Diagnostic triage and the role of natriuretic peptide testing and echocardiography for suspected heart failure:

an appropriateness ratings evaluation by UK GPs

Abstract

Background

Some UK GPs are acquiring access to natriuretic peptide (NP) testing or echocardiography as diagnostic tests for heart failure. This study developed appropriateness ratings for the diagnostic application of these tests in routine general practice.

Aim

To develop appropriateness ratings for the diagnostic application of NP testing or echocardiography for heart failure in general practice.

Design and setting

An appropriateness ratings evaluation in UK general practice.

Method

Four presenting symptoms (cough, bilateral ankle swelling, dyspnoea, fatigue), three levels of risk of cardiovascular disease (low, intermediate, high), and dichotomous categorisations of cardiovascular/chest examination and electrocardiogram result, were used to create 540 appropriateness scenarios for patients in whom NP testing or echocardiography might be considered. These were rated by a 10-person expert panel, consisting of GPs and GPs with specialist interests in cardiology, in a two-round RAND Appropriateness Method.

Results

Onward referral for NP testing or echocardiography was rated as an appropriate next step in 217 (40.2%) of the 540 scenarios; in 194 (35.9%) it was rated inappropriate. The ratings also show where NP testing or echocardiography were ranked as equivalent next steps and when one test was seen as the more appropriate than the other.

Conclusion

NP testing should be the routine test for suspected heart failure where referral for diagnostic testing is considered appropriate. An abnormal electrocardiogram status makes referral to echocardiography an accompanying, or more appropriate, next step alongside NP testing, especially in the presence of dyspnoea. Abnormal NP testing should subsequently be followed up with referral for echocardiography.

Keywords

appropriateness; diagnosis; heart failure; natriuretic peptides; primary care; quality of healthcare.

INTRODUCTION

Diagnosing heart failure is difficult; it requires some objective evidence of cardiac dysfunction in the presence of symptoms and signs. Unfortunately, key indicators, such as dyspnoea, cough, or bilateral ankle swelling (BAS), are common and non-specific, with no single symptom or cluster of symptoms having sufficient predictive value for making a diagnosis.¹⁻³ Moreover, interpreting such symptoms can be difficult in, for example, older patients or those who are obese.^{1,4}

Reliance on symptoms and signs alone, which has been a standard way of labelling patients as having heart failure, can lead to both under- and over-diagnosis.⁵ As an example, many patients with peripheral oedema are, in the absence of testing, diagnosed as having heart failure when the oedema has another, unrelated aetiology.⁵ Another difficulty for GPs is distinguishing respiratory causes from cardiac causes of dyspnoea, in which symptoms of heart failure may be attributed to chronic obstructive pulmonary disease and, therefore, treated incorrectly.

In many parts of Europe, GPs are gaining enhanced direct access to investigations that improve the accuracy of diagnosing heart failure,⁶ namely natriuretic peptide (NP) testing and echocardiography; Box 1 outlines how these tests work. However, guidelines are not clear as to how these technologies should be applied in triaging patients with

symptoms; as a result, their use is inconsistent across, and within, countries, including the UK. The 'appropriateness' of the use of NP testing and echocardiography is also largely opinion based, with hospital specialists being viewed as the ultimate arbiters of the appropriateness of testing decisions made by GPs.⁷

Some specialists have expressed concern about the overdiagnosis of heart failure by GPs who rely simply on routine clinical assessment^{8,9} when it might be more appropriate to arrange NP testing or echocardiography. However, others believe that such enhanced access might simply lead to more indiscriminate use of NP testing, with a consequent rise in false positives.¹¹ As one of the symptoms of heart failure is fatigue, it is also quite possible to imagine that the current (and growing) battery of tests being ordered by many GPs for patients who are 'tired all the time' might be extended to include such items as C-reactive protein and NP.¹²

However, the precise definition and derivation of the term 'appropriateness' in studies examining GP testing is rather elusive.¹³ Appropriateness is not just about overuse or misuse, but should also consider underuse. For example, in a recent study of access to coronary angiography, it was demonstrated that many patients deemed appropriate for angiography failed to receive it and, as a result, experienced worse outcomes.¹⁴

SM Campbell, PhD, senior research fellow, National Primary Care Research and Development Centre, University of Manchester, Manchester; **A Fuat**, PhD, FRCGP, PGDip (Cardiology), research fellow, Wolfson Research Institute, Durham University, Stockton-on-Tees; **N Summerton**, MA, DM, MPH, MRCGP, FFP, DRCOG, GP; **N Lancaster**, BSc, research assistant, The Medical Centre, Driffield; **FD Hobbs**, FRCGP, FRCP, FESC, FMedSci, professor, Department of Primary Care Clinical Sciences, University of Birmingham, Edgbaston.

Address for correspondence

Dr Nick Summerton, The Medical Centre, Cranwell Road, Driffield, East Yorkshire, YO25 6UH.

E-mail: nsummerton@doctors.org.uk

Submitted: 19 August 2010; **Editor's response:**

15 September 2010; **final acceptance:** 30 September 2010.

©British Journal of General Practice

This is the full-length article (published online 27 Jun 2011) of an abridged version published in print. Cite this article as: **Br J Gen Pract 2011; DOI: 10.3399/bjgp11X583218.**

How this fits in

In many parts of Europe, including the UK, GPs are gaining enhanced access to natriuretic peptide testing and echocardiography to seek to improve the accuracy of diagnosing heart failure. Unfortunately as current guidance is unclear about how these technologies should be applied in triaging patients with specific symptoms or symptom combinations their use is inconsistent. This study uses the RAND Appropriateness Method to identify clusters of symptoms, patient characteristics and electrocardiogram results that GPs perceive as appropriate or inappropriate for referral for either natriuretic peptide testing or echocardiography as part of the diagnostic approach for heart failure.

In the absence of definitive guidance, the purpose of this study was, therefore, to derive appropriateness ratings for the use of heart-failure diagnostic triage in routine primary care. The aim was to identify the clusters of symptoms, patient characteristics, and electrocardiogram (ECG) results that GPs perceive as appropriate and inappropriate for referral for either NP testing or echocardiography as a diagnostic approach for heart failure. In addition, the study sought to identify those combinations of symptoms, patient characteristics, and ECG results for which uncertainty exists regarding the diagnostic pathway.

METHODS

The RAND Corporation has developed a definition of appropriateness in relation to GP-initiated testing:

'For an average group of patients presenting to an average GP, clinically appropriate testing means that the expected health benefit(s) exceed the expected negative consequences by a sufficiently wide margin that the test is worth doing'.¹⁵

In conjunction with the University of California, RAND has developed a validated method to generate appropriateness ratings; this involves combining a summary of scientific evidence with the collective judgement of experts.¹⁶ For diagnostic testing, this presents an opportunity to produce a series of ratings of clinical scenarios regarding the appropriateness of performing a particular test in relation to patient-specific symptoms, patient

characteristics, medical history, and previous investigations.¹⁷

This study adhered to the RAND Appropriateness Method^{17,18} by, initially, reviewing and collating the existing evidence base on heart-failure diagnosis, which included two recently published systematic reviews and their consultant papers.¹² From this review, together with discussions with cardiologists and GPs, a list of clinical scenarios was generated that covered the major symptom-initiated presentations within general practice for patients with possible heart failure (but not already known to have heart failure), in whom NP testing or echocardiography might be considered (Table 1).

Four common presenting symptoms (cough, BAS, dyspnoea, and fatigue/tiredness), three levels of perceived risk of cardiovascular disease (CVD) (low, intermediate, or high), and dichotomous categorisations of CVD/chest examination and ECG result were identified. The symptom-initiated presentations were then used to develop a list of 540 clinical scenarios (Table 1).

Each scenario represented a separate testing indication consisting of variable combinations of the presenting symptoms, CVD risk, chest examination, and ECG result. Although sex was not explicitly included as a separate variable, it was incorporated in the assessment of CVD risk. The decision to use three levels of CVD risk was based on considerations about the nature of primary care diagnostics and study feasibility. More detail on the characteristics of these variables and the associated instructions given to panellists is provided in Table 1.

Each scenario was then independently linked to one of following three potential decisions that a GP might make:

- no further investigations (Nil) — to carry out neither echocardiography nor an NP test;
- echocardiography — to directly refer for echocardiography without carrying out an NP test; or
- B-type NP (BNP)/NT-proBNP triage (NP) — to first carry out an NP test in order to rule out those individuals at low likelihood of having heart failure. The remaining patients would subsequently go on to be referred for echocardiography.

A rating scale was used to rate these potential decisions in relation to the clinical presentation on a 9-point integer scale, as follows:

Box 1. Natriuretic peptide and echocardiography diagnostic tests for heart failure

- Natriuretic peptide (NP) tests show how well a heart is working by measuring the amount of natriuretic peptide hormone in the blood. The plasma concentration is increased in patients with asymptomatic and symptomatic left ventricular dysfunction, where the heart has to work harder than usual over a long period of time, such as from heart failure. There is no difference in the diagnostic accuracy of B-type NP (BNP) and NT-proBNP.¹
- Echocardiography: echocardiogram/cardiac echocardiogram: an ultrasound scan of the heart giving detailed and accurate pictures of the heart muscle, the heart chambers, and structures within the heart such as the valves.

- scores 1–3: inappropriate next step (that is, no benefit, possible harms);
- scores 4–6: uncertainty about next step (that is, when harms and benefits are judged as approximately equal, or when the best available evidence does not support a judgement either way);
- scores 7–9: appropriate next step (that is, benefits were judged to outweigh harms).

As an example, a rating of 1 would mean that the decision would be an extremely inappropriate next step, whereas a rating of 9 would mean that it was extremely appropriate.

Panel

Panel composition is a crucial factor in ensuring the legitimacy of consensus techniques.²¹ As the aim of this study was to create specific appropriateness ratings for routine clinical general practice, the 10-person panel was identified and selected in discussion with the Primary Care Cardiovascular Society and the Royal College of General Practitioners. It

consisted of practising GPs, five of whom were nationally recognised experts in this area. These five individuals had particular interests in cardiology in general, and heart failure more specifically. Two panel members also had a significant publication record of original research in relation to NP and heart failure. The panel was co-chaired by one researcher, who is expert in the RAND Appropriateness Method, and one who is expert in primary care diagnostics.

Round one was conducted by email. Panel members were sent the two systematic reviews, the literature summary, and the list of scenarios, together with a list of definitions for all the terms used. They were asked to independently consider and rate the appropriateness of each decision for every scenario on the 9-point integer scale, basing their ratings on the summarised evidence and their own clinical experience as GPs.

In round two, panellists met for a 1.5-day face-to-face meeting, in which they were given rating sheets that included feedback from the first round. This feedback included the frequency distribution of ratings of all

Table 1. Variable characteristics and instructions given to panellists

Variable	Characteristics and instructions
Presenting symptoms (onset within the preceding 3 months)	Cough — productive/non productive, day and/or night Ankle swelling — bilateral Dyspnoea — including exertional dyspnoea, dyspnoea at rest, orthopnoea, and/or paroxysmal nocturnal dyspnoea Fatigue/tiredness — related to exertion
CVD/chest examination normal/abnormal]	An abnormal examination was defined as that which identified one or more of the following features: ^a <ul style="list-style-type: none"> • Elevated jugular venous pressure • Arrhythmia — possible atrial fibrillation or tachycardia (rate 100/minute at rest) • Additional heart sounds — including S3, S4, and murmurs • Laterally displaced apical impulse • Lung crackles/crepitations
Risk of CVD ¹⁹	The panellists were explicitly asked to base their understanding of low, intermediate, or high risk of CVD on the following considerations: <ul style="list-style-type: none"> • Known JBS2 score (that is, taking into account age, sex, smoking status, blood pressure, cholesterol, and, perhaps, family history and ethnic origin). Based on the JBS2 score a 10-year CVD risk of <10% is categorised as 'low' whereas a 10-year CVD risk of >20% is categorised as 'high'. Consideration could also be given to alcohol intake, body mass/build, and activity levels in estimating risk • Current treatments for hypertension, hyperlipidaemia, HIV and anti-psychotic medication affect risk of CVD; in general, individuals undergoing these treatments will be, at least, at an intermediate level of risk • Comorbidities: individuals with chronic kidney disease, autoimmune disorders (such as systemic lupus erythematosus and rheumatoid arthritis), and women after a premature menopause are, at least, at an intermediate level of risk • High risk: The following were assumed to be at high risk: <ul style="list-style-type: none"> — Cardiovascular history: that is, angina, myocardial infarction, stroke, transient ischaemic attack, or peripheral arterial disease — Diabetes (type 1 or type 2) in individuals over the age of 40 years — Familial hypercholesterolaemia
ECG (normal/none or abnormal)	In considering this categorisation, panellists were asked to appreciate that a 'normal' ECG recording will, in most cases, exclude heart failure, but abnormalities might be subtle and unrecognised by generalists. ²⁰ For the sake of this exercise, the panel was, therefore, asked to only consider an ECG as normal if it had (a) been reported by an individual with specialist skills in ECG interpretation, or (b) reported as 'normal' by an ECG machine with interpretative software

^aAnkle oedema has not been listed here as it is covered in the presenting symptoms. CVD = cardiovascular disease. ECG = electrocardiogram. JBS = Joint British Societies.

panellists across the 9-point scale, together with the overall panel median rating, for each of the 540 scenarios. Each panellist was also presented with their own first-round rating sheet as a reminder of how they had initially rated each scenario. During the meeting, panellists then discussed each symptom-based grouping of scenarios in turn, focusing on areas of disagreement, before rating each scenario on their individual blinded rating sheets.

Data entry and analysis

Within a consensus technique, there are two aspects to each scale's rating process for every scenario: the overall panel median rating and the level of agreement within a panel.²¹ The RAND Appropriateness Method definition of consensus (overall panel median of 7–9) was adhered to without disagreement. Disagreement was defined to exist where $\geq 33\%$ of panellists rated a scenario in the 1–3 range and $\geq 33\%$ of panellists rated the same scenario in the 7–9 range of the 1–9 point scale. In consequence, a scenario rated 7–9 without disagreement was considered to be an appropriate next step, and a rating of 1–3 without disagreement was considered to be an inappropriate next step.

RESULTS

Results are presented for the ratings from the second round only and are summarised in Tables 2 and 3. The appropriate next steps are presented as matrices in Appendix 1.

Table 2 shows a summary of the panel ratings for the 540 separate scenarios and indicates a high level of agreement within the panel, with only 42 scenarios (7.7%) rated with disagreement. Of the 540 scenarios, 217 (40.2%) were rated as appropriate next steps without disagreement and 194 (35.9%) scenarios were rated as inappropriate next steps without disagreement. In all other scenarios ($n=129$) the appropriate next

step was considered equivocal.

Table 3 presents the ratings for all combinations of symptoms and tests, and shows, for each combination, which of the three next steps studied were rated as either appropriate or inappropriate; the overall panel median rating is given in parentheses. When the panel rated the cluster of symptoms and tests as an equivocal marker for referral for either NP testing or echocardiography, no next step is recorded.

Table 3 shows examples of scenarios in which the panel felt referral for either NP testing or echocardiography would be an inappropriate next step and that no investigation was an appropriate next step; for example, where there was a normal CVD examination, low CVD risk, and normal/no ECG status with either cough, fatigue, or BAS presenting as single symptoms. However, in this case where dyspnoea presented in combination with at least two other symptoms, NP testing was never rated as inappropriate.

There was only one example of NP testing being rated an appropriate next step and echocardiography as an inappropriate next step (normal CVD examination, low CVD risk, normal/no ECG status and presenting symptoms of cough, BAS, and dyspnoea). There were also examples of scenarios in which referral to NP testing was rated an appropriate next step but referral to echocardiography was considered equivocal (that is, normal CVD examination, low CVD risk, abnormal ECG status and presenting symptoms of dyspnoea and fatigue or cough or BAS [data not shown]). In general, the presence of dyspnoea was more likely to be associated with NP testing being an appropriate next step compared to no testing.

Table 3 also shows that there are no scenarios in which echocardiography was rated an appropriate next step but NP testing was not also considered

Table 2. Summary of round-two appropriateness ratings

	Appropriateness
Number of scenarios	540
Median 1–3	194 (35.9%)
Median 4–6	114 (21.1%)
Median 7–9	232 (43.0%)
Inappropriate (median 1–3 without disagreement)	194 (35.9%)
Appropriate (median 7–9 without disagreement)	217 (40.2%)
Inappropriate (median 1–3 with disagreement)	5 (0.9%)
Appropriate (median 7–9 with disagreement)	3 (0.6%)
Disagreement (33% of scores in 1–3 and 7–9)	42 (7.7%)

Table 3. Summary appropriateness ratings for all combinations of symptoms and tests

CVD risk/CVD exam/ ECG status	Cough	BAS	Dyspnoea	Fatigue	Cough/ BAS	Cough/ dyspnoea	Cough/ fatigue	BAS/ dyspnoea	BAS/ fatigue	Cough/ BAS/ dyspnoea/ fatigue	Cough/ dyspnoea/ fatigue	BAS/ dyspnoea/ fatigue	Cough/BAS/ dyspnoea/ fatigue
Normal CVD exam; low CVD risk; normal/no ECG status	Nil = A (9) BNP = I (1) E = I (1)	Nil = A (8) BNP = I (2) E = I (1)	E = I (2)	Nil = A (9) BNP = I (2) E = I (1)	Nil = A (8) BNP = I (2) E = I (2)	Nil = A (7) BNP = I (3) E = I (1)	Nil = A (7.5) BNP = I (2) E = I (1)	E = I (2) BNP = A (7.5) E = I (1)	Nil = A (7) BNP = A (7.5) E = I (1)	Nil = I (2) BNP = A (7) E = I (3)	E = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (1) BNP = A (8) E = A (7)
Normal CVD exam; low CVD risk; abnormal ECG status	E = I (3)	E = I (3)	NA = I (2) BNP = A (7.5) E = A (9)	E = I (2) BNP = A (8) E = I (3)	E = I (3) BNP = A (8) E = I (3)	NA = I (3) BNP = A (8) E = I (3)	E = I (3) BNP = A (8) E = I (3)	NA = I (2) BNP = A (7.5) E = I (3)	Nil = I (3) BNP = A (7.5) E = I (1)	Nil = I (2) BNP = A (7.5) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (1) BNP = A (7.5) E = A (7)	Nil = I (1) BNP = A (8) E = A (7.5)
Normal CVD exam; intermediate CVD risk; normal/no ECG status	Nil = A (8.5) BNP = I (1) E = I (1)	Nil = A (8) BNP = I (2) E = I (1)	E = I (3) BNP = I (2) E = I (1)	Nil = A (8) BNP = I (2) E = I (1)	Nil = A (8) BNP = I (3) E = I (2)	Nil = A (7) BNP = I (3) E = I (2)	Nil = A (7) BNP = I (3) E = I (2)	Nil = I (3) BNP = I (2) E = I (1)	Nil = A (7) BNP = A (7.5) E = I (1)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (1) BNP = A (8) E = A (7.5)
Normal CVD exam; intermediate CVD risk; abnormal ECG status	-	BNP = A (7) E = I (2)	Nil = I (2) BNP = A (8) E = I (3)	E = I (3) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	-	Nil = I (2) BNP = A (7.5) E = I (1)	Nil = I (2) BNP = A (7.5) E = I (1)	Nil = I (2) BNP = A (8) E = I (3)	Nil = I (2) BNP = A (7.5) E = I (3)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7.5)
Normal CVD exam; high CVD risk; normal/no ECG status	Nil = A (7) BNP = I (2) E = I (2)	E = I (3) BNP = A (7) E = I (2)	Nil = I (2) BNP = A (7) E = I (3)	E = I (2) BNP = A (7) E = I (3)	E = I (3) BNP = A (7) E = I (3)	E = I (3) BNP = A (7) E = I (3)	E = I (3) BNP = A (7) E = I (3)	Nil = I (3) BNP = A (7.5) E = I (1)	E = I (3) BNP = A (7.5) E = I (1)	Nil = I (2) BNP = A (7.5) E = I (3)	Nil = I (2) BNP = A (7.5) E = I (3)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7.5)
Normal CVD exam; high CVD risk; abnormal ECG status	Nil = I (3) BNP = A (7.5) E = A (7)	Nil = I (3) BNP = A (7) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (3) BNP = A (7) E = A (7)	Nil = I (3) BNP = A (7) E = A (7)	Nil = I (2) BNP = A (7) E = A (7)	Nil = I (2) BNP = A (7) E = A (7)	Nil = I (1) BNP = A (7.5) E = A (7)	Nil = I (2) BNP = A (7.5) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7.5)
Abnormal CVD exam; low CVD risk; normal/no ECG status	-	Nil = I (2) BNP = A (7.5) E = A (8)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7) E = I (3)	Nil = I (2) BNP = A (7.5) E = A (7)	Nil = I (2) BNP = A (7.5) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7.5)
Abnormal CVD exam; low CVD risk; abnormal ECG status	Nil = I (1) BNP = A (7.5) E = A (7.5)	Nil = I (1) BNP = A (8) E = A (7.5)	Nil = I (1) BNP = A (8) E = A (7.5)	Nil = I (1) BNP = A (7) E = A (7.5)	Nil = I (2) BNP = A (7) E = A (7)	Nil = I (2) BNP = A (7) E = A (7)	Nil = I (2) BNP = A (7) E = A (7)	Nil = I (1) BNP = A (7.5) E = A (7)	Nil = I (1) BNP = A (7.5) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7)	Nil = I (1) BNP = A (8) E = A (7.5)
Abnormal CVD exam; intermediate CVD risk; normal/no ECG status	-	Nil = I (2) BNP = A (7.5) E = A (9)	Nil = I (2) BNP = A (7) E = A (9)	Nil = I (2) BNP = A (7) E = A (9)	Nil = I (2) BNP = A (7) E = A (9)	Nil = I (2) BNP = A (7) E = A (9)	Nil = I (2) BNP = A (7) E = A (9)	Nil = I (2) BNP = A (7.5) E = A (9)	Nil = I (2) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (8) E = A (9)	Nil = I (1) BNP = A (8) E = A (9)	Nil = I (1) BNP = A (8) E = A (9)	Nil = I (1) BNP = A (8) E = A (9)
Abnormal CVD exam; high CVD risk; abnormal ECG status	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (7.5) E = A (9)	Nil = I (1) BNP = A (8) E = A (9)	Nil = I (1) BNP = A (8) E = A (9)	Nil = I (1) BNP = A (8) E = A (9)	Nil = I (1) BNP = A (8) E = A (9)

*Overall second-round panel median ratings are given in parentheses. A = appropriate. BAS = B-type natriuretic peptide. CVD = cardiovascular disease. ECG = electrocardiogram. E = echocardiogram. Nil = no investigation. I = inappropriate.

Funding body

UK Department of Health (Pathology Programme).

Ethics committee

The GPRD Group has ethical approval from a multicentre research ethics committee for all purely observational research using GPRD data; namely, studies that do not include patient involvement, as here. No individual patients are identifiable through this research.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

FD Richard Hobbs has received occasional sponsorship or speaking fees from a variety of pharmaceutical and diagnostic companies with interests in heart failure, including Astra-Zeneca, Pfizer and Roche Diagnostics. The other authors have declared no competing interests.

Acknowledgments

We thank the panellists — Terry McCormack, Nigel Rowell, Ahmet Fuat, Peter Savill, Richard Falk, Nick Taylor, Chris Arden, Richard Hobbs, Raghu Raghunath, and Michael Norton — in addition to the following colleagues from the Department of Health: Sue Hill, Roger Boyle, and Ian Barnes.

Discuss this article

Contribute and read comments about this article on the Discussion Forum: <http://www.rcgp.org.uk/bjgp-discuss>

appropriate. For many scenarios, echocardiography and NP testing were rated as equally appropriate next steps; however, in some scenarios, one test was considered more appropriate than the other (Table 3). In general, the presence of an abnormal CVD examination result, high CVD risk, and abnormal ECG status made referral to echocardiography a more appropriate next step than NP testing. In most scenarios representing normal or abnormal CVD examination results in combination with any risk of CVD and normal/no ECG status, NP testing was rated as more appropriate than echocardiography.

These ratings have been used to generate appropriateness matrices for single symptoms and combinations of symptoms (Appendix 1). Furthermore, in developing the appropriateness matrices, it also made clinical sense to combine the low and intermediate CVD risk categories. These matrices show where NP and echocardiography were rated the most appropriate next step, and where NP and echocardiography were rated as equally appropriate next steps.

DISCUSSION

Summary

This study has identified situations in which GP referral for NP testing or echocardiography would be considered appropriate, inappropriate, or equivocal on the basis of peer consensus by generalist and specialist GPs. Given the lack of definitive guidance on such diagnostic triaging, these data could form a basis for how 'open-access' diagnostic testing services should be commissioned for primary care. Although every patient who is eventually diagnosed with heart failure will likely have received an NP assay and echocardiography as part of their work-up, in which patient phenotypes the tests might be considered, and in what order, is not determined.

Strengths and limitations

This study adhered to a validated systematic consensus method for developing appropriateness scenarios.^{17,22} Panel composition is the most important influence on panel ratings and, as advised, the study's panel members reflected the issue, evidence base, and likely users of the output of the ratings.²¹ The study panel, comprising generalists and generalists with a special interest in heart failure, considered the overall evidence base in these appropriateness ratings; however

they did not explicitly consider cost effectiveness, which is not part of the RAND Appropriateness Method.¹⁵

Comparison with existing literature

These findings have resonance with the conclusions of the recent systematic review of Mant *et al*, who proposed that patients with suspected heart failure should be offered BNP (or NT-proBNP) as a triage test prior to echocardiography, unless presenting with any one of a history of myocardial infarction, basal crepitations, or, in men, ankle oedema.¹

Implications for practice and research

This study sought to generate appropriateness ratings and appropriateness matrices that are clinically precise, specific, and applicable to routine clinical practice in relation to an area in which GPs express a lack of confidence.²³ The importance of these findings should also be considered in the context of the variable access that primary care currently has to NP, or even echocardiography, testing within the UK.

Although the RAND Approach has been applied successfully in a number of specialist areas of diagnostic testing,²⁴ the authors are not aware of any attempt to apply it to assessing the appropriateness of diagnostic testing requests by generalists. This study describes the development of appropriateness ratings for GPs in relation to NP and echocardiography diagnostic testing for possible heart failure. In cases where the study panel rated all 'next steps' as equivocal, it does not follow that NP testing or echocardiography should be ruled out, but that the GP should take other issues, such as comorbidities, patient preferences, and other clinical information, into account when making treatment decisions.

Appropriateness ratings that are developed using consensus techniques have high face validity. However, this is a minimum prerequisite for any quality measure and subsequent developmental work is required to provide empirical evidence and to test for acceptability, feasibility, reliability, sensitivity to change, and validity.^{21,25,26} Although the findings illuminate the decision-making process for diagnosing heart failure in routine general practice, they should now be the subject of prospective validation and empirical testing using an observational design.

These results indicate that NP testing should be used as part of the routine decision-making process for diagnosing

suspected heart failure within primary care in all cases in which referral for any diagnostic testing is considered appropriate. Subsequently, all patients with abnormal NP testing should be followed up by referral for echocardiography. However, in general,

the presence of an abnormal ECG status makes referral to echocardiography a more appropriate next step or an accompanying next step alongside NP testing, especially in the presence of dyspnoea.

REFERENCES

- Mant J, Doust J, Roalfe A, *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): 1–207, iii.
- Madhok V, Falk G, Rogers A, *et al*. The accuracy of symptoms, signs and diagnostic tests in the diagnosis of left ventricular dysfunction in primary care: a diagnostic accuracy systematic review. *BMC Fam Pract* 2008; **9**: 56.
- Davie AP, McMurray JJV. Symptoms and signs of heart failure. In: Dargie HJ, McMurray JJV, Poole-Wilson PA (eds). *Managing heart failure in primary care*. London: Blackwell Healthcare Communications, 1996: 58–68.
- Cleland JGF, Erdmann E, Ferrari R, *et al*. Guidelines for the diagnosis of heart failure. *Eur Heart J* 1995; **16**(6): 741–751.
- Blankfield RP, Finkelhor RS, Alexander JJ, *et al*. Etiology and diagnosis of bilateral leg edema in primary care. *Am J Med* 1998; **105**(3): 192–197.
- National Institute for Health and Clinical Excellence. *Chronic heart failure: management of chronic heart failure in adults in primary and secondary care*. London: NICE, 2003.
- Summerton N. *Patient-centred diagnosis*. Abingdon: Radcliffe Medical Press, 2007.
- Remes J, Miettinen H, Reunanen A, Pyörälä K. Validity of clinical diagnosis of heart failure in primary health care. *Eur Heart J* 1991; **12**(3): 315–321.
- Wheeldon NM, MacDonald TM, Flucker CJ, *et al*. Echocardiography in chronic heart failure in the community. *Q J Med* 1993; **86**(1): 17–23.
- Mogelvang R, Goetze JP, Schnohr P, *et al*. Discriminating between cardiac and pulmonary dysfunction in the general population with dyspnoea by plasma pro-B-type natriuretic peptide. *J Am Coll Cardiol* 2007; **50**(17): 1694–1701.
- Armstrong-Klein S. *Brain-type natriuretic peptide (BNP): an information resource for cardiac networks*. NHS Improvement, 2008.
- Gialamas A, Beilby JJ, Pratt NL, *et al*. Investigating tiredness in Australian general practice. Do pathology tests help in diagnosis? *Aust Fam Physician* 2003; **32**(8): 663–666.
- Smellie WS, Lowrie R, Wilkinson E. A laboratory based intervention to improve appropriateness of lipid tests and audit cholesterol lowering in primary care. *BMJ* 2001; **323**(7323): 1224–1227.
- Hemingway H, Chen R, Junghans C, *et al*. Appropriateness criteria for coronary angiography in angina: reliability and validity. *Ann Intern Med* 2008; **149**(4): 221–231.
- Brook RH. Appropriateness: the next frontier. *BMJ* 1994; **308**(6923): 218–219.
- Hicks NR. Some observations on attempts to measure appropriateness of care. *BMJ* 1994; **309**(6956): 730–733.
- Shekelle P. Appropriateness criteria: a useful tool for the cardiologist. *Heart* 2009; **95**(7): 517–520.
- Brook RH, Chassin MR, Fink A, *et al*. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 1986; **2**(1): 53–63.
- British Cardiac Society; British Hypertension Society; Diabetes UK, *et al*. JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. *Heart* 2005; **91**(Suppl 5): v1–52.
- Mant J, Fitzmaurice DA, Hobbs FD, *et al*. Accuracy of diagnosing atrial fibrillation on electrocardiogram by primary care practitioners and interpretative diagnostic software: analysis of data from screening for atrial fibrillation in the elderly (SAFE) trial. *BMJ* 2007; **335**(7616): 380 doi:10.1136/bmj.39227.551713.AE.
- Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. *BMJ* 2003; **326**(7393): 816–819.
- Naylor CD. What is appropriate care? *N Eng J Med* 1998; **338**(26): 1918–1920.
- Fuat A, Hungin AP, Murphy JJ. Barriers to accurate diagnosis and effective management of heart failure in primary care: qualitative study. *BMJ* 2003; **326**(7382): 196 doi:10.1136/bmj.326.7382.196.
- Brindis RG, Douglas PS, Hendel RC, *et al*. ACCF/ASNC appropriateness criteria for single-photon emission computed tomography myocardial perfusion imaging (SPECT MPI): a report of the American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness Criteria Working Group and the American Society of Nuclear Cardiology endorsed by the American Heart Association. *J Am Coll Cardiol* 2005; **46**(8): 1587–1605.
- HEDIS. *Technical specifications for physician measurement*. Washington, DC: NCQA, 2009.
- Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use*. Oxford: Oxford Medical Publications, 1995.

Symptom combinations:

Cough and ankle swelling:

Risk	Exam	ECG	Next steps
			Nil
			NP
			NP
			NP
			NP/echo
			NP/echo*
			Echo/NP

Cough and dyspnoea:

Risk	Exam	ECG	Next steps
			Nil
			NP
			NP
			NP
			NP/echo*
			NP/echo*
			Echo/NP

Cough and fatigue:

Risk	Exam	ECG	Next steps
			Nil
			NP
			NP/echo*
			Echo/NP

Fatigue and dyspnoea:

Risk	Exam	ECG	Next steps
			NP
			NP
			NP
			NP/echo
			NP/echo
			NP/echo
			Echo/NP

Ankle swelling and dyspnoea:

Risk	Exam	ECG	Next steps
			NP
			NP
			NP
			NP
			NP/echo
			NP/echo*
			NP/echo
			Echo/NP

Ankle swelling and fatigue:

Risk	Exam	ECG	Next steps
			Nil
			NP
			NP
			NP
			NP/echo
			Echo/NP

Cough, ankle swelling and dyspnoea:

Risk	Exam	ECG	Next steps
			NP
			NP
			NP
			NP
			NP/echo
			NP/echo
			NP/echo
			Echo/NP

Cough, ankle swelling and fatigue:

Risk	Exam	ECG	Next steps
			NP
			NP
			NP
			NP
			NP/echo
			NP/echo
			NP/echo
			Echo/NP

Cough, fatigue and dyspnoea:

Risk	Exam	ECG	Next steps
			NP
			NP
			NP
			NP/echo
			NP/echo
			NP/echo
			Echo/NP

Fatigue, ankle swelling and dyspnoea:

Risk	Exam	ECG	Next steps
			NP
			NP
			NP
			NP/echo
			NP/echo
			NP/echo*
			NP/echo
			Echo/NP

Cough, fatigue, ankle swelling and dyspnoea:

Risk	Exam	ECG	Next steps
			NP
			NP
			NP
			NP/echo
			NP/echo
			NP/echo*
			Echo/NP
			Echo/NP
			Echo/NP

Single symptoms:

Cough:

Risk	Exam	ECG	Next steps
			Nil
			Nil
			NP
			NP
			NP
			NP/echo
			Echo/NP

Dyspnoea:

Risk	Exam	ECG	Next steps
			NP
			NP
			NP
			NP
			NP/echo
			NP/echo*
			Echo/NP

Ankle swelling:

Risk	Exam	ECG	Next steps
			Nil
			NP
			NP
			NP
			NP
			NP
			NP
			NP/echo*

Fatigue:

Risk	Exam	ECG	Next steps
			Nil
			NP
			NP
			NP
			NP/echo*
			Echo/NP

- = Abnormal finding or high risk of cardiovascular disease
- = Normal finding or low/intermediate risk of cardiovascular disease
- Echocardiogram [Echo]: to directly refer for an echocardiogram without undertaking a BNP test
- B-type NP(BNP)/NT-proBNP triage (NP): to first undertake a BNP/NT-proBNP test in order to 'rule out' those individuals at low likelihood of heart failure

In the matrices, the following sequence applies:
 NP/Echo = NP is the most appropriate next step
 Echo/NP = Echo is the most appropriate next step
 NP/Echo* = NP or echo are seen as equally appropriate next steps

ECG = electrocardiogram. Exam = cardiovascular disease/chest examination.
 Risk = risk of cardiovascular disease.