDR – with natriuretic peptide testing – for diagnosing heart failure in primary care.

The economic evaluation aimed to assess the cost-effectiveness of using the MICE CDR in heart failure diagnosis in general practice from a NHS and Personal Social Services perspective.

Methods

Primary care patients aged ≥ 55 years presenting to their general practitioner (GP) with symptoms suggestive of heart failure were recruited across 28 general practices in central England. All consenting patients underwent a full clinical assessment, which included a NT-proBNP test, an echocardiogram and a quality-of-life questionnaire, at a research clinic within 1 week of recruitment. Follow-up quality-of-life and resource-use questionnaires were mailed to the patients at 6 and 12 months after attending the clinic.

The diagnosis of 'heart failure' or 'no heart failure' was determined by an expert panel of cardiologists using the European Society of Cardiology 2012 definition. Clinical information, including the variables of the MICE rule and NT-proBNP level, was presented in stages to quantify any incorporation bias.

For the economic evaluation, a decision tree was developed comparing different diagnostic strategies using data from REFER participants to determine which symptomatic patients would receive the correct diagnostic decision. The model used a lifetime horizon and a UK NHS perspective.

Results

In total, 304 participants were recruited; the mean age of participants was 73.9 years (standard deviation 8.8 years) and 124 (40.8%) were male. In total, 104 participants [34.2%, 95% confidence interval (CI) 28.9% to 39.8%] had a confirmed diagnosis of heart failure. The CDR had a sensitivity of 90% (95% CI 83% to 95%) and a specificity of 46% (95% CI 39% to 53%). NT-proBNP level alone with a cut-off point of < 400 pg/ml had a sensitivity of 77% (95% CI 68% to 85%) and specificity of 92% (95% CI 87% to 95%). At the lower cut-off point of 125 pg/ml, sensitivity was 94% (95% CI 88% to 98%) and specificity was 49% (95% CI 42% to 56%).

The economic model used a lifetime horizon and a UK NHS perspective. The results suggest that use of the current recommended NICE guidelines for identifying patients with heart failure is the most cost-effective option, with a cost of £4400 per quality-adjusted life-year (QALY) gained compared with a do nothing strategy. That is, patients presenting with symptoms suggestive of heart failure should be referred straight for echocardiography if they have a history of myocardial infarction or if their NT-proBNP level is \geq 400 pg/ml. The MICE rule was more expensive and less effective than the other comparators. The base-case results were robust to sensitivity analysis.

Conclusions

Natriuretic peptide testing alone performed as well as the validated CDR in determining which patients presenting with symptoms went on to have a diagnosis of heart failure. The current NT-proBNP cut-off level of 400 pg/ml used in the UK is too high and means that one in five patients with heart failure may not be appropriately referred for further investigation and diagnosis.

The economics study represents the first cost–utility analysis comparing heart failure diagnostic strategies for symptomatic patients. Current practice in England is the most cost-effective option for identifying patients for confirmatory heart failure diagnosis. The low number of heart failure with reduced ejection fraction patients (12%) in the REFER patient population limited the benefits of early detection.

Trial registration

This trial is registered as ISRCTN17635379.

Funding

This project was funded by the Efficacy and Mechanism Evaluation (EME) programme, a MRC and NIHR partnership.

Chapter 1 The REFER study: background

Much of this report has been reproduced from Taylor *et al.*,¹ this article is Open Access: CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>), and Monahan *et al.*,² this article is Open Access: CC BY 4.0 license (<https://creativecommons.org/licenses/by/4.0/>).

Heart failure diagnosis

Heart failure is a chronic disease associated with significant mortality and poor quality of life for patients.³⁻⁵ Making an accurate and timely diagnosis is crucial and requires referral for objective testing but deciding who to refer can be challenging.⁶ The symptoms of heart failure are often non-specific and include gradual-onset breathlessness, fatigue and ankle swelling.⁷ However, these symptoms are not unique to heart failure and can be associated with other conditions or patients may have several coexisting diseases.⁸⁻¹⁰ Patients with symptoms that may be associated with heart failure often present to primary care. Identifying the patients likely to have heart failure, and therefore requiring referral for diagnostic testing, can be difficult.^{11,12}

Clinical decision rules

Clinical decision rules (CDRs) can help clinicians to assess the probability that a patient has a particular condition.¹³ They are used widely in medicine to inform decisions about investigation and management. For example, the Ottawa ankle rule can be useful when deciding if a patient with an ankle injury is likely to have a fracture and, therefore, whether radiography would be beneficial or unnecessary.¹⁴ Similarly, the CHA2DS2-VASc score can be used in patients with atrial fibrillation to determine the likely benefit of anticoagulation to reduce the risk of stroke.¹⁵ There is not currently a CDR for patients with heart failure presenting to primary care.

The 'MICE' rule was developed from a systematic review of the evidence for symptoms of heart failure.¹⁶ The review identified 11 prospective heart failure studies set in primary care. The decision rule was derived from an individual patient data set from one of these studies (Zaphiriou *et al.*¹⁷) and externally validated on four others that included relevant variables.¹⁸⁻²¹

The MICE rule comprised four clinical elements – Male, history of myocardial Infarction, Crepitations at the lung bases and oEdema – and was combined with natriuretic peptide levels to identify those likely to have heart failure and who should be referred for further diagnostic testing. The elements of the CDR are shown in *Box 1*.

Natriuretic peptide testing

Natriuretic peptides can be used to identify those patients with symptoms who have an increased likelihood of heart failure.²² B-type natriuretic peptide (BNP) and N-terminal pro-BNP (NT-proBNP) are released from the chambers of the heart in response to pressure or fluid overload, acting on the kidneys to induce a diuresis and on the vasculature to cause both arterial and venous dilatation, thereby reducing preload and afterload. Heart failure is associated with an increased natriuretic peptide level. A rising natriuretic peptide level can be an early sign of heart failure; however, other factors such as renal impairment and angiotensin-converting enzyme (ACE) inhibitors can influence the results, which therefore need to be interpreted in light of these other factors. The European Society of Cardiology (ESC)⁶ suggests a NT-proBNP threshold of 125 pg/ml to rule out heart failure, whereas the National Institute for Health and Care Excellence (NICE)⁷ in England suggests a NT-proBNP threshold of 400 pg/ml.

BOX 1 The MICE CDR¹⁶

Refer straight for echocardiography if the patient has any one of:

- a history of myocardial infarction
- basal crepitations
- ankle oedema in a male.

Otherwise, carry out a B-type natriuretic peptide (BNP) [or N-terminal pro-BNP (NT-proBNP)] test and refer straight for echocardiography if the BNP/NT-proBNP level is above one of three cut-off points set by sex/symptoms recorded in the clinical rule:

1. female without ankle oedema: refer if BNP is > 210–360 pg/ml depending on local availability of echocardiography (or NT-proBNP is > 620–1060 pg/ml)
2. male without ankle oedema: refer if BNP is > 130–220 pg/ml (or NT-proBNP is > 390–660 pg/ml)
3. female with ankle oedema: refer if BNP is > 100–180 pg/ml (or NT-proBNP is > 190–520 pg/ml).

Aim of the study

This study aimed to assess the performance of the CDR, NT-proBNP and their combination in identifying patients with heart failure presenting to primary care.

Chapter 2 The REFER study: methods

The full protocol for the REFer for Echocardiogram (REFER) study has been published previously²³ and a copy of the clinical record form used for each participant is available in *Appendix 1*.

Participants

The REFER study used a prospective, observational, diagnostic validation design to assess the performance of the MICE rule and NT-proBNP levels in identifying patients with heart failure. The study population was primary care patients aged ≥ 55 years presenting with recent new-onset shortness of breath, lethargy or peripheral ankle oedema of > 48 hours' duration for which there was no other obvious cause. Patients were excluded if they were unable to give consent or had a previous confirmed diagnosis (i.e. with objective evidence) of heart failure, an obvious alternative diagnosis, severe symptoms requiring immediate management or recent (within 60 days) acute coronary syndrome.

Participants were recruited from 28 general practices around the West Midlands region of the UK (central England) based on their presenting symptoms. Participating practices were asked to invite all presenting patients who met the inclusion criteria to join the study consecutively. If a patient expressed an interest in taking part, the general practitioner (GP) completed an online portal with contact details and printed out a patient information sheet. Patients could choose whether to contact the study team using the designated telephone number or provide their own number to be contacted by a member of the study team. An appointment for participants to attend the research clinic was arranged during the initial telephone call.

Test methods

Participants attended the research assessment clinic within 7 days of recruitment. The clinics were held at two practices – one in the north and one in the south of the region. Assessments were carried out by trained research nurses and a British Society of Echocardiography (BSE)-accredited echocardiographer. The purpose and process of the study was explained and written consent was taken from patients. A research nurse then completed a clinical record form containing a clinical history and examination findings. Blood was sampled for NT-proBNP testing and renal function. Two attempts at blood-taking were allowed. The NT-proBNP level was determined using a point-of-care device (Roche Diagnostics, Burgess Hill, UK). The echocardiographer then carried out an electrocardiogram (ECG) and an echocardiogram. Finally, participants were asked to complete two quality-of-life questionnaires: the Short Form questionnaire-12 items (SF-12) (version 1) and the European Quality of Life-5 Dimensions (EQ-5D) (three-level version). The SF-12 is a validated questionnaire that has been used to measure health-related quality of life.²⁴ The EQ-5D focuses on five domains: mobility, self-care, activity, pain and anxiety.²⁵

The test results were faxed back to the GP on the day of the research assessment clinic. Patients were encouraged to make a follow-up appointment with their GP to discuss the results and allow any further management to be arranged.

The reference standard was an expert consensus panel of three cardiology specialists, who reviewed each case blinded to the assessments by other panel members. The ESC 2012 guideline was used to define heart failure.⁶ The panel was presented with clinical information and investigation results in three separate stages to assess incorporation bias. At step 1 the clinical assessment (excluding the MICE variables) and ECG and echocardiogram findings were presented. At step 2 the CDR components (male, history of myocardial infarction, crepitations and oedema) were added and, finally, at step 3 the NT-proBNP result was included. The cardiology specialists were asked to record whether the patient did or did not have heart failure at each of the three steps. At least two cardiologists needed to agree the diagnosis

independently at each step before the panel could move on to the next step. In the case of disagreement a third cardiologist adjudicated the case, blind to the other panel members' assessments, before the panel could move to the next step.

Statistical methods

A sample of 500 symptomatic patients attending their GP for breathlessness, lethargy or ankle swelling was proposed. This sample size was sufficient to estimate the sensitivity of the CDR to within 4% and specificity to within 6% at the 95% confidence level. Calculations were based on a sensitivity of 94% and specificity of 48% obtained from the previous individual patient data meta-analysis²² and prevalence of heart failure in a symptomatic population of 30%.

Participants with and without a diagnosis of heart failure at step 3 were compared using independent *t*-tests or Wilcoxon rank-sum tests for continuous measures and chi-squared tests for categorical variables. The main outcome measures were test performance of the CDR and natriuretic peptide test – alone and in combination – in estimating a diagnosis of heart failure. The findings of the expert consensus panel determined whether heart failure – the 'observed disease' – was present or absent. The CDR and NT-proBNP results were also used to determine whether heart failure was likely to be present – the 'test disease' – and referral for echocardiography would have been indicated. Observed compared with test disease status was then cross-tabulated to determine the sensitivity and specificity, positive predictive value (PPV) and negative predictive value (NPV) for the CDR, NT-proBNP and CDR and NT-proBNP in combination and also by the NT-proBNP cut-off points of 125 pg/ml and 400 pg/ml suggested by the ESC⁴ and NICE⁵ guidelines, respectively; 95% confidence intervals (CIs) were calculated using the binomial exact method. Receiver operating characteristic (ROC) curves were generated to determine the overall discriminatory ability of each test in predicting a diagnosis of heart failure. Comparisons were made between the performance characteristics of the current cohort and those observed in the original derivation data set.²²

Chapter 3 The REFER study: results

Participants

The recruitment phase of the REFER study started on 1 May 2011 and was completed on 31 August 2013. Participants were recruited from a random sample of 28 general practices in central England, stratified by practice list size and deprivation quartile.²⁶ Assessment was undertaken at the research clinic within 7 days of participants presenting to their GP.

Figure 1 shows a flow diagram of patient recruitment. In total, 397 patients were eligible for inclusion, of whom 45 were excluded: 37 did not provide consent, three did not attend the appointment, four lost contact with the study team and one was hospitalised before attending the study appointment. Of the 37 patients who did not provide consent, it was normally the case that they expressed an interest to their GP at the time of the consultation but either did not subsequently contact the research team or did not want to take part when the study team called to arrange an appointment at the assessment clinic. Of the 352 remaining participants, 48 did not have a blood test because of failed venepuncture or, in a small number of cases, a nurse not being available. These participants did not proceed to step 3 for the reference standard diagnosis of heart failure because of an unknown level of NT-proBNP and were excluded from the analysis presented in this paper.

The remaining 304 participants formed our validation cohort; participants were similar to those excluded with respect to demography and medical history except for previous record of heart failure, with those without a NT-proBNP test having a higher prevalence (2.3% vs. 8.3%). These heart failure labels from the routine clinical records were, however, not necessarily confirmed with objective evidence or a formal diagnosis.

The most common presenting symptom was breathlessness; however, over half of the participants had all three symptoms of breathlessness, ankle oedema and lethargy, as shown in Figure 2.

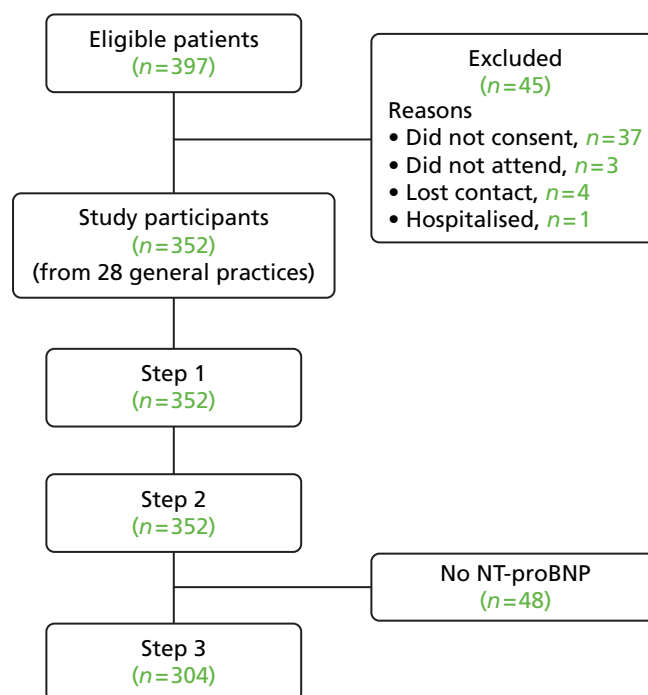


FIGURE 1 Flow diagram showing the number of participants in the REFER study. Step 1 = clinical information + ECG + echocardiogram; step 2 = CDR variables; step 3 = NT-proBNP result.

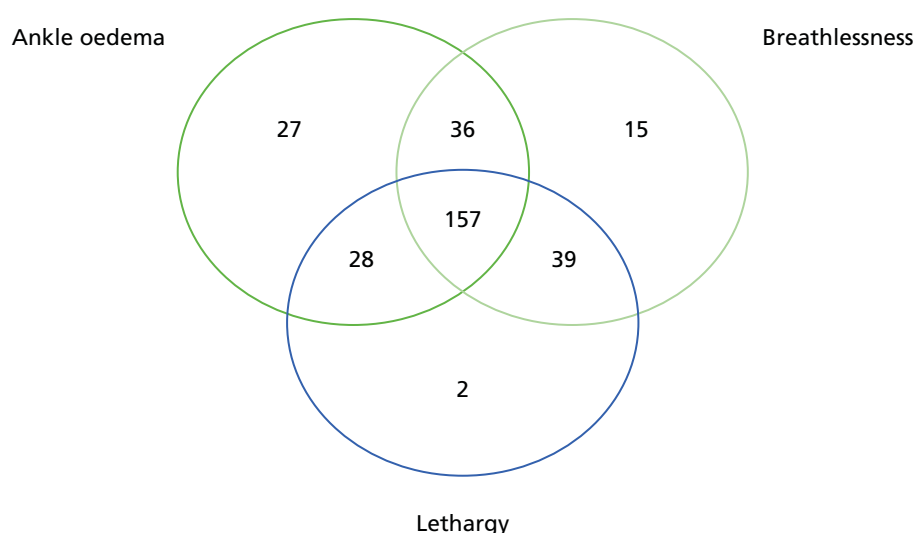


FIGURE 2 Venn diagram showing the presenting symptoms of participants.

The clinical and demographic characteristics of the study population are shown in *Table 1*. The mean age was 73.9 years [standard deviation (SD) 8.8 years] and 124 (40.8%) participants were male. The cohort had a mix of ethnicities, including 18.4% Asian or Asian British participants. Cardiovascular risk factors such as hypertension and diabetes mellitus were prevalent: 221 (72.7%) participants reported being hypertensive and 86 (28.3%) had diabetes mellitus. Comorbidities were common: 183 (60.2%) participants had arthritis and 73 (24.0%) had depression. Seven participants had a record of previous heart failure but this was not confirmed with objective evidence or a formal diagnosis. Cardiovascular medications were commonly prescribed, possibly relating to the high rate of hypertension in the cohort.

TABLE 1 Characteristics of the REFER study participants

Characteristic	REFER study participants, n (%)	Derivation data set, ¹⁷ n (%)
Age (years), mean (SD)	73.9 (8.8)	71.5 (11.5)
Male	124 (40.8)	122 (41)
Ethnicity		
White	214 (70.4)	–
Asian/Asian British	56 (18.4)	–
Black/black British	16 (5.3)	–
Other	18 (5.6)	–
Presenting symptom		
Ankle oedema	248 (81.6)	191 (64)
Breathlessness	247 (81.3)	283 (95)
Lethargy	226 (74.3)	185 (62.1)
Previous MI	34 (11.2)	42 (14)
Basal crepitations	16 (5.3)	81 (27)
Hypertension	221 (72.7)	166 (55.7)
Diabetes mellitus	86 (28.3)	58 (19.0)
COPD	17 (5.6)	58 (19.0)

TABLE 1 Characteristics of the REFER study participants (*continued*)

Characteristic	REFER study participants, <i>n</i> (%)	Derivation data set, ¹⁷ <i>n</i> (%)
Depression	73 (24.0)	–
Arthritis	183 (60.2)	–
Medications		
ACE inhibitors	98 (32.2)	71 (23.2)
Beta-blockers	82 (27.0)	71 (23.2)
ARBs	58 (19.1)	–
Diuretics	136 (44.7)	190 (63.8)
NT-proBNP (pg/ml), median (IQR)	214 (79–494)	381.5 (135–1187)

ARB, angiotensin receptor blocker; COPD, chronic obstructive pulmonary disease; IQR, interquartile range; MI, myocardial infarction; SD, standard deviation.

Comparison with the clinical decision rule derivation data set

The REFER cohort, although similar in age and sex to the derivation data set,¹⁷ had fewer referrals because of shortness of breath and more referrals because of ankle oedema and lethargy. Hypertension and diabetes mellitus were observed in greater frequency in the REFER population but a lower proportion of patients had chronic obstructive pulmonary disease (COPD). The prescribing of diuretics was less frequent in the REFER cohort but a higher proportion was prescribed ACE inhibitors.

Rates of confirmed heart failure

The expert panel reviewed the data for each participant in three steps and determined whether or not a heart failure diagnosis was present at each stage. Forty-eight participants did not have a NT-proBNP level measured and so could not progress to step 3. Of the 304 cases with a NT-proBNP level recorded, 66 (21.7%, 95% CI 17.2% to 26.8%) were diagnosed with heart failure at step 1, 89 (29.3%, 95% CI 24.2% to 34.7%) at step 2 and 104 (34.2%, 95% CI 28.9% to 39.8%) at step 3. The objective abnormalities found on ECG and echocardiogram are shown in *Table 2*.

The characteristics of participants with and without heart failure are shown in *Table 3*. Participants with heart failure were older. Half of those with heart failure were male, whereas about one-third of those

TABLE 2 Objective abnormalities found on ECG and echocardiogram in participants with and without heart failure

Abnormality ^a	Heart failure, <i>n</i> (%)	No heart failure, <i>n</i> (%)
Moderate to severe LVSD: ejection fraction < 40%	3 (2.9)	0 (0)
Borderline LVSD: ejection fraction 41–50%	9 (8.7)	1 (0.5)
Diastolic dysfunction	15 (14.4)	6 (3.0)
Significant valve disease	47 (45.2)	17 (8.5)
Atrial fibrillation	33 (31.7)	0 (0)
Total in group	104	200

LVSD, left ventricular systolic dysfunction.

a Some participants had more than one abnormality.

TABLE 3 Characteristics of participants with and without heart failure

Characteristic	Heart failure (<i>N</i> = 104), <i>n</i> (%)	No heart failure (<i>N</i> = 200), <i>n</i> (%)	<i>p</i> -value
Age (years), mean (SD)	77.4 (7.4)	72.1 (9.0)	< 0.0001
Male	52 (50.0)	72 (36.0)	0.02
BMI (kg/m ²), mean (SD)	29.1 (5.7)	31.1 (6.7)	0.008
Presenting symptom			
Breathlessness	84 (80.8)	163 (81.5)	0.88
Ankle oedema	87 (83.7)	161 (80.5)	0.50
Lethargy	72 (69.2)	154 (77.0)	0.14
Basal crepitations	4 (3.8)	12 (6.0)	0.42
Previous MI	17 (16.3)	17 (8.5)	0.04
Hypertension	79 (76.0)	142 (71.0)	0.36
Diabetes mellitus	29 (27.9)	57 (28.5)	0.91
Depression	17 (16.3)	56 (28.0)	0.02
COPD	7 (6.7)	10 (5.0)	0.53
Arthritis	55 (52.9)	128 (64.0)	0.06
Medication			
ACE inhibitors	38 (36.5)	60 (30.0)	0.25
Beta-blockers	46 (44.2)	36 (18.0)	< 0.0001
ARBs	19 (18.3)	39 (19.5)	0.80
Diuretics	61 (58.7)	75 (37.5)	0.0004
NT-proBNP (pg/ml), median (IQR)	715.5 (413–1559)	126 (60–233)	< 0.0001
NT-proBNP ≥ 125 pg/ml	98 (94.2)	102 (51.0)	< 0.0001
NT-proBNP ≥ 400 pg/ml	80 (76.9)	17 (8.5)	< 0.0001

ARB, angiotensin receptor blocker; BMI, body mass index; IQR, interquartile range; MI, myocardial infarction.

without heart failure were male. The presenting symptom profile was similar between the groups. Proportionately more patients with heart failure had a history of myocardial infarction (16.4% vs. 8.5%) but there was no significant difference in other comorbidities such as hypertension, COPD and arthritis. Depression was more common in the non-heart failure group. Cardiovascular medications were more likely to be prescribed in those with heart failure than in those without heart failure. The median NT-proBNP level was significantly higher in the heart failure group. At the lower 125 pg/ml cut-off point, over half of patients without heart failure had a NT-proBNP level above the threshold for referral to echocardiography.

Diagnostic accuracy estimates

The diagnostic accuracy of the CDR, NT-proBNP level and CDR and NT-proBNP level in combination is shown in *Table 4*. The clinical information (MICE symptoms) of the CDR had a sensitivity of 44.2% (95% CI 34.5% to 54.3%) but, with the addition of the NT-proBNP level at the lower cut-off point, the sensitivity improved to 90.4% (95% CI 83.0% to 95.3%) and specificity was 45.5% (95% CI 38.5% to 52.7%). The NT-proBNP level alone with a cut-off point of 400 pg/ml had a sensitivity of 76.9% (95% CI 67.6% to 84.6%) and a specificity of 91.5% (95% CI 86.7% to 95.0%). At the lower cut-off point of 125 pg/ml, sensitivity was 94.2% (95% CI 87.9% to 97.9%) and specificity was 49.0% (95% CI 41.9% to

TABLE 4 Performance characteristics of the CDR and NT-proBNP level

Characteristic	AUROC (95% CI)	Sensitivity (95% CI) (%)	Specificity (95% CI) (%)	PPV (95% CI) (%)	NPV (95% CI) (%)	LR+ (95% CI)	LR- (95% CI)
Derivation data set¹⁷							
CDR: lower cut-off point ^a	0.74 (0.70 to 0.79)	90.2 (82.7 to 95.2)	58.2 (50.9 to 65.2)	52.9 (42.6 to 64.8)	91.9 (75.8 to 100)	2.16 (1.81 to 2.57)	0.17 (0.09 to 0.31)
CDR: upper cut-off point ^b	0.75 (0.70 to 0.80)	87.3 (79.2 to 93.0)	62.2 (55.5 to 69.1)	54.6 (46.6 to 62.4)	90.4 (84.1 to 94.8)	2.31 (1.9 to 2.81)	0.21 (0.12 to 0.34)
REFER study participants							
CDR – MICE variables	0.54 (0.48 to 0.60)	44.2 (34.5 to 54.3)	64.0 (56.9 to 70.6)	39.0 (30.1 to 48.4)	68.8 (61.6 to 75.4)	1.23 (0.93 to 1.63)	0.87 (0.71 to 1.06)
CDR and NT-proBNP: lower cut-off points ^a	0.68 (0.64 to 0.72)	90.4 (83.0 to 95.3)	45.5 (38.5 to 52.7)	46.3 (39.3 to 53.4)	90.1 (82.5 to 95.1)	1.66 (1.44 to 1.91)	0.21 (0.12 to 0.39)
CDR and NT-proBNP: upper cut-off points ^b	0.71 (0.66 to 0.76)	78.8 (69.7 to 86.2)	63.5 (56.4 to 70.2)	52.9 (44.7 to 61.0)	85.2 (78.5 to 90.5)	2.16 (1.75 to 2.66)	0.33 (0.23 to 0.49)
NT-proBNP ≥ 125 pg/ml alone	0.72 (0.67 to 0.76)	94.2 (87.9 to 97.9)	49.0 (41.9 to 56.1)	49.0 (41.9 to 56.1)	94.2 (87.9 to 97.9)	1.85 (1.6 to 2.13)	0.12 (0.05 to 0.26)
NT-proBNP ≥ 400 pg/ml alone	0.84 (0.80 to 0.89)	76.9 (67.6 to 84.6)	91.5 (86.7 to 95.0)	82.5 (73.4 to 89.4)	88.4 (83.2 to 92.4)	9.05 (5.67 to 14.4)	0.25 (0.18 to 0.36)
AUROC, area under the ROC curve; LR, likelihood ratio.							
a Post-test probability of 20%.							
b Post-test probability of 30%.							

56.1%). These performance characteristics were mostly lower in magnitude than the corresponding values observed in the derivation data set. However, comparison of the CIs suggests that the differences were not statistically different at the 5% level.

Figure 3 shows the ROC curves for each index test for predicting heart failure. Significant differences ($p < 0.0001$) were observed between the areas under the ROC curves (AUROCs) shown in Table 4. NT-proBNP had the best discriminatory power, with an AUROC of 0.91 (95% CI 0.88 to 0.95), and the clinical element of the MICE CDR the poorest, with an AUROC of 0.54 (95% CI 0.48 to 0.60).

There were no significant differences in diagnostic accuracy measures of the CDR between those with and those without missing NT-proBNP results, as shown in Table 5.

Assessing incorporation bias

The performance characteristics for the CDR and NT-proBNP at steps 1–3 are shown in Table 6. The diagnostic accuracy of all tests increased at each step, with the largest changes observed when NT-proBNP was used without the clinical element of the CDR. A NT-proBNP cut-off point of 400 pg/ml showed a statistically significant increase in the detection of cases without heart failure from step 2 to step 3 ($p < 0.05$).

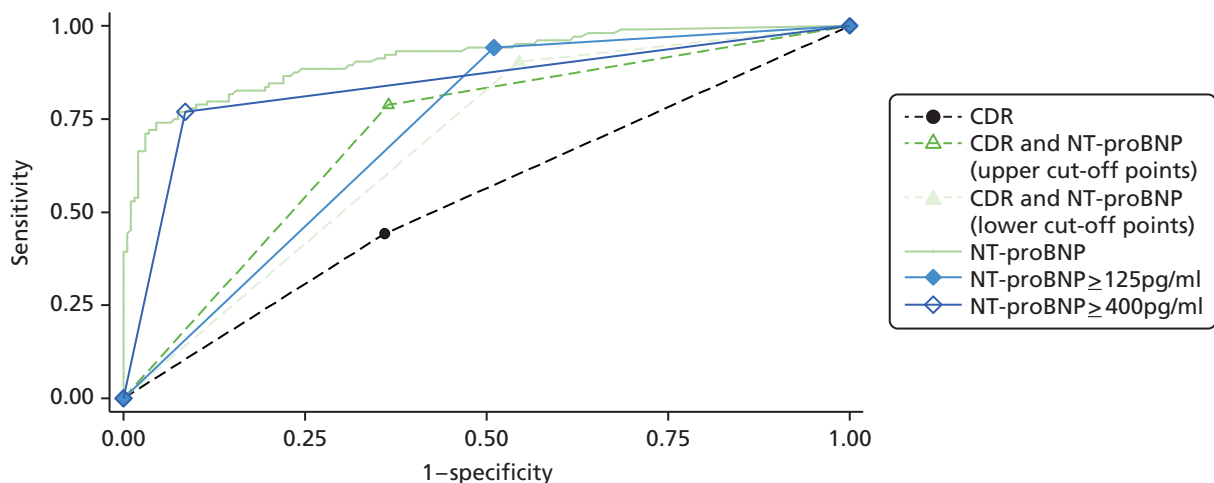


FIGURE 3 Receiver operating characteristic curves for CDR and NT-proBNP for predicting heart failure.

TABLE 5 Diagnostic accuracy measures for participants with and without a NT-proBNP measurement

Population	AUROC (95% CI)	Sensitivity (95% CI) (%)	Specificity (95% CI) (%)
With NT-proBNP	0.59 (0.53 to 0.65)	51.7 (40.8 to 62.4)	66.5 (59.8 to 72.8)
Without NT-proBNP	0.63 (0.49 to 0.78)	58.8 (32.9 to 81.6)	67.7 (48.6 to 83.3)

TABLE 6 Performance characteristics of CDR and NT-proBNP at steps 1–3

Step	Sensitivity (95% CI) (%)	Specificity (95% CI) (%)	PPV (95% CI) (%)	NPV (95% CI) (%)
CDR and NT-proBNP (lower cut-off points)				
1	78.8 (67 to 87.9)	36.6 (30.4 to 43.0)	25.6 (19.8 to 32.2)	86.1 (77.8 to 92.2)
2	85.4 (76.3 to 92.0)	40.9 (34.3 to 47.8)	37.4 (30.8 to 44.5)	87.1 (79.0 to 93.0)
3	90.4 (83.0 to 95.3)	45.5 (38.5 to 52.7)	46.3 (39.3 to 53.4)	90.1 (82.5 to 95.1)
CDR and NT-proBNP (upper cut-off points)				
1	60.6 (47.8 to 72.4)	51.7 (45.1 to 58.2)	25.8 (19.1 to 33.4)	82.6 (75.5 to 88.3)
2	74.2 (63.8 to 89.2)	58.6 (51.7 to 65.3)	42.6 (34.7 to 50.8)	84.6 (77.7 to 90.0)
3	78.8 (69.7 to 86.2)	63.5 (56.4 to 70.2)	52.9 (44.7 to 61.0)	85.2 (78.5 to 90.5)
NT-proBNP \geq 125 pg/ml				
1	81.8 (70.4 to 90.2)	38.7 (32.4 to 45.2)	27 (21.0 to 33.7)	88.5 (80.7 to 93.9)
2	84.3 (75.0 to 91.1)	41.9 (35.2 to 48.8)	37.5 (30.8 to 44.6)	86.5 (78.4 to 92.4)
3	94.2 (87.9 to 97.9)	49 (41.9 to 56.1)	49 (41.9 to 56.1)	94.2 (87.9 to 97.9)
NT-proBNP \geq 400 pg/ml				
1	48.5 (36.0 to 61.1)	72.7 (66.6 to 78.2)	33 (23.8 to 43.3)	83.6 (77.8 to 88.3)
2	58.4 (47.5 to 68.8)	79.1 (73.0 to 84.3)	53.6 (43.2 to 63.8)	82.1 (76.2 to 87.1)
3	76.9 (67.6 to 84.6)	91.5 (86.7 to 95.0)	82.5 (73.4 to 89.4)	88.4 (83.2 to 92.4)

Chapter 4 The REFER study: discussion

Summary of findings

At a low cut-off point of 125 pg/ml, NT-proBNP testing alone showed similar diagnostic performance to the CDR (clinical features and natriuretic peptide level) in determining which patients presenting with possible heart failure should be referred for diagnostic testing. The current NT-proBNP cut-off point of 400 pg/ml recommended by NICE⁷ in the UK is too high. Sensitivity for a diagnosis of heart failure at this level was 77% (95% CI 68% to 85%), meaning that one in five patients with heart failure presenting to primary care may not be appropriately referred for further investigation and diagnosis.

Strengths and limitations

This study included patients presenting prospectively to their GP. A large proportion of health care in the UK is provided through general practice and testing the CDR in a real-life clinical setting where most patients are managed allows accurate validation of the rule.²⁷

Participants underwent thorough phenotyping including clinical and objective assessment and then review by a panel of three experienced cardiologists, using a staged system to allow for assessment of inclusion bias, to agree a formal diagnosis so that the 'observed disease' was accurate.²⁸

The study was slow to recruit and failed to meet the initial target of 500 patients. The main reason for this slow recruitment, in common with many studies requiring opportunistic recruitment during routine consultations, was the additional administrative burden on busy GPs. This was at a time of unparalleled increased workloads in UK general practice. Furthermore, when the study was designed, natriuretic peptide assays were not routinely available to UK GPs and, therefore, the provision of a rapid natriuretic peptide and echocardiography service might have provided some 'compensation' for the additional administration needed for the study. However, shortly after the study commenced, natriuretic peptide assays became an open-access diagnostic for practices, alongside one-stop heart failure diagnostic clinics.⁷

The number of participants with heart failure with reduced ejection fraction (HFREF) was unexpectedly low in the cohort. This may reflect the nature of heart failure presentation, with those with HFREF perhaps more likely to present acutely directly to secondary care or perhaps already being under the care of a cardiologist for a known cardiovascular comorbidity such as coronary artery disease.²⁹

The increase in performance across the stepped diagnosis suggests that the sensitivity of the index tests may have been overestimated because of incorporation bias. However, evaluation of the results at step 2 (at which point NT-proBNP testing is excluded from clinical diagnosis) confirms that the diagnostic accuracy of the NT-proBNP test alone at the lower cut-off is similar to that of the CDR (in combination with NT-proBNP testing).

Comparison with existing literature

Heart failure can be a difficult diagnosis to make and the idea of a CDR to help primary care clinicians make the decision of who to refer for objective testing is justifiable.³⁰ The role of CDRs as an aid to clinical decision-making however remains controversial. There are many examples of CDRs being generated and validated with the hope of improving clinical accuracy but performance characteristics are

often modest at best.^{13,31–33} Furthermore, remembering the components of a CDR and applying them during the consultation can be challenging for busy generalist clinicians seeing patients with undifferentiated illness.³⁴

The reason why the CDR performed no better than NT-proBNP testing alone may be because the diagnosis of heart failure in the cohort was largely heart failure with preserved ejection fraction (HFPEF). The derivation and initial validation of the CDR relied predominantly on epidemiological studies that included heart failure as a result of left ventricular systolic dysfunction, or HFREF, and so may not directly apply to this population.¹⁶ In addition, the way that symptoms were recorded may have differed – the study used to derive the CDR was carried out by cardiologists in a secondary care clinic, whereas the REFER study data were collected by research nurses; however, both studies relied on referral from primary care.

The prevalence of atrial fibrillation and valvular disease was also very high in the REFER cohort. However, this is likely to be an increasing reality in primary care, with the number of patients with HFPEF and/or other cardiovascular comorbidities increasing.³⁵

The threshold for NT-proBNP below which heart failure can be reasonably excluded is also an area of ongoing research.³⁶ Guidelines differ in the threshold currently recommended. For example, the ESC⁴ suggests that a NT-proBNP level of < 125 pg/ml should be used to rule out heart failure but, in the UK, NICE⁷ suggests a much higher NT-proBNP threshold of 400 pg/ml. For any test there must always be a trade-off between sensitivity and specificity.²² A high sensitivity will ensure that fewer cases are missed, but at the expense of more patients undergoing echocardiography, which is an expense to the health-care system and may be an unnecessary inconvenience for the patient. However, accepting a test with a sensitivity that is too low could result in a diagnosis of heart failure being missed.

Implications for practice and research recommendations

Echocardiography remains the most commonly used modality for providing objective evidence of heart failure,³⁷ but all health-care systems must work within the constraints of finite resources and in many countries there is a limited supply of adequately trained echocardiography technicians. Access to a CDR and/or point-of-care blood test that can reliably rule heart failure in or out quickly would be attractive to clinicians. This study shows that, in patients suspected of having heart failure, a negative NT-proBNP blood test alone, at a threshold of 125 pg/ml, means that heart failure is unlikely and, thus, this could be used as a 'rule out' test. Provision of point-of-care testing in a primary care setting could aid clinicians in their decision on whether or not to refer for echocardiography, which could improve care for patients and reduce the burden on echocardiography services.

In the REFER study, a large proportion of the participants with heart failure had HFPEF. Estimates of the proportion of patients with HFREF compared with HFPEF vary but these findings may reflect the contemporary population presenting to general practice with symptoms suggestive of heart failure. This has important implications for how patients in general practice might be managed, given the current lack of effective treatments. Further research to find strategies to improve quality of life, reduce hospitalisations or extend survival in HFPEF patients is needed.

Natriuretic peptide levels are influenced by multiple factors. Arrhythmias such as atrial fibrillation are associated with high circulating levels, whereas medications such as ACE inhibitors reduce natriuretic peptide production. Interpreting the NT-proBNP level in clinical practice needs to be carried out in the context of a patient's comorbid conditions and treatment regime. Further research to provide guidance on natriuretic peptide correction factors could be helpful to clinicians.

Conclusions

Natriuretic peptide testing alone performed as well as a validated CDR, based on the best available epidemiological evidence, in determining which patients went on to have a diagnosis of heart failure. The current guidelines, which suggest measuring NT-proBNP in patients with symptoms suggestive of heart failure, are justified, but the cut-off point must be low enough to ensure that GPs are not falsely reassured that referral is not required.

Chapter 5 The REFER study health economic evaluation: background

Heart failure is a common clinical condition that has a major impact on patients and is associated with high costs for health systems, but it is not easy to diagnose accurately or early in primary care.^{4,12,28} The optimal cost-effective strategy for diagnosing heart failure patients at primary care level is not known.

The main symptoms suggestive of heart failure are shortness of breath, tiredness and swollen ankles, but these complaints are common in primary care and most patients presenting with them will not have heart failure.³⁸ Furthermore, referring symptomatic patients on for confirmatory investigations such as echocardiography is expensive. However, early detection of heart failure is important as evidence-based therapies can substantially improve quality of life, reduce premature mortality and reduce avoidable hospital admissions.^{6,7}

There are very few data on diagnostic strategies in patients presenting in primary care with symptoms suggestive of heart failure in which the population are rigorously phenotyped for heart failure. Economic analyses of diagnostic triage for this patient population are rarer still, but the REFER study provides appropriate data to undertake an economic evaluation.

The REFER study data set

The REFER study was a prospective, observational, diagnostic validation study of the MICE CDR, with natriuretic peptide testing, for diagnosing heart failure in primary care. The full methods for the REFER study have been published previously.²³ The CDR was developed from an individual patient data meta-analysis of epidemiological studies of heart failure screening in primary care, commissioned as a health technology assessment by NICE.⁴

Briefly, primary care patients aged ≥ 55 years presenting to their GP with symptoms suggestive of heart failure were recruited across 28 central England practices in the UK. All consenting patients underwent a full clinical assessment, which included a NT-proBNP test, an echocardiogram and a quality-of-life questionnaire, at a research clinic within 1 week of recruitment. Follow-up quality-of-life and resource-use questionnaires were mailed to the patients at 6 and 12 months after attending the clinic.

The diagnosis of heart failure or no heart failure was determined by an expert panel of cardiologists using the ESC 2012 definition.⁶ Clinical information, including the variables of the MICE rule and NT-proBNP level, was presented in stages to quantify any incorporation bias.

Economic evaluation

The aim of this study was to assess the cost-effectiveness of using the MICE CDR in heart failure diagnosis in general practice from a NHS and Personal Social Services perspective. To do so, a decision tree was developed comparing different diagnostic strategies against the CDR. The economic analysis utilised the REFER data set to determine which symptomatic patients would receive the correct diagnostic decision. The cost and quality of life consequences of a correct and incorrect referral were taken from the data set and the literature. The economic evaluation took a lifetime horizon and all costs and outcomes were discounted at 3.5% in accordance with NICE guidance on the methods of technology appraisal.³⁹

Chapter 6 The REFER study health economic evaluation: methods

The six comparators are described in the following section. As the MICE rule has lower and upper cut-off points for NT-proBNP referral levels, we treat it here as two different diagnostic comparators: MICE upper cut-off points and MICE lower cut-off points. The strategies differ in terms of immediate actions. All patients with true heart failure who are not referred at this stage are assumed to return 6 months later and such patients will be referred immediately for echocardiography.

Economic evaluation diagnostic pathways

MICE clinical decision rule

The MICE CDR was developed from a previous individual patient data meta-analysis.⁴⁰ The MICE CDR states that patients presenting with heart failure symptoms to their GP will be referred straight for echocardiography if they have a history of myocardial infarction or basal crepitations or are male with ankle oedema. Otherwise, a NT-proBNP test is carried out and the patient is referred straight for echocardiography if the test results are above one of three cut-off points set by sex/symptoms recorded in the clinical rule (with the upper MICE NT-proBNP cut-off points in parentheses):

1. A female patient without ankle oedema should be referred if NT-proBNP is ≥ 620 pg/ml (≥ 1060 pg/ml).
2. A male patient without ankle oedema should be referred if NT-proBNP is ≥ 390 pg/ml (≥ 660 pg/ml).
3. A female patient with ankle oedema should be referred if NT-proBNP is ≥ 190 pg/ml (≥ 520 pg/ml).

National Institute for Health and Care Excellence-recommended strategy

The NICE guidelines for the management of chronic heart failure⁷ suggest that a patient presenting with symptoms suggestive of heart failure should be referred straight for echocardiography if they have a history of myocardial infarction. Otherwise, a NT-proBNP test should be carried out and the patient referred for an echocardiograph if the NT-proBNP level is ≥ 400 pg/ml.

Echo all strategy

In the echo all strategy, all patients presenting to their GP with heart failure symptoms would be referred straight for echocardiography. We made a simplifying assumption that there would be no problems with access to echocardiography and that there would be a sufficient number of trained echocardiographers for all patients.

NT-proBNP 125 strategy

In the NT-proBNP 125 strategy, all patients presenting to their GP with heart failure symptoms would have a NT-proBNP test carried out and be referred for echocardiography if their NT-proBNP level was ≥ 125 pg/ml.

Do nothing strategy

In the do nothing strategy, no patients presenting to their GP with heart failure symptoms would be referred straight for echocardiography or undergo a NT-proBNP test.

Decision tree structure

A decision tree, presented in TreeAge Pro 2014 (TreeAge Software, Inc., Williamstown, MA, USA) and developed in Microsoft Excel®, version 2010 (Microsoft Corporation, Redmond, WA, USA), was structured to represent the various diagnostic strategies (Figure 4). A decision tree was the most appropriate method to use here as a 'one-off' decision is involved (i.e. whether or not to send patients on for a confirmatory

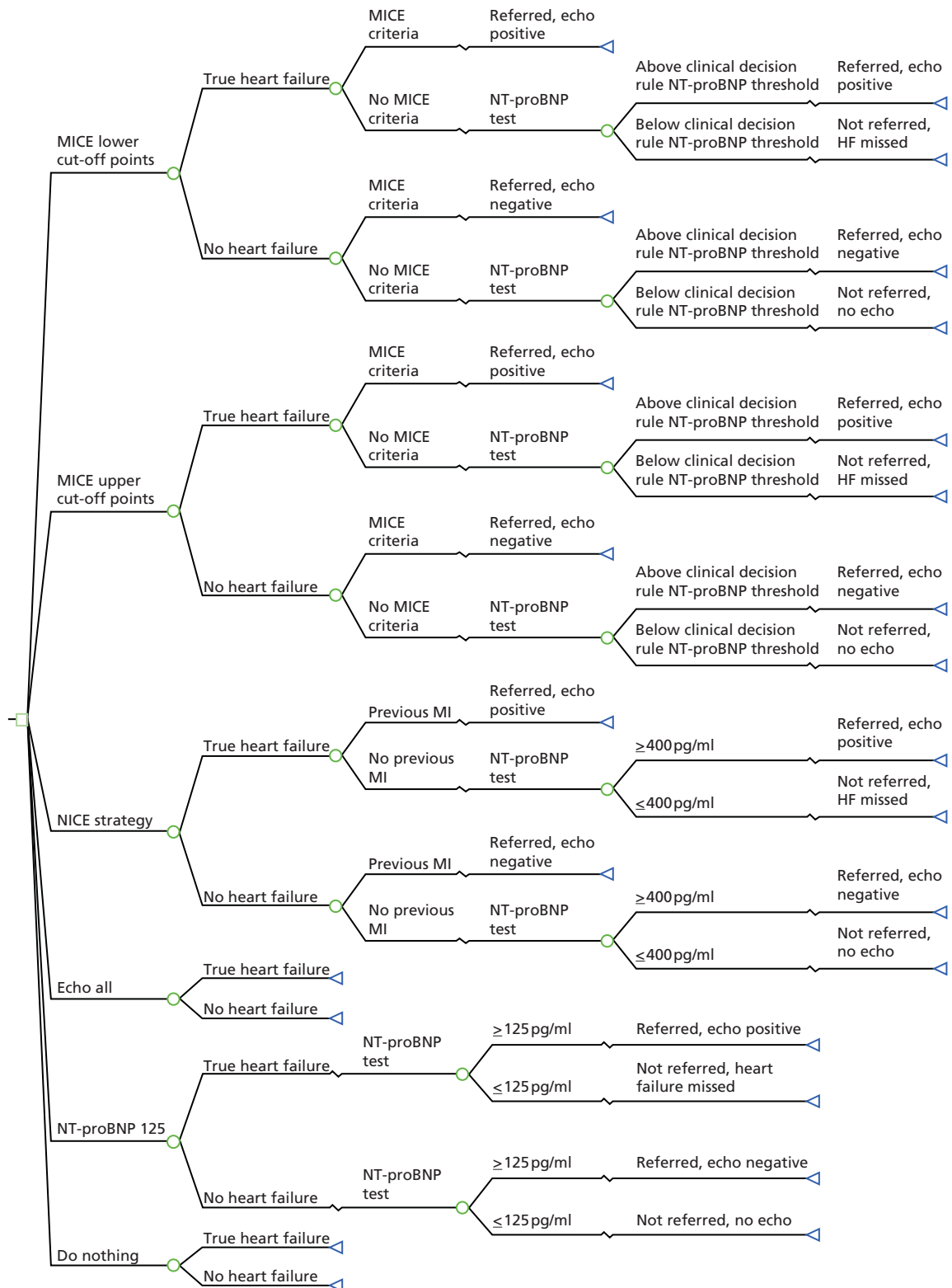


FIGURE 4 Decision tree of the different diagnostic strategies for patients presenting with heart failure symptoms in primary care. Echo, echocardiogram; HF, heart failure; MI, myocardial infarction.

diagnosis for heart failure). The branch probabilities were estimated from the REFER data set. To ensure correct representation of the statistical uncertainty in the model, patients were categorised into groups on the principle that two patients would be in the same group if and only if they followed the same pathway in all strategies considered. For example, consider four patients of the same sex and with the same clinical signs, with no previous history of myocardial infarction but with NT-proBNP levels of 110 pg/ml (A),

220 pg/ml (B), 330 pg/ml (C) and 440 pg/ml (D). Under the NT-proBNP 125 strategy, patients B, C and D would be referred for echocardiography but patient A would not, whereas under the NICE strategy only patient D would be referred. However, no strategy in the model has a cut-off point between 220 pg/ml and 330 pg/ml and so patients B and C can be placed in the same group. *Table 7* shows the breakdown of patients into groups based on this principle.

Effects of early heart failure diagnosis

The methodology behind the benefits of each strategy was drawn largely from a previous Heart Failure Health Technology Assessment report.⁴⁰ It was assumed that a confirmed heart failure diagnosis leads to patients being initiated on heart failure drug treatment if they have HFREF. Current trials do not suggest a survival advantage of ACE inhibitors or beta-blockers for patients with HFPEF.¹⁵ Thus, the benefits from early detection of heart failure were weighted by the number of heart failure patients in the population (11.5%) who have a reduced ejection fraction.

TABLE 7 Patient clinical characteristics

Clinical parameters	Heart failure (N = 104), n (%)	No heart failure (N = 200), n (%)
Patients with a previous MI and a NT-proBNP level of < 125 pg/ml	1 (1.0)	10 (5.0)
Patients with a previous MI and a NT-proBNP level of ≥ 125 pg/ml	16 (15.4)	7 (3.5)
Patients with basal crepitations without a previous MI and with a NT-proBNP level of < 125 pg/ml	0 (0.0)	5 (2.5)
Patients with basal crepitations without a previous MI and with a NT-proBNP level between 125 pg/ml and 399 pg/ml	0 (0.0)	4 (2.0)
Patients with basal crepitations without a previous MI and with a NT-proBNP level of ≥ 400 pg/ml	3 (2.9)	2 (1.0)
Male patients with ankle oedema without basal crepitations or a previous MI and with a NT-proBNP level of < 125 pg/ml	0 (0.0)	22 (11.0)
Male patients with ankle oedema without basal crepitations or a previous MI and with a NT-proBNP level between 125 pg/ml and 399 pg/ml	4 (3.8)	17 (8.5)
Male patients with ankle oedema without basal crepitations or a previous MI and with a NT-proBNP level of ≥ 400 pg/ml	22 (21.2)	5 (2.5)
Female patients without ankle oedema and without the MICE criteria and with a NT-proBNP level of < 125 pg/ml	0 (0.0)	9 (4.5)
Female patients without ankle oedema and without the MICE criteria and with a NT-proBNP level between 125 pg/ml and 399 pg/ml	0 (0.0)	6 (3.0)
Female patients without ankle oedema and without the MICE criteria and with a NT-proBNP level between 400 pg/ml and 619 pg/ml	0 (0.0)	2 (1.0)
Female patients without ankle oedema and without the MICE criteria and with a NT-proBNP level between 620 pg/ml and 1059 pg/ml	1 (1.0)	0 (0.0)
Female patients without ankle oedema and without the MICE criteria and with a NT-proBNP level of ≥ 1060 pg/ml	0 (0.0)	0 (0.0)
Female patients with ankle oedema and without the MICE criteria and with a NT-proBNP level of < 125 pg/ml	4 (3.8)	43 (21.5)
Female patients with ankle oedema and without the MICE criteria and with a NT-proBNP level between 125 pg/ml and 189 pg/ml	3 (2.9)	15 (7.5)

continued

TABLE 7 Patient clinical characteristics (continued)

Clinical parameters	Heart failure (N = 104), n (%)	No heart failure (N = 200), n (%)
Female patients with ankle oedema and without the MICE criteria and with a NT-proBNP level between 190 pg/ml and 399 pg/ml	4 (3.8)	31 (15.5)
Female patients with ankle oedema and without the MICE criteria and with a NT-proBNP level between 400 pg/ml and 519 pg/ml	4 (3.8)	5 (2.5)
Female patients with ankle oedema and without the MICE criteria and with a NT-proBNP level of ≥ 520 pg/ml	30 (28.8)	0 (0.0)
Male patients without ankle oedema and without the MICE criteria and with a NT-proBNP level of < 125 pg/ml	1 (1.0)	9 (4.5)
Male patients without ankle oedema and without the MICE criteria and with a NT-proBNP level between 125 pg/ml and 389 pg/ml	2 (1.9)	7 (3.5)
Male patients without ankle oedema and without the MICE criteria and with a NT-proBNP level between 390 pg/ml and 399 pg/ml	1 (1.0)	0 (0.0)
Male patients without ankle oedema and without the MICE criteria and with a NT-proBNP level between 400 pg/ml and 659 pg/ml	2 (1.9)	0 (0.0)
Male patients without ankle oedema and without the MICE criteria and with a NT-proBNP level of ≥ 660 pg/ml	6 (5.8)	1 (0.5)

The type of heart failure drug and dose used were obtained from a UK retrospective cohort study of patients treated for heart failure in primary care; 36.6% of patients were treated with beta-blockers, 58.9% with ACE inhibitors and 13.4% with angiotensin receptor blockers (ARBs).⁴⁰ Patients can be on more than one heart failure drug therapy; to determine the proportion of patients on ACE inhibitors but not on beta-blockers (not reported in the study) we took a simple average of the upper bound and lower bound range of the possible values. This gave a value of 36.3% for patients on ACE inhibitors but not on a beta-blocker.

The initiation of heart failure treatment after correct early detection was assumed to have a survival benefit for the heart failure patient compared with a missed heart failure diagnosis.

Effects of delayed heart failure diagnosis

A missed heart failure diagnosis was assumed to delay the diagnosis by 6 months. The delayed diagnosis patients were assumed to incur a further GP visit and an echocardiogram scan to confirm the diagnosis. We assumed that the delayed diagnosis patients had the same prognosis as untreated patients for the first 6 months. After 6 months, the delayed diagnosis patients were put on treatment and thereafter had the same survival probability as someone who was already on treatment.

Treatment effects on mortality

As in a previous model,⁴¹ we took the survival data from patients with heart failure in the Framingham Heart Study to estimate prognosis in the absence of drug therapy. An individual patient data meta-analysis indicated that beta-blockers compared with placebo gave a hazard ratio of 0.73 (95% CI 0.67 to 0.80) for all-cause mortality; 95% of these beta-blocker patients were also on ACE inhibitors/ARBs.⁴² The results of a systematic review showed that the mortality odds ratio for patients taking ACE inhibitors compared with a placebo was 0.80 (95% CI 0.74 to 0.87).⁴³ A Cochrane review found no significant effect of ARBs on mortality and we made a conservative assumption based on this analysis that patients on ARBs have the same survival rate as untreated patients.⁴⁴

The survival data and the drug efficacies were used to plot sex-specific survival curves for the untreated patients and patients on different drug therapies (assuming no temporal changes in drug efficacy). The survival curves were extended beyond the 10-year survival data by linear extrapolation to achieve a lifetime horizon. *Tables 8 and 9* depict the estimated probabilities associated with treatment and no treatment of men and women, respectively. The patients who were put on treatment earlier after correct detection of heart failure (treat early) and the patients with a delayed diagnosis (treat late) were weighted by the proportion of patients on the different drug therapies.

TABLE 8 Estimated undiscounted survival probabilities for male patients with heart failure on and off therapy

Years	Untreated	ACE inhibitors	Beta-blockers	Treated early	Treated late
0	1.00	1.00	1.00	1.00	1.00
0.25	0.73	0.77	0.79	0.77	0.73
0.5	0.68	0.72	0.75	0.72	0.68
1	0.57	0.62	0.66	0.62	0.59
2	0.46	0.52	0.57	0.52	0.49
5	0.25	0.29	0.36	0.31	0.29
10	0.11	0.13	0.20	0.15	0.14
11	0.08	0.10	0.16	0.12	0.11
12	0.05	0.07	0.12	0.08	0.07
13	0.03	0.03	0.07	0.04	0.04
14	0.00	0.00	0.02	0.01	0.00
15	0.00	0.00	0.00	0.00	0.00

TABLE 9 Estimated undiscounted survival probabilities for female patients with heart failure on and off therapy

Years	Untreated	ACE inhibitors	Beta-blockers	Treated early	Treated late
0	1.00	1.00	1.00	1.00	1.00
0.25	0.72	0.76	0.79	0.76	0.72
0.5	0.69	0.74	0.77	0.74	0.69
1	0.64	0.69	0.72	0.69	0.65
2	0.56	0.61	0.65	0.61	0.58
5	0.38	0.43	0.49	0.44	0.41
10	0.21	0.25	0.32	0.26	0.25
11	0.18	0.21	0.28	0.23	0.21
12	0.14	0.17	0.24	0.19	0.17
13	0.11	0.13	0.20	0.15	0.14
14	0.07	0.09	0.15	0.11	0.10
15	0.04	0.05	0.10	0.06	0.06
16	0.01	0.01	0.02	0.01	0.01
17	0.00	0.00	0.00	0.00	0.00

Figures 5 and 6 show the discounted survival curves of patients treated early and treated late for men and women, respectively. The survival benefit from early treatment is the area between the treated early and treated late curves. These curves were discounted at 3.5% to obtain discounted life-years gained from early treatment. The sex-specific survival benefit was then combined by the weighted average number of men and women in the data set with heart failure.

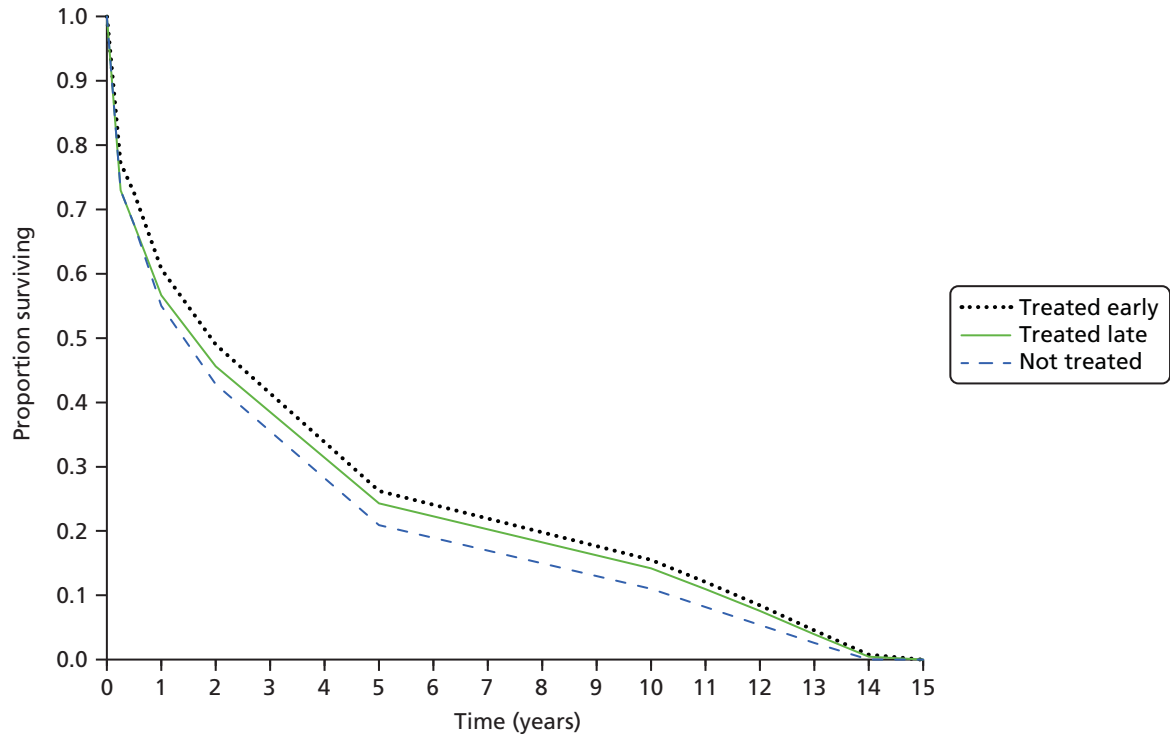


FIGURE 5 Discounted survival curves for male patients on and off treatment.

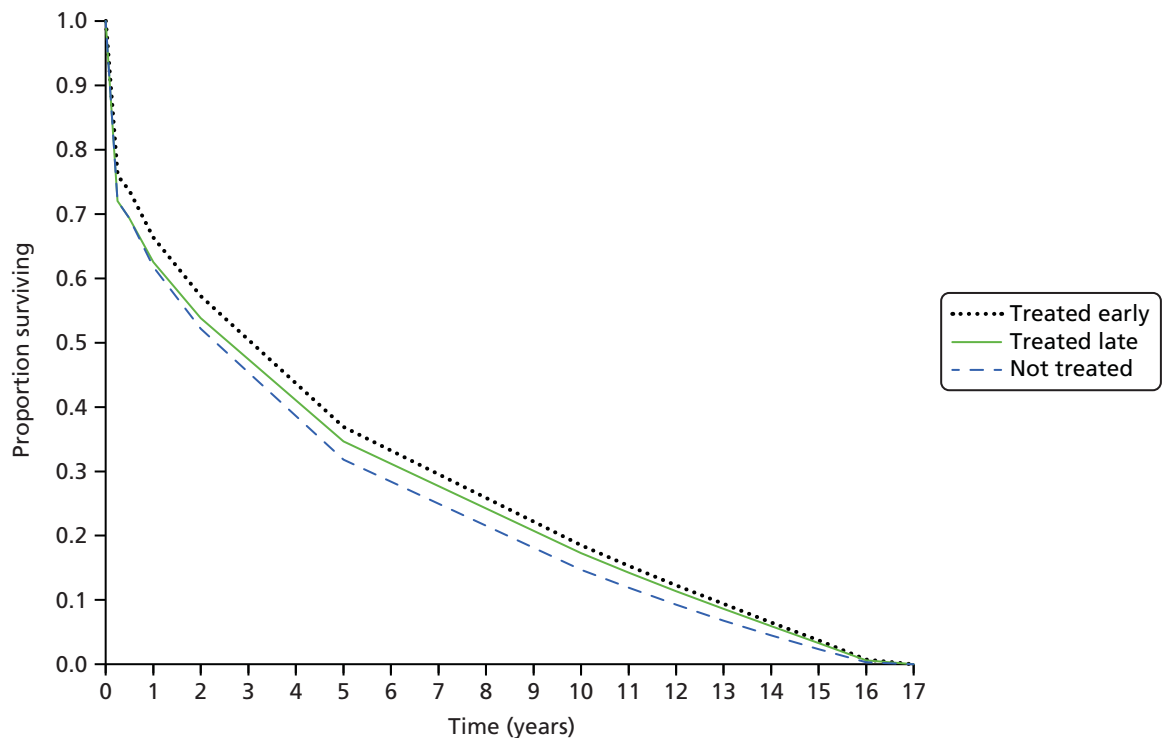


FIGURE 6 Discounted survival curves for female patients on and off treatment.

Patients with heart failure in the REFER data set who answered the 6-month follow-up questionnaire gave a mean EQ-5D score of 0.615 (SD 0.31) (Table 10). The discounted quality-adjusted life-year (QALY) gain can be estimated from the product of the EQ-5D and the discounted life-years gained. Current trials do not suggest a survival advantage of ACE inhibitors or beta-blockers for patients with HFPEF.⁴⁵ Thus, the survival gain from early detection of heart failure is weighted by the number of heart failure patients in the population (11.5%) who have a reduced ejection fraction (HFREF).

Early heart failure treatment effect on hospital admissions

The early detection of heart failure also has the benefit of a reduction in hospitalisations. Patients on beta-blockers compared with placebo have a hazard ratio of 0.71 (95% CI 0.65 to 0.77) for first heart failure-related hospital admission. Similarly, treatment with ACE inhibitors reduces the likelihood of a hospital admission, with an odds ratio of 0.67 (95% CI 0.61 to 0.74).⁴²

For the purpose of the model we were interested in the increased rate of admission to hospital for an untreated population because of a 6-month delay in heart failure diagnosis. In a treated population following a new diagnosis of heart failure, 59% of people in a south London heart failure incidence study had a subsequent hospital admission over a span of 19 months.⁴⁶ This would result in a 19-month 'hospital-free survival' probability of 41% (odds 0.70). We assumed that the 'hospital-free survival' curve followed the same pattern (constant odds ratio) as the survival curve for untreated males. From this, the estimated survival at 19 months was 50.6% (odds 1.02 in favour of survival), with survival at 6 months of 67.7% (odds 2.09). This gave estimated odds of no admission at 6 months if treated of 1.42, which gave a probability of no admission of 58.7%, hence a probability of admission if treated of 41.3%.

Using the odds ratio of hospital admission with ACE inhibitor treatment, the estimated rate of admission within 6 months if untreated is $0.41/0.67 = 0.62$. Using the hazard ratio for treatment with beta-blockers (0.71), assuming a constant hazard over the 6 months, the estimated rate of admission within 6 months if untreated is 71.2%. A simple average was taken to calculate the estimated rate of admission with 6 months if untreated (66.4%). Thus, the increased number of hospitalisation cases if untreated is 25.1% (41.3–66.4%). This was then weighted by the number of HFREF patients in the population as only these patients receive a prognostic benefit from treatment.

TABLE 10 Health utility states for the REFER patient population

Health utility states	Heart failure (<i>n</i> = 104)	No heart failure (<i>n</i> = 200)
EQ-5D 6-month utility		
Number of observations	80 (missing 24)	158 (missing 42)
Mean	0.615	0.556
Median	0.690	0.620
SD	0.305	0.329
EQ-5D 12-month utility		
Number of observations	76 (missing 28)	147 (missing 53)
Mean	0.634	0.553
Median	0.675	0.660
SD	0.287	0.333

Costs associated with the diagnostic strategies

A list of costs and their sources are shown in *Table 11*. The cost of an avoidable heart failure hospital admission was calculated from weighted average 2013/14 reference costs of heart failure admissions.⁴⁸ The cost of an echocardiography referral, consisting of a simple echocardiogram and a consultant-led outpatient first attendance, was taken from the Payment by Results mandatory tariff 2013/14.⁴⁷

The cost of a NT-proBNP test was taken from a NICE costing report,⁷ which based it on a provider contract price. As there was no mention of a cost year, for inflation purposes the NT-proBNP test cost was assumed to have been obtained when the report was published (2010). This cost was then inflated to 2013/14 prices using the Hospital & Community Health Service (HCHS) pay and prices index.⁵⁰ The early treatment heart failure drug cost was estimated from the drug resource usage combined with drug prices from the *British National Formulary* (BNF).⁴⁹ The heart failure drug therapy prescription mix was taken from a 5-year retrospective cohort study of general practices involved in the Doctors Independent Network (DIN-LINK) database.⁴¹ As the typical drug dose for a heart failure patient is variable, we pragmatically assumed that the same percentage of patients reached the target dose, as defined by the 2005 ESC guidelines⁵¹ and the remainder achieved half the target dose. When the drug listed was not recommended by the 2012 ESC guidelines,⁶ the target dose indication for heart failure in the BNF was used. The drug therapy costs were weighted by the percentage of patients on the different regimes (ACE inhibitors, beta-blockers and ARBs). For the base-case analysis, the cost of the generic drugs for the first 6 months was used (£10); the cost of the equivalent drug therapy with branded drugs was substituted in the sensitivity analysis (£30).

The results are presented as the total costs and effects of each strategy. Effectiveness was measured in QALYs. The total costs and outcomes of each strategy are ordered by increasing effectiveness. Incremental cost-effectiveness ratios (ICERs) were calculated from the difference in costs and effects between two options. Options that are more costly and less effective (dominated) are excluded from consideration. Likewise, options that suffer from extended dominance are removed from consideration. Extended dominance occurs when an option would be dominated compared with a mixed option of two other strategies.

The incremental analysis was designed to generate the cost per additional QALY gained of using one diagnostic strategy over another. Cost-effectiveness was assessed in relation to the lower NICE cost-effectiveness threshold of £20,000 per QALY gained.

To assess model robustness, deterministic and probabilistic sensitivity analysis was performed.

TABLE 11 Cost parameters

Cost parameters	Value (£)	Source
NT-proBNP testing	30	2010 NICE costing report ⁷
Echocardiograph plus specialist assessment	241	Payment by Results 2013/14 tariff (currency code RA60Z and treatment function 320 Cardiology) ⁴⁷
Avoidable heart failure hospital admission	2107	Weighted average 2013/14 reference costs of heart failure admission (currency codes EB03A–EB03E) ⁴⁸
Early diagnosis heart failure generic drug therapy	10	<i>BNF</i> ⁴⁹ combined with typical heart failure doses for heart failure patients in the UK ⁴²
Early diagnosis heart failure branded drug therapy	30	<i>BNF</i> ⁴⁹ combined with typical heart failure doses for heart failure patients in the UK ⁴²
Discount rate for costs and outcomes	3.5%	NICE ³⁹
GP visit	46	Curtis ⁵⁰

Deterministic sensitivity analysis

For the deterministic sensitivity analysis the following scenarios were explored to test the robustness of the base-case results:

- doubling and halving the cost of a BNP test
- altering the drug efficacies to their lower and upper CIs
- substituting in branded drug prices for drug therapy prices
- increasing the proportion of HFREF patients from 12% to 24%, 50% and 100%.

Probabilistic sensitivity analysis

For the probabilistic sensitivity analysis, distributions were attached to the clinical parameters, drug efficacies and heart failure utility (*Table 12*). The model was run for 10,000 iterations and the results presented as a cost-effectiveness acceptability frontier (CEAF). The CEAF shows the uncertainty associated with the optimal diagnostic option across a range of different cost-effectiveness thresholds.

Positive count data in groups formed the parameters for the Dirichlet distribution (see *Table 12*). On each replication a vector of probabilities was sampled from the appropriate distribution. When there were no count data (no patients) in a patient group, we assumed that the appropriate combination was impossible and no distribution was attached to clinical parameters with zero occurrences in the data set. Treating the probability as fixed for these empty patient groups will slightly underestimate the uncertainty in the model rather than the alternative of adding an occurrence and positively biasing the amount of occurrences for these groups.⁵²

TABLE 12 Parameter distributions for the probabilistic sensitivity analysis

Parameters	Distribution	Parameter estimates ^a	Source
Patient's true diagnosis	Beta	$\alpha = 104, \beta = 200$	REFER data set
Heart failure utility (EQ-5D)	Beta	$\alpha = 0.95, \beta = 0.59$	REFER data set
Beta-blocker effect on mortality	Log-normal	$\mu = -0.31, \sigma = 0.05$	Patient-level meta-analysis by Kotecha <i>et al.</i> ⁴²
Beta-blocker effect on hospitalisation risk	Log-normal	$\mu = -0.34, \sigma = 0.04$	Patient-level meta-analysis by Kotecha <i>et al.</i> ⁴²
ACE inhibitor effect on mortality	Log-normal	$\mu = -0.22, \sigma = 0.04$	Flather <i>et al.</i> ⁴³
ACE inhibitor effect on hospitalisation risk	Log-normal	$\mu = -0.40, \sigma = 0.05$	Flather <i>et al.</i> ⁴³
Patients on each drug therapy (beta-blockers, ACE inhibitors, other)	Dirichlet	$(\alpha_1, \alpha_2, \alpha_3) = (3403, 3378, 1908)$	Calvert <i>et al.</i> ⁴¹

^a For the log-normal distributions, μ is the mean and σ is the SD of the underlying normal distribution, which gives the logarithm of the model parameter.

Chapter 7 The REFER study health economic evaluation: results

The total costs and effectiveness of each strategy under base-case assumptions and parameters are shown in *Table 13*. Both CDRs of MICE are excluded because of being dominated outright or weakly dominated (extended dominance). Given a willingness-to-pay threshold of £20,000 per QALY, only the NICE strategy would be cost-effective. The mean total costs and effects (measured in QALYs) of each strategy are depicted on a cost-effectiveness plane in *Figure 7*.

TABLE 13 Base-case results

Strategy	Proportion of true heart failure detected (%)	Proportion of not heart failure ruled out (%)	Cost (£)	QALY gain vs. 'do nothing'	ICER (£ per QALY)
Do nothing	0.00	100.00	119	–	
NICE	78.85	63.50	142	0.0051	(Dominated)
MICE upper cut-off points	81.73	84.00	167	0.0050	4400
MICE lower cut-off points	90.38	45.50	191	0.0057	(Extended dominance)
NT-proBNP 125	94.23	49.00	196	0.0059	69,000
Echo all	100.00	0.00	241	0.0063	125,100

Echo, echocardiogram.

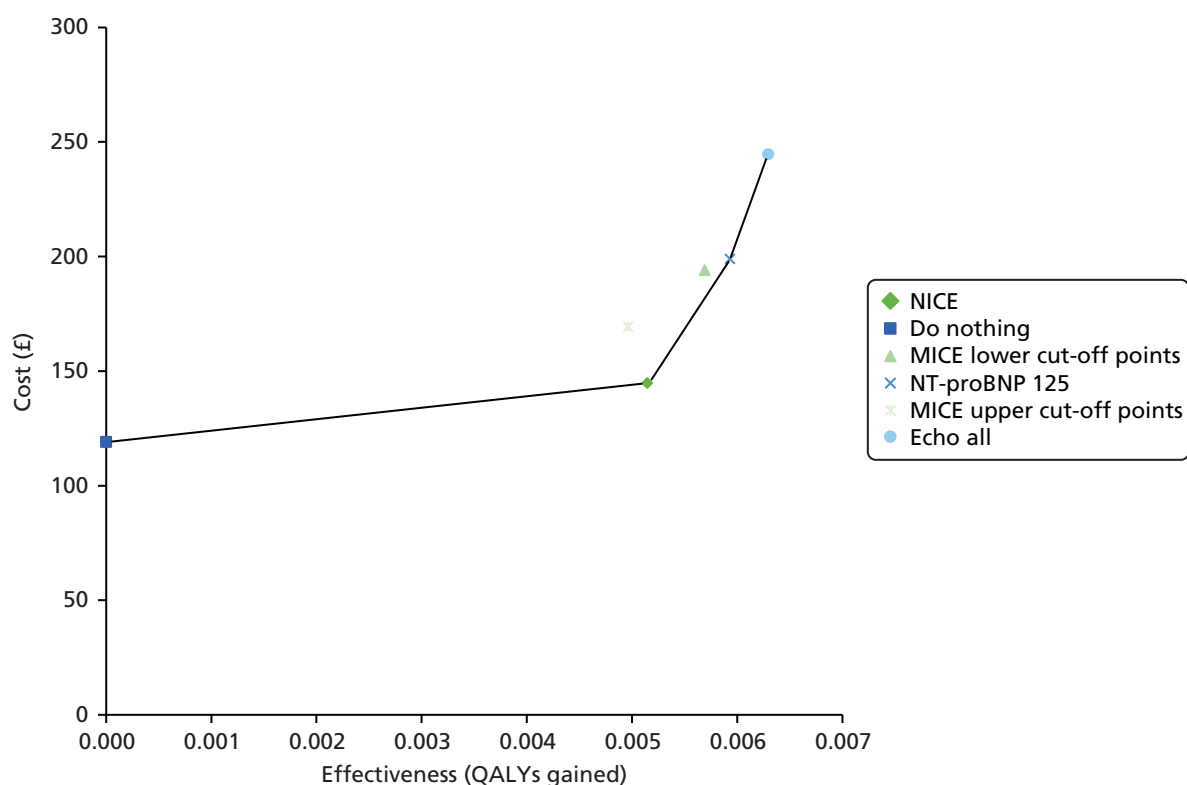


FIGURE 7 Cost-effectiveness plane. Echo, echocardiogram.

The MICE upper cut-off point strategy is dominated by the NICE strategy. Similarly, the MICE lower cut-off point strategy is excluded by extended dominance between the NICE strategy and the NT-proBNP 125 strategy. ICERs (rounded to the nearest multiple of £100 per QALY) have been given relative to the previous non-dominated strategy.

Deterministic sensitivity analysis results

The results of the sensitivity analysis are shown in *Table 14*. The MICE cut-off point options are excluded from the table as they remained dominated in each scenario. As in the base-case results, the NICE strategy remained the most cost-effective option for each sensitivity analysis scenario except when the proportion of HFREF changed to $\geq 50\%$. When the proportion of HFREF patients is 50%, the NT-proBNP 125 strategy becomes cost-effective, and when this proportion reaches 100% it becomes cost-effective to refer all patients for immediate echocardiography. Threshold analysis indicated that the NT-proBNP 125 strategy becomes cost-effective compared with the NICE strategy when the proportion of HFREF rises above 36%.

Probabilistic sensitivity analysis results

Figure 8 illustrates the overall uncertainty related to the optimal decision across a range of plausible willingness-to-pay values, where willingness to pay is measured in cost per additional QALY. Taking NICE's willingness to pay for an additional QALY as £20,000, the likelihood of the NICE strategy being the optimal option (i.e. highest net monetary benefits) is 99.9%. As the willingness-to-pay threshold increases beyond £68,000, the NT-proBNP 125 option becomes more likely to be the optimal option.

TABLE 14 Sensitivity analysis scenario results

Scenario ^a	Cost per additional QALY (£)		
	NICE strategy vs. do nothing strategy	NT-proBNP 125 strategy vs. NICE strategy	Echo all strategy vs. NT-proBNP 125 strategy
Base-case results	4400	69,000	125,100
Double NT-proBNP test cost	9600	73,300	42,100
Halve NT-proBNP cost	1800	66,800	166,600
Branded drug price therapy	9700	68,900	125,200
Higher drug efficacy for mortality	3300	52,000	94,200
Lower drug efficacy for mortality	6600	104,200	189,100
Proportion of HFREF patients doubled to 24%	600	32,900	60,900
Proportion of HFREF patients increased to 50%	Dominates do nothing strategy	13,300	26,400
Proportion of HFREF patients increased to 100%	Dominates do nothing strategy	5000	11,600

Echo, echocardiogram.
 a The MICE cut-off point options were excluded from the table as they remained dominated in each scenario.

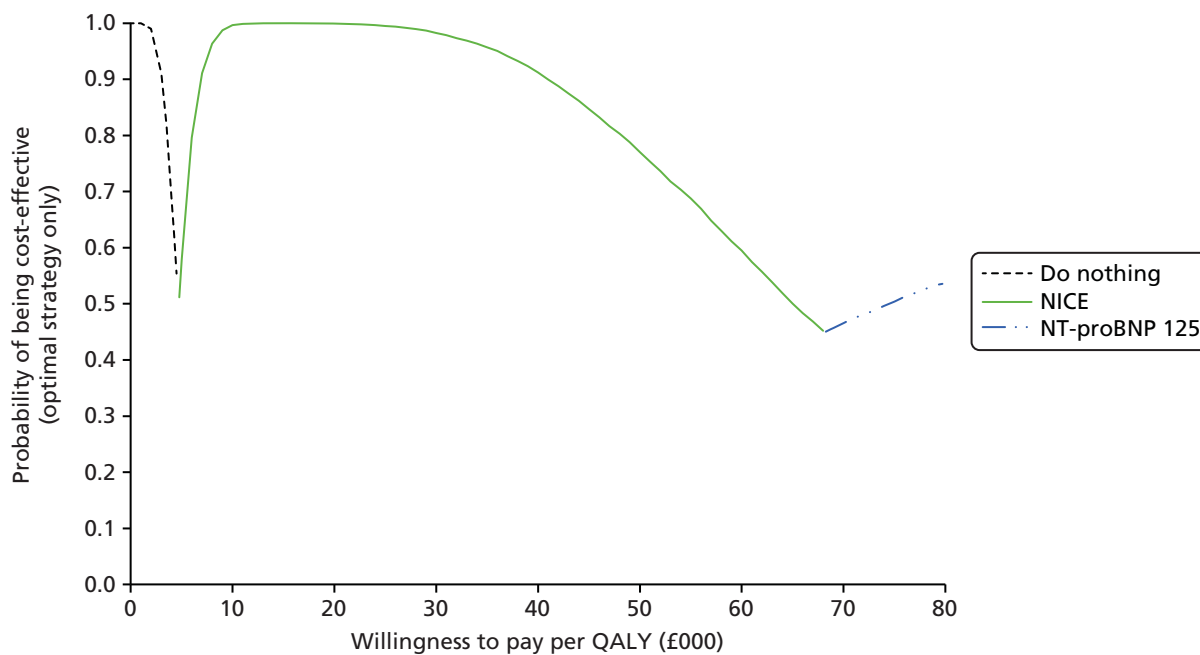


FIGURE 8 Cost-effectiveness acceptability frontier showing the optimal diagnostic strategy across a range of willingness-to-pay thresholds.

Chapter 8 The REFER study health economic evaluation: discussion

Summary of findings

The results indicate that, for the REFER population, the current NICE guidelines on diagnosing heart failure provide the most cost-effective strategy compared with the MICE CDRs and other diagnostic strategies. The MICE CDRs with upper and lower cut-off points were excluded from the base-case and the sensitivity analyses because of dominance by the other strategies. The cost-effectiveness results were robust to deterministic and probabilistic sensitivity analyses.

The modest gains in effectiveness for each strategy, even with a lifetime horizon, can be largely explained by the patient characteristics in the REFER data set. The benefits of early detection of heart failure and the subsequent prognostic window of improvement (treated early vs. treated late) used in the model are limited as only 12% of the heart failure participants in the REFER study were found to have a reduced left ventricular ejection fraction. In this analysis, survival benefit and reduced hospitalisation risk from early detection are restricted to this subgroup, that is, we model no benefits from earlier detection in the 88% of the REFER population detected with non-HFREF. This is an appropriate conservative assessment but it is unlikely that these patients will derive zero benefit from detection.

In a sensitivity analysis scenario, increasing the proportion of HFREF patients raised the total QALYs produced by each diagnostic strategy because of the higher rewards of correct early detection. Most population surveys suggest that approximately half of patients with heart failure have HFREF,⁵³ in contrast to the 12% detected in this study, and in this scenario the dominant cost-effective strategy is the reduced NT-proBNP cut-off point of 125 pg/ml for referral for echocardiography rather than the NICE level of 400 pg/ml. The echo all strategy becomes the most cost-effective option when we assume that all of the heart failure patients will receive a prognostic benefit from treatment. However, it must be acknowledged that there may be practical barriers to the implementation of such a strategy. Delays to echocardiography may potentially offset the advantage of early detection and referring all patients for an echocardiogram will put pressure on local diagnostic services. The cost used for a NT-proBNP test was an inflated cost from 2010.

Strengths and limitations

Since NT-proBNP testing came into more general use in the health-care system, the cost per test may have dropped and so the cost-effectiveness of natriuretic peptide testing may be greater than the estimates presented. This was explored in the sensitivity analysis by halving the cost of a NT-proBNP test, which resulted in the NT-proBNP 125 strategy becoming more cost-effective than the echo all and do nothing strategies.

The heart failure therapy drug mix for patients used in the economic evaluation relates to a 2009 publication.⁴¹ As it is very likely that there are now higher rates of drug use, the proportion of patients on ACE inhibitors and beta-blockers may be more than assumed and so the benefits of diagnosis may be greater.

The main strength of this analysis is that the diagnostic accuracy of the various strategies tested was calculated based on a consistent primary data set. This was reinforced by the process used in the clinical study to develop an appropriate gold standard against which to compare imperfect diagnostic strategies. The main limitations are that the REFER data set may not be typical of patients in other geographical areas and that assumptions have had to be made in projecting the lifetime costs and outcomes of a correct diagnosis.

Comparison with existing literature

This is the first cost–utility analysis of heart failure diagnosis in this patient population, making comparison with other studies difficult. A cost–consequence analysis, in which only costs are considered, suggested that a diagnostic strategy using a BNP referral level of 100 pg/ml had a higher test accuracy and was only marginally more expensive than a consultant-led ECG strategy.⁵⁴ Other studies^{55,56} have looked at the costs of using NT-proBNP testing as a means to rule out heart failure and to reduce the levels of echocardiography referrals but these do not come close to answering the question of the optimal cost-effective strategy and also do not place a cost on a missed heart failure diagnosis.

Implications for practice and research recommendations

The findings of this study suggest that further research around the cost-effectiveness of different natriuretic peptide thresholds in specific heart failure populations (HFREF vs. HFPEF) is needed to determine the optimal strategy for diagnosis in primary care. The organisation of health-care systems globally is associated with differing costs of and access to investigations. Cost-effectiveness analyses must be carried out in the context of individual health-care economies to determine the most cost-effective local strategy.

Conclusions

Overall, this analysis provides evidence based on primary data that the current strategy recommended by NICE for the diagnosis of heart failure is appropriate, that is, patients presenting with symptoms suggestive of heart failure should be referred straight for echocardiography if they have a history of myocardial infarction or if their NT-proBNP level is ≥ 400 pg/ml. However, based on sensitivity analysis, as the proportion of patients with HFREF increases from the 12% seen in the REFER study, it becomes more cost-effective to change the NT-proBNP threshold from 400 pg/ml to 125 pg/ml.

Acknowledgements

The REFER study was funded by a grant from the National Institute for Health Research (NIHR) Efficacy and Mechanism Evaluation (EME) programme. The study received approval from the Midlands Research Ethics Committee (09/H1207/121). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

FDRH is part funded by the NIHR School for Primary Care Research; NIHR Oxford Biomedical Research Centre; NIHR Collaboration for Leadership in Applied Health Research and Care, Oxford; and Harris Manchester College, University of Oxford. CT is funded by a NIHR Academic Clinical Lectureship at the University of Oxford.

The authors would also like to acknowledge Ms Marites Derit, echocardiography technician, and the trial steering committee – Professor Kelvin Jordan (chairperson), Professor Tom Quinn, Professor Christian Mallen and Barry Clark (patient representative). Thanks also to the 28 recruiting practices: Grange Hill Medical Centre, Dovecote Medical Practice, Bellevue Medical Centre, Burbury St Surgery, Fernley Medical Centre, The Oaks Medical Centre, Sutton Park Surgery, Balsall Heath Medical Centre, Stockland Green Primary Care Centre, Ridgeacre Surgery, Grove Road and Shirley Medical Centre, Maypole Health Centre, Riverbrook Medical Centre, West Heath Surgery, Vicarage Road Surgery, Woodgate Valley Health Centre, Greenridge Health Centre, Fernbank Medical Centre, Downsfield Surgery, Laurie Pike Health Centre, Shanklin House Surgery, Apollo Surgery, Kingstanding Health Centre, Tudor Practice, Highgate Health Centre, Bourneville Surgery, Eden Court Medical Practice and Yardley Wood Medical Practice.

The REFER study investigators

Investigators also included Professor Martin R Cowie (Professor of Cardiology, Faculty of Medicine, National Heart and Lung Institute, Imperial College London, UK), Dr Russell Davis (Consultant Cardiologist, Department of Cardiology, Sandwell and West Birmingham Hospitals, UK), Professor Jon Deeks (Professor of Biostatistics, Institute of Applied Health Research, University of Birmingham, UK), Professor Jonathan Mant (Professor of Primary Care, Department of Public Health and Primary Care, University of Cambridge, UK), Dr Deborah McCahon (Lecturer in Primary Care, Institute of Applied Health Research, University of Birmingham, UK), Professor Theresa McDonagh (Professor of Heart Failure, Department of Cardiology, King's College Hospital, UK), Dr George Sutton (Consultant Cardiologist, Faculty of Medicine, National Heart and Lung Institute, Imperial College London, UK) and Dr Lynda Tait (Senior Research Fellow, School of Health Sciences, University of Nottingham, UK).

Contributions of authors

Clare J Taylor (NIHR Academic Clinical Lecturer and GP) was a grant applicant, provided clinical input throughout the study and first drafted the main results paper for the REFER study.

Mark Monahan (Health Economist) was the health economist on the study and drafted the health economic evaluation paper for the REFER study.

Andrea K Roalfe (Statistician) conceived of and designed the study, was a grant applicant and study statistician and contributed to writing the draft papers.

Pelham Barton (Reader in Mathematical Modelling) was a grant applicant and senior health economist on the study and contributed to writing the health economic evaluation paper for the REFER study.

Rachel Iles (Research Fellow) was the study co-ordinator throughout the study and commented on the draft papers.

FD Richard Hobbs (Professor of Primary Care and GP) had the original idea, conceived of and designed the study, was a grant applicant and principal investigator, provided clinical input throughout the study, compiled the report, co-drafted the papers and is corresponding author and guarantor.

Contributions of others

Martin Cowie (Professor of Cardiology) conceived of and designed the study, was a grant applicant, was on the expert panel and commented on the draft papers.

Russell Davis (Cardiologist) was on the expert panel and commented on the draft papers.

Jon Deeks (Professor of Biostatistics) was a grant applicant, provided statistical advice and commented on the draft papers.

Jonathan Mant (Professor of Primary Care) conceived of and designed the study, was a grant applicant and commented on the draft papers.

Deborah McCahon (Lecturer in Primary Care) was study co-ordinator for part of the study and commented on the draft papers.

Theresa McDonagh (Professor of Cardiology) was on the expert panel.

George Sutton (Cardiologist) was on the expert panel.

Lynda Tait (Research Fellow) conceived of and designed the study, was a grant applicant, was senior study co-ordinator for most of the study and commented on the draft papers.

Publications

One paper has been published and two further papers have been submitted to peer-reviewed journals. Standards for Reporting of Diagnostic Accuracy (STARD) and Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklists were completed for the REFER study and health economics analysis papers, respectively, before submission.

Tait L, Roalfe AK, Mant J, Cowie M, Deeks JJ, Iles R, *et al*. The REFER (REfer for Echocardiogram) protocol: a prospective validation of a clinical decision rule, NT-proBNP, or their combination, for the diagnosis of heart failure in primary care. Rationale and design. *BMC Cardiovasc Disord* 2012;**12**:97.

Taylor CJ, Roalfe AK, Iles R, Hobbs FR, REFER investigators, Barton P, *et al*. Primary care REFerral for Echocardiogram (REFER) in heart failure: a diagnostic accuracy study. *Br J Gen Pract* 2017;**67**:e94–102.

Monahan M, Barton P, Taylor CJ, Roalfe AK, Hobbs FDR. MICE or NICE? An economic evaluation of clinical decision rules in the diagnosis of heart failure in primary care [published online ahead of print 2 March 2017]. *Int J Cardiol* 2017.

Data sharing statement

The original data used in the analysis are available from the corresponding author, Professor FD Richard Hobbs (Nuffield Department of Primary Care Health Sciences, Radcliffe Primary Care Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG; richard.hobbs@phc.ox.ac.uk).

References

1. Taylor CJ, Roalfe AK, Iles R, Hobbs FR, REFER investigators, Barton P, *et al.* Primary care REFerral for Echocardiogram (REFER) in heart failure: a diagnostic accuracy study. *Br J Gen Pract* 2017;**67**:e94–102. <http://dx.doi.org/10.3399/bjgp16X688393>
2. Monahan M, Barton P, Taylor CJ, Roalfe AK, Hobbs FDR. MICE or NICE? An economic evaluation of clinical decision rules in the diagnosis of heart failure in primary care [published online ahead of print 2 March 2017]. *Int J Cardiol* 2017. <http://dx.doi.org/10.1016/j.ijcard.2017.02.149>
3. Zapka JG, Moran WP, Goodlin SJ, Knott K. Advanced heart failure: prognosis, uncertainty, and decision making. *Congest Heart Fail* 2007;**13**:268–74. <http://dx.doi.org/10.1111/j.1527-5299.2007.07184.x>
4. Stewart S, Jenkins A, Buchan S, McGuire A, Capewell S, McMurray JJ. The current cost of heart failure to the National Health Service in the UK. *Eur J Heart Fail* 2002;**4**:361–71. [http://dx.doi.org/10.1016/S1388-9842\(01\)00198-2](http://dx.doi.org/10.1016/S1388-9842(01)00198-2)
5. Taylor CJ, Roalfe AK, Iles R, Hobbs FD. Ten-year prognosis of heart failure in the community: follow-up data from the Echocardiographic Heart of England Screening (ECHOES) study. *Eur J Heart Fail* 2012;**14**:176–84. <http://dx.doi.org/10.1093/eurjhf/hfr170>
6. McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, *et al.* ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2012;**33**:1787–847. <http://dx.doi.org/10.1093/eurheartj/ehs104>
7. National Institute for Health and Care Excellence. *Chronic Heart Failure – Management of Chronic Heart Failure in Adults in Primary and Secondary Care*. London: NICE; 2010.
8. Rutten FH, Cramer MJ, Grobbee DE, Sachs AP, Kirkels JH, Lammers JW, Hoes AW. Unrecognized heart failure in elderly patients with stable chronic obstructive pulmonary disease. *Eur Heart J* 2005;**26**:1887–94. <http://dx.doi.org/10.1093/eurheartj/ehi291>
9. Bertens LC, Reitsma JB, van Mourik Y, Lammers JW, Moons KG, Hoes AW, Rutten FH. COPD detected with screening: impact on patient management and prognosis. *Eur Respir J* 2014;**44**:1571–8. <http://dx.doi.org/10.1183/09031936.00074614>
10. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;**380**:37–43. [http://dx.doi.org/10.1016/S0140-6736\(12\)60240-2](http://dx.doi.org/10.1016/S0140-6736(12)60240-2)
11. Le Jemtel TH, Padeletti M, Jelic S. Diagnostic and therapeutic challenges in patients with coexistent chronic obstructive pulmonary disease and chronic heart failure. *J Am Coll Cardiol* 2007;**49**:171–80. <http://dx.doi.org/10.1016/j.jacc.2006.08.046>
12. Fuat A, Hungin AP, Murphy JJ. Barriers to accurate diagnosis and effective management of heart failure in primary care: qualitative study. *BMJ* 2003;**326**:196. <http://dx.doi.org/10.1136/bmj.326.7382.196>
13. van Ierland Y, Elshout G, Berger MY, Vergouwe Y, de Wilde M, van der Lei J, *et al.* Translation of clinical prediction rules for febrile children to primary care practice: an observational cohort study. *Br J Gen Pract* 2015;**65**:e224–33. <http://dx.doi.org/10.3399/bjgp15X684373>

14. Stiell I, Wells G, Laupacis A, Brison R, Verbeek R, Vandemheen K, Naylor CD. Multicentre trial to introduce the Ottawa ankle rules for use of radiography in acute ankle injuries. Multicentre Ankle Rule Study Group. *BMJ* 1995;**311**:594–7. <http://dx.doi.org/10.1136/bmj.311.7005.594>
15. Lip GY. Using the CHADS2 and CHA2DS2-VASc scores for stroke risk prediction as well as the identification of stroke outcomes and cardiac complications in patients with and without atrial fibrillation. *Cerebrovasc Dis* 2013;**36**:281–2. <http://dx.doi.org/10.1159/000355981>
16. Roalfe AK, Mant J, Doust JA, Barton P, Cowie MR, Glasziou P, *et al*. Development and initial validation of a simple clinical decision tool to predict the presence of heart failure in primary care: the MICE (Male, Infarction, Crepitations, Edema) rule. *Eur J Heart Fail* 2012;**14**:1000–8. <http://dx.doi.org/10.1093/eurjhf/hfs089>
17. Zaphiriou A, Robb S, Murray-Thomas T, Mendez G, Fox K, McDonagh T, *et al*. The diagnostic accuracy of plasma BNP and NTproBNP in patients referred from primary care with suspected heart failure: results of the UK natriuretic peptide study. *Eur J Heart Fail* 2005;**7**:537–41. <http://dx.doi.org/10.1016/j.ejheart.2005.01.022>
18. Hobbs FD, Davis RC, Roalfe AK, Hare R, Davies MK, Kenkre JE. Reliability of N-terminal pro-brain natriuretic peptide assay in diagnosis of heart failure: cohort study in representative and high risk community populations. *BMJ* 2002;**324**:1498. <http://dx.doi.org/10.1136/bmj.324.7352.1498>
19. Cowie MR, Struthers AD, Wood DA, Coats AJ, Thompson SG, Poole-Wilson PA, Sutton GC. Value of natriuretic peptides in assessment of patients with possible new heart failure in primary care. *Lancet* 1997;**350**:1349–53. [http://dx.doi.org/10.1016/S0140-6736\(97\)06031-5](http://dx.doi.org/10.1016/S0140-6736(97)06031-5)
20. Cost B. *Heart Failure in the Elderly*. PhD thesis. Rotterdam: Erasmus University Rotterdam; 2000.
21. Wright SP, Doughty RN, Pearl A, Gamble GD, Whalley GA, Walsh HJ, *et al*. Plasma amino-terminal pro-brain natriuretic peptide and accuracy of heart-failure diagnosis in primary care: a randomized, controlled trial. *J Am Coll Cardiol* 2003;**42**:1793–800. <http://dx.doi.org/10.1016/j.jacc.2003.05.011>
22. Hobbs FD, Davis RC, Roalfe AK, Hare R, Davies MK. Reliability of N-terminal proBNP assay in diagnosis of left ventricular systolic dysfunction within representative and high risk populations. *Heart* 2004;**90**:866–70. <http://dx.doi.org/10.1136/hrt.2003.014258>
23. Tait L, Roalfe AK, Mant J, Cowie MR, Deeks JJ, Iles R, *et al*. The REFER (REFer for Echocardiogram) protocol: a prospective validation of a clinical decision rule, NT-proBNP, or their combination, in the diagnosis of heart failure in primary care. Rationale and design. *BMC Cardiovasc Disord* 2012;**12**:97. <http://dx.doi.org/10.1186/1471-2261-12-97>
24. Müller-Nordhorn J, Roll S, Willich SN. Comparison of the short form (SF)-12 health status instrument with the SF-36 in patients with coronary heart disease. *Heart* 2004;**90**:523–7. <http://dx.doi.org/10.1136/hrt.2003.013995>
25. EuroQol Group. EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:199–208 [http://dx.doi.org/10.1016/0168-8510\(90\)90421-9](http://dx.doi.org/10.1016/0168-8510(90)90421-9)
26. Noble M, McLennan D, Wilkinson K, Whitworth A, Dibben C, Barnes H. *English Indices of Deprivation 2007*. London: Department for Communities and Local Government; 2008.
27. NHS England. *Transforming Primary Care in London*. London: NHS England Primary Care Transformation Programme; 2013.
28. Taylor CJ, Hobbs FD. Heart failure therapy in patients with coronary artery disease. *Curr Opin Pharmacol* 2013;**13**:205–9. <http://dx.doi.org/10.1016/j.coph.2013.01.009>
29. Ho JE, Gona P, Pencina MJ, Tu JV, Austin PC, Vasan RS, *et al*. Discriminating clinical features of heart failure with preserved vs. reduced ejection fraction in the community. *Eur Heart J* 2012;**33**:1734–41. <http://dx.doi.org/10.1093/eurheartj/ehs070>

30. Bertens LC, van Mourik Y, Rutten FH, Cramer MJ, Lammers JW, Hoes AW, *et al.* Staged decision making was an attractive alternative to a plenary approach in panel diagnosis as reference standard. *J Clin Epidemiol* 2015;**68**:418–25. <http://dx.doi.org/10.1016/j.jclinepi.2014.09.020>
31. Adams ST, Leveson SH. Clinical prediction rules. *BMJ* 2012;**344**:d8312. <http://dx.doi.org/10.1136/bmj.d8312>
32. Lucassen W, Geersing GJ, Erkens PM, Reitsma JB, Moons KG, Büller H, van Weert HC. Clinical decision rules for excluding pulmonary embolism: a meta-analysis. *Ann Intern Med* 2011;**155**:448–60. <http://dx.doi.org/10.7326/0003-4819-155-7-201110040-00007>
33. Le Marechal F, Martinot A, Duhamel A, Pruvost I, Dubos F. Streptococcal pharyngitis in children: a meta-analysis of clinical decision rules and their clinical variables. *BMJ Open* 2013;**3**:e001482. <http://dx.doi.org/10.1136/bmjopen-2012-001482>
34. Balla J, Heneghan C, Thompson M, Balla M. Clinical decision making in a high-risk primary care environment: a qualitative study in the UK. *BMJ Open* 2012;**2**:e000414. <http://dx.doi.org/10.1136/bmjopen-2011-000414>
35. Owan TE, Hodge DO, Herges RM, Jacobsen SJ, Roger VL, Redfield MM. Trends in prevalence and outcome of heart failure with preserved ejection fraction. *N Engl J Med* 2006;**355**:251–9. <http://dx.doi.org/10.1056/NEJMoa052256>
36. Taylor CJ, Roalfe AK, Iles R, Hobbs FD. The potential role of NT-proBNP in screening for and predicting prognosis in heart failure: a survival analysis. *BMJ Open* 2014;**4**:e004675. <http://dx.doi.org/10.1136/bmjopen-2013-004675>
37. Storey P. Imaging for cardiac disease: a practical guide for general practitioners. *Aust Fam Physician* 2014;**43**:260–3.
38. Brenner S, Güder G. The patient with dyspnea. Rational diagnostic evaluation. *Herz* 2014;**39**:8–14. <http://dx.doi.org/10.1007/s00059-014-4057-6>
39. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.
40. Mant J, Doust J, Roalfe A, Barton P, Cowie MR, Glasziou P, *et al.* Systematic review and individual patient data meta-analysis of diagnosis of heart failure, with modelling of implications of different diagnostic strategies in primary care. *Health Technol Assess* 2009;**13**(32). <http://dx.doi.org/10.3310/hta13320>
41. Calvert MJ, Shankar A, McManus RJ, Ryan R, Freemantle N. Evaluation of the management of heart failure in primary care. *Fam Pract* 2009;**26**:145–53. <http://dx.doi.org/10.1093/fampra/cmn105>
42. Kotecha D, Holmes J, Krum H, Altman DG, Manzano L, Cleland JG, *et al.* Efficacy of β blockers in patients with heart failure plus atrial fibrillation: an individual-patient data meta-analysis. *Lancet* 2014;**384**:2235–43. [http://dx.doi.org/10.1016/S0140-6736\(14\)61373-8](http://dx.doi.org/10.1016/S0140-6736(14)61373-8)
43. Flather MD, Yusuf S, Køber L, Pfeffer M, Hall A, Murray G, *et al.* Long-term ACE-inhibitor therapy in patients with heart failure or left-ventricular dysfunction: a systematic overview of data from individual patients. ACE-Inhibitor Myocardial Infarction Collaborative Group. *Lancet* 2000;**355**:1575–81. [http://dx.doi.org/10.1016/S0140-6736\(00\)02212-1](http://dx.doi.org/10.1016/S0140-6736(00)02212-1)
44. Heran BS, Musini VM, Bassett K, Taylor RS, Wright JM. Angiotensin receptor blockers for heart failure. *Cochrane Database Syst Rev* 2012;**4**:CD003040. <http://dx.doi.org/10.1002/14651858.cd003040.pub2>
45. Paulus WJ, van Ballegoij JJ. Treatment of heart failure with normal ejection fraction: an inconvenient truth! *J Am Coll Cardiol* 2010;**55**:526–37. <http://dx.doi.org/10.1016/j.jacc.2009.06.067>

46. Cowie MR, Fox KF, Wood DA, Metcalfe C, Thompson SG, Coats AJ, *et al.* Hospitalization of patients with heart failure: a population-based study. *Eur Heart J* 2002;**23**:877–85. <http://dx.doi.org/10.1053/euhj.2001.2973>
47. Department of Health. *Payment by Results in the NHS: Tariff for 2013 to 2014*. 2013. URL: www.gov.uk/government/publications/payment-by-results-pbr-operational-guidance-and-tariffs (accessed 25 April 2016).
48. Department of Health. *NHS Reference Costs 2013/2014*. 2014. URL: www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014 (accessed 25 April 2016).
49. Joint Formulary Committee. *British National Formulary*. 68th edn. London: BMJ Group and Pharmaceutical Press; 2014.
50. Curtis L. *Unit Costs of Health and Social Care 2014*. Canterbury: Personal Social Service Research Unit, University of Kent; 2014.
51. Swedberg K, Cleland J, Dargie H, Drexler H, Follath F, Komajda M, *et al.* Guidelines for the diagnosis and treatment of chronic heart failure: executive summary (update 2005). *Eur Heart J* 2005;**26**:1115–40.
52. Cooper NA, Barton PM, Breijer M, Caffrey O, Opmeer BC, Timmermans A, *et al.* Cost-effectiveness of diagnostic strategies for the management of abnormal uterine bleeding (heavy menstrual bleeding and post-menopausal bleeding): a decision analysis. *Health Technol Assess* 2014;**18**(24). <http://dx.doi.org/10.3310/hta18240>
53. Borlaug BA. The pathophysiology of heart failure with preserved ejection fraction. *Nat Rev Cardiol* 2014;**11**:507–15. <http://dx.doi.org/10.1038/nrcardio.2014.83>
54. Scott MA, Price CP, Cowie MR, Buxton MJ, Scott M. Cost–consequences analysis of natriuretic peptide assays to refute symptomatic heart failure in primary care. *Br J Cardiol* 2008;**15**:199–206.
55. Ferrandis MJ, Ryden I, Lindahl TL, Larsson A. Ruling out cardiac failure: cost–benefit analysis of a sequential testing strategy with NT-proBNP before echocardiography. *Ups J Med Sci* 2013;**118**:75–9. <http://dx.doi.org/10.3109/03009734.2012.751471>
56. Goode KM, Clark AL, Cleland JG. Ruling out heart failure in primary-care: the cost–benefit of pre-screening using NT-proBNP and QRS width. *Int J Cardiol* 2008;**130**:426–37. <http://dx.doi.org/10.1016/j.ijcard.2007.08.131>

Appendix 1 The REFER study clinical record form

REFER STUDY: Nurse CRF

PATIENT INITIALS

ID NUMBER

--	--

--	--	--	--	--	--

General instructions

Each patient is provided with a unique 5 -digit ID number. Please enter this code at the top right hand corner of every page.

CRF completion

Units are provided for all measurements. Data should be entered in the units provided. If necessary please convert values to the unit specified.

Dates must be entered in the format DD MM YYYY using a leading zero if necessary.

Any errors should be corrected with a single strike through without obliterating the original entry. Any corrections should be initialled and dated.

Please check the calibration of all instruments prior to start

Stadiometer: Check recorded height of standard 1 metre rule once the instrument is set up and record the result.

Scales: The zero setting on the scales should be checked by pressing the reset button with the scales empty. This should read 0.00.

Taking measurements

Height: Participants should be asked to stand on the stadiometer. Please check:

- *Feet:* Ankles should be together and resting on the bar at the back. Ensure the participant doesn't stand on tiptoe.
- *Arms:* Should be resting at sides.
- *Head:* Should face straight ahead.

Weight: Participants should stand as reasonable straight if possible.

Blood Pressure: Subject should sit at the measurement table and rest their arm on the table. The cuff should be placed around the upper arm with the bladder centre over the artery. Ensure that the upper arm is at chest level.

Blood sampling

- Prepare label for vacutainer
- Wash hands between patients and wear rubber gloves.
- Ensure patient is sitting comfortably and check for previous problems with blood sampling.
- Clean the area with an alcohol wipe if obviously unclean. Allow skin to dry.
- After venepuncture raise subjects arm and ask patient to apply cotton wool to avoid bruising
- Dispose of needle in sharps box.

Any second attempt must be in the opposite arm.

REFER STUDY: Nurse CRF: PATIENT DETAILS 1

PATIENT INITIALS ID NUMBER

Consent form

Date consented __ / __ /

Please tick to confirm:

- | | Yes | No |
|-------------------------------------|--------------------------|--------------------------|
| • Consent form completed and signed | <input type="checkbox"/> | <input type="checkbox"/> |
| • One copy given to patient | <input type="checkbox"/> | <input type="checkbox"/> |
| • One copy filed in medical notes | <input type="checkbox"/> | <input type="checkbox"/> |
| • One copy kept for research file | <input type="checkbox"/> | <input type="checkbox"/> |

Date of birth

DD MM YY

Gender

Male Female

Post Code

.....

REFER STUDY: Nurse CRF: PATIENT DETAILS 1

PATIENT INITIALS

ID NUMBER

Marital status

Single

Divorced

Married

Separated

Cohabiting

Widowed

Your occupation (if retired – occupation before retirement)

.....

If married or cohabiting, partners occupation

.....

Your ethnic origin

White British

White Irish

White - Other Background

Mixed – White & Black Caribbean

Mixed – White & Black African

Mixed – White & Asian

Asian / Asian British – Indian

Asian / Asian British - Bangladeshi

Asian / Asian British – Other background

Black / Black British - Caribbean

Black / Black British - African

Black / Black British – Other background

Other – Chinese

Other – Any other ethnic group

Please specify:.....

REFER STUDY: Nurse CRF: CLINICAL HISTORY 1

PATIENT INITIALS

ID NUMBER

Clinical History

Previous MI	<input type="checkbox"/>	Angina	<input type="checkbox"/>
Arrhythmias (specify type below)	<input type="checkbox"/>	Heart Failure	<input type="checkbox"/>
Cardiac Devices (PPM / ICD)	<input type="checkbox"/>	Revascularisation (CABG/PTCA)	<input type="checkbox"/>
High cholesterol	<input type="checkbox"/>	Hypertension	<input type="checkbox"/>
H/O of rheum/scarlet fever	<input type="checkbox"/>	Valve Surgery	<input type="checkbox"/>
Diabetes	<input type="checkbox"/>	Peripheral artery disease	<input type="checkbox"/>
Stroke	<input type="checkbox"/>	TIA	<input type="checkbox"/>
COPD	<input type="checkbox"/>	Stomach ulcer	<input type="checkbox"/>
Arthritis	<input type="checkbox"/>	Depression	<input type="checkbox"/>
Other	<input type="checkbox"/>	Kidney Problems	<input type="checkbox"/>

Shortness of Breath

Other *please specify*:

.....

Do you have a family history of :**Yes****No****Don't****know**

Heart attack (<65 years of age)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
High blood pressure (<65 years of age)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diabetes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Do you get short of breath when you walk?If yes, how far can you manage when walking on level

ground at your own pace:

..... yards

For how long can you keep walking: mins

If no, is there any other reason that you have to stop walking?

.....

REFER STUDY: Nurse CRF: CLINICAL HISTORY 2

PATIENT INITIALS

ID NUMBER

	Yes	No
Do you get shortness of breath?	<input type="checkbox"/>	<input type="checkbox"/>
a) walking up hill	<input type="checkbox"/>	<input type="checkbox"/>
b) climbing upstairs	<input type="checkbox"/>	<input type="checkbox"/>
c) walking on the level with someone of you own age	<input type="checkbox"/>	<input type="checkbox"/>
d) during heavy housework (lifting/moving heavy furniture)	<input type="checkbox"/>	<input type="checkbox"/>
e) during moderate housework (vacuuming, sweeping)	<input type="checkbox"/>	<input type="checkbox"/>
f) during leisure activities	<input type="checkbox"/>	<input type="checkbox"/>
If <u>yes</u> , please specify activity:		
g) during light housework (dusting, washing dishes)	<input type="checkbox"/>	<input type="checkbox"/>
h) while washing or dressing	<input type="checkbox"/>	<input type="checkbox"/>
i) sitting at rest	<input type="checkbox"/>	<input type="checkbox"/>
j) lying down in bed	<input type="checkbox"/>	<input type="checkbox"/>
k) Do you wake during the night with shortness of breath	<input type="checkbox"/>	<input type="checkbox"/>

If yes, how many pillows do you sleep on?**How long have you had these symptoms** months years**Are your symptoms?** Improving stable getting worse **Do you get tired easily** yes no **Do your ankles/feet swell** yes no **Do you get chest pain** yes no

REFER STUDY: Nurse CRF: CLINICAL HISTORY 3

--	--

--	--	--	--	--	--

PATIENT INITIALS

ID NUMBER

Do you have a cough in the daytime yes no

Do you have a cough at night yes no

Do you get wheezy yes no

Do you cough up phlegm yes no

Which one is your main symptom that brought you to this consultation?

Shortness of breath

Tiredness

Ankle swelling

Chest pain

Cough

None

What diagnosis were you given?

None

Heart trouble

Heart failure

Angina

Bronchitis/asthma

Other

If other please specify:

REFER STUDY: Nurse CRF: TREATMENT 1

PATIENT INITIALS ID NUMBER

Are you under the care of a Consultant Cardiologist?Yes No

Name:.....

Hospital:.....

Date (approximately) referred: __ / __ / ____

Number of previous hospitalisations with Cardiac Events

Details with dates:

1)..... Date:

2)..... Date:

3)..... Date:

4)..... Date:

5)..... Date:

Has your GP recently (e.g since the consultation when this study was discussed) referred you for a chest x-ray?Yes No
 Before this consultation with your GP: -**Have you ever had?**

Yes No

An electrical recording of your heart (ECG)

A chest X-ray because of ill health

 If yes, was itnormal abnormal don't know

Yes No

An ultrasound recording of your heart (echo)

Do you have a Cardiac device?

No cardiac device

Standard pacemaker

ICD (defibrillator)

REFER STUDY: Nurse CRF: TREATMENT 2

--	--

--	--	--	--	--	--

PATIENT INITIALS

ID NUMBER

Medication

	Name	Dose / Day
Diuretic		
Digoxin		
ACE Inhibitor		
Angiotensin Receptor Blocker		
Beta Blocker		
Oral Nitrate		
Calcium Antagonist		
Other Hypertensive		
Anti - arrhythmic Drug		
Lipid Lowering Drug		
Anti Platelet		
Anticoagulant		
Others		
Over The Counter		

REFER STUDY: Nurse CRF: LIFESTYLE 1

PATIENT INITIALS

ID NUMBER

Have you ever smokedyes no If you have smoked - Do you still smokeyes no **If NO**, how long ago did you give up smoking months years

For how many years did you smoke years

How many did you smoke each day:
Tobacco

cigarettes

cigars

Grams

If YES, for how many years have you smoked yearsHow much do you smoke each day
Tobacco

cigarettes

cigars

Grams

Do you drink alcoholyes no **If yes**, how much do you drink in an average week (number)

½ pints of beer

glasses of wine

measure of spirits

REFER STUDY: Nurse CRF: PHYSICAL EXAM 1

PATIENT INITIALS

ID NUMBER

Physical Examination**Pulmonary function**
l/minFEV¹ l/minFVC **Height** . metres**Weight** kg**Waist Circumference** cm**Hip Circ** cm**Blood pressure** / mm Hg
systolic / *diastolic*After 5 mins or more mm Hg
systolic / *diastolic***Resting Pulse** Beats per minute**Rhythm** Regular Irregular **Peripheral oedema** Present Absent **Lung Auscultation**Clear Basal Creps Extensive Wheeze Reduced Air entry / dullness Other Please specify

REFER STUDY: Nurse CRF: QUESTIONNAIRE 1

PATIENT INITIALS ID NUMBER

Questionnaires

SF 12 Completed

Yes No

Reason if no:

EQ-5D questionnaire completed

Yes No

Reason if no:

Blood sampling

Consent for blood sampling obtained?

Yes No

Reason if no:

Blood sample obtained?

Yes No

Reason if no:

Appendix 2 Patient and public involvement

Heart failure is a prevalent condition associated with poor quality of life and reduced survival rates and, therefore, it represents a major public health problem. Effective treatments are available that improve both quantity and quality of life but patients must first receive a diagnosis. The symptoms of heart failure overlap with those of other conditions and can be gradual in onset, making the diagnostic process more challenging than for some other chronic diseases.

The importance of heart failure, the concept for the study and its design were discussed with a small group of people who later contributed as our trial steering committee; this included lay input (a non-patient representative with ethics committee experience), methodological expertise (statistics and trials) and clinical content expertise (cardiologist). As well as sitting on the trial steering committee, the lay representative also piloted our patient materials after their initial piloting with other experienced research staff not working on the REFER study.

The REFER study aimed to investigate the use of a CDR, NT-proBNP testing or their combination in helping clinicians make the diagnosis of heart failure. Eligible participants were those aged ≥ 55 years presenting to their GP with new-onset breathlessness, tiredness or ankle swelling of < 48 hours' duration, with no obvious recurrent, acute or self-limiting cause. The study involved prospective recruitment of patients during the GP consultation. To successfully recruit participants, the study relied on GPs first considering patient eligibility within each consultation and second inviting eligible patients to take part in the study. If a patient was willing to participate, the GP completed a simple online form, which included inclusion (and exclusion) criteria and asked for brief clinical details.

Recruitment rates were initially quite low and so we encouraged practices to display the study poster in their waiting rooms so that patients could be made aware of the study and could ask their GP about taking part. We also discussed the possibility of using social media such as Twitter (Twitter, Inc., San Francisco, CA, USA; www.twitter.com) and Facebook (Facebook, Inc., Menlo Park, CA, USA; www.facebook.com) to raise public awareness of the study but, given the fairly complex eligibility criteria and the requirement for recruitment by a GP within a patient symptomatic consultation, the study team agreed that this would not be helpful. Mouse mats and drinks coasters with study information were also produced to encourage GPs and patients to consider the study during consultations.

The results of the study have been written up for publication in open-access peer-review journals. The results broadly show that a NT-proBNP blood test is most helpful in determining the likelihood of a diagnosis of heart failure and therefore the clinical decision to refer for echocardiography (which provides objective evidence of cardiac function required for a formal diagnosis). Public awareness of the symptoms and signs of heart failure, and the appropriate use of blood testing by GPs, are important to facilitate earlier diagnosis and we will be publicising the results of the study once complete via our website (which has between 8000 and 18,000 hits per week, mostly from the public) and Twitter feeds. We will also prepare a press release aimed at raising public awareness of heart failure, the difficulties of accurate early diagnosis and the implications of the research for patients and their carers.

A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library