

Research article

Open Access

Heart failure diagnosis in primary health care: clinical characteristics of problematic patients. A clinical judgement analysis study

Ylva Skånér*¹, Johan Bring², Bengt Ullman³ and Lars-Erik Strender¹

Address: ¹Family Medicine Stockholm, Karolinska Institutet, Alfred Nobels allé 12, SE-141 83 Huddinge, Sweden, ²Statisticon AB, Östra Ågatan 31, SE-753 22 Uppsala, Sweden and ³Department of Cardiology, Stockholm Söder Hospital, SE-118 83 Stockholm, Sweden

Email: Ylva Skånér* - ylva.skanner@klinvet.ki.se; Johan Bring - johan.bring@statisticon.se; Bengt Ullman - bengt.ullman@sos.sll.se; Lars-Erik Strender - lars-erik.strender@klinvet.ki.se

* Corresponding author

Published: 18 September 2003

Received: 11 April 2003

BMC Family Practice 2003, 4:12

Accepted: 18 September 2003

This article is available from: <http://www.biomedcentral.com/1471-2296/4/12>

© 2003 Skånér et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Early detection of chronic heart failure has become increasingly important since the introduction of effective treatment. However, clinical diagnosis of heart failure is known to be difficult, especially in mild cases or early in the course of the disease. The purpose of this study is to analyse how patient characteristics contribute to difficulties in diagnosing systolic heart failure.

Methods: Design: A Clinical Judgement Analysis study of 40 case vignettes based on authentic patients, including relevant clinical data except echocardiography. Setting: Primary health care and two cardiology outpatient clinics in Stockholm. Subjects: 70 participants with different types of clinical experience; 27 specialists in general practice, 22 cardiologists, and 21 medical students. Main outcome measures: The assessed probability of heart failure for each case vignette, and the disagreement between the participants. The number of clinical variables (cues) indicative of heart failure in the case vignettes.

Results: The ten case vignettes with the *least* diverging assessments more often had increased relative cardiac volume and atrial fibrillation. No further specific clinical patterns could be found in subgroups of the case vignettes. The ten case vignettes with the *most* diverging assessments were those with an intermediate number of clinical variables. The case vignettes with the least diverging assessments more often represented patients with cardiac enlargement and atrial fibrillation.

Conclusion: Diagnosing mild heart failure is difficult, as these patients are not easy to characterise. In our study, a larger number of positive cues resulted in more diagnostic conformity among the participants, and the most important information was cardiac enlargement. The importance of more objective diagnostic methods in diagnosing suspected cases of heart failure should be emphasised.

Background

Heart failure is a major cause of morbidity and mortality, and the prevalence seems to be increasing. In Stockholm the prevalence has been estimated to 1.3–2.5 % [1] and in

a recent study, 0.7 % of a population of 100 000 inhabitants registered at 13 health centres in Stockholm had the diagnosis heart failure [2].

Early detection of heart failure has become increasingly important, as modern drug treatment has the potential to improve symptoms and quality of life, slow down the rate of disease progression, and improve survival. Unfortunately, the clinical diagnosis of chronic heart failure is known to be difficult, especially in mild cases, as many features of the condition are not organ specific, and there may be few clinical features in the early stages of the disease [3–5]. An echocardiogram gives an objective measurement of heart function, and is now often considered to be decisive for the diagnosis of heart failure. However, it often represents a late step in the diagnostic process in primary health care, and the clinical evaluation of symptoms and signs is still crucial. Knowledge about various aspects of diagnostic thinking is an important research area. Better insight into how doctors diagnose heart failure can help us to improve education and guidelines, and to construct decision support systems.

In a Clinical Judgement Analysis (CJA) study we analysed differences between general practitioners, cardiologists and medical students [6]. We found that the groups were strikingly similar with respect to diagnostic accomplishment, and they used similar diagnostic strategies for assessing the probability of systolic heart failure. However, on an individual level they differed considerably in both respects.

In this study the aim was to analyse how patient characteristics contribute to difficulties in diagnosing systolic heart failure. Our hypothesis was that patients who were either typical heart failure patients (a large number of variables indicating heart failure) or typical non-heart failure patients (few or no variables indicating heart failure) would be easier to diagnose correctly than the intermediate cases.

Methods

The procedure regarding participants and case vignettes, has been described in detail elsewhere [6].

A total of 70 physicians with different types of clinical experience participated in the study: 27 general practitioners (GPs), 22 cardiologists, and 21 medical students. In the above-mentioned study we had found the three groups to be similar regarding strategies and diagnostic accomplishment when assessing heart failure probability, and we therefore decided to treat them as a single group in this context.

The case vignettes were based on authentic patients, as we wanted them to be realistic, i.e. we wanted the associations between values of the variables to represent associations found in real patients [7]. In order to assess the participants' diagnostic accomplishment, we needed

patients with a valid diagnosis (heart failure or not heart failure). We used 40 patients referred from GPs to the cardiology outpatient clinic at the Stockholm Söder Hospital for problems related to heart failure. The data were collected from patient records at the cardiology clinic. The diagnoses were made by the attending cardiologists, and were based on all available clinical information including echocardiography. Diastolic dysfunction is difficult to measure in old patients and in the presence of atrial fibrillation, and therefore diastolic dysfunction will not be discussed in this study.

The case vignettes were presented in a booklet. Eight variables (cues) were used to characterise each patient: a history of myocardial infarction, dyspnoea, atrial fibrillation, leg oedema, rales, systolic blood pressure, signs of minor pulmonary stasis, and relative cardiac volume on chest X-ray. The cues were chosen because of their relevance, according to articles, textbooks, interviews with GPs and prediction validity in relevant populations, and because of their availability in the patients' medical records. Information about the left-ventricular ejection fraction was not given in the case vignettes, as we wanted to capture an earlier step in the judgement process. For each vignette the participants assessed the probability of the patient having some degree of heart failure. The assessments were made on a visual analogue scale, anchored at the ends by 0 % (with certainty, not heart failure), and 100 % (with certainty, heart failure), respectively.

We dichotomised the non-binary cue values in order to count the number of positive cues for each case vignette. A positive cue for systolic blood pressure was ≤ 140 mm Hg (since a low value had been found to be associated with heart failure in our group of patients), and for relative cardiac volume a positive cue was ≥ 490 ml/m² for men and ≥ 450 ml/m² for women (corresponding to the local reference values).

To get a measurement of how difficult it was to diagnose each case vignette, we used the interquartile range (IQR, the difference between the third and the first quartile) of the participants' assessments of the probability of heart failure. A large interquartile range indicates that the participants have divergent opinions about whether the patient has heart failure or not. The ten vignettes with the highest interquartile ranges were considered the most difficult cases and the ten vignettes with the lowest values the least difficult cases.

For each vignette the median value of the participants' assessments was calculated. A high median value indicates that the participants consider it probable that the patient has heart failure.

Table 1: Some characteristics of the case vignettes.

	Heart failure n = 26	Non-heart failure n = 14
Age, mean and range	77 (56–92)	76 (56–84)
Sex, number of men	11	7
Number of patients with a history of myocardial infarction	10	2
Number of patients with dyspnoea	23	9
Number of patients with atrial fibrillation *)	14	3
Number of patients with leg oedema	9	4
Number of patients with rales	8	4
Number of patients with a systolic blood pressure \leq 140 mmHg *)	17	4
Number of patients with signs of pulmonary stasis	11	5
Number of patients with cardiac enlargement **)	25	8
Number of patients with cardiac volume >700 ml/m ² ***)	15	2
Relative cardiac volume (ml/m ²), mean and range ***)	722 (470–940)	529 (350–880)
Echocardiogram: ejection fraction (%), mean and range (not presented as a cue) ***)	37 (20–55)	55 (50–65)

*) $p < 0.05$ **) $p < 0.01$ ***) $p < 0.001$

T-test, chi-square test and Fisher exact test were used for analysing differences between groups. The association between the interquartile ranges and the number of positive cues for all the vignettes was studied with a regression line plot. All calculations were done using Minitab version 11.12, except Fisher exact test, which was done in EpiInfo version 6.

Approval for the study was obtained from the local Ethics Committee.

Results

Characteristics of the case vignettes

Table 1 presents the characteristics of the case vignettes. Twenty-six of the cases had heart failure. Vignettes representing patients with and without heart failure differed regarding relative cardiac volume, atrial fibrillation, and systolic blood pressure. They also differed regarding ejection fraction values (not presented in the vignettes). The numbers of positive cues for the two groups were 4.5 and 2.7, respectively. The average age was 76 years.

Five of the patients without heart failure had hypertension, two had valvular diseases, two had angina pectoris, two had diabetes, and one had chronic obstructive pulmonary disease. Patients with and without heart failure did not differ in this respect. One patient with and four patients without heart failure were suspected of having diastolic cardiac dysfunction.

Case vignettes with the least and the most divergent assessments

Table 2 presents the characteristics of the case vignettes with the least and the most divergent assessments. The

two groups differed only regarding relative cardiac volume and atrial fibrillation.

For all of the vignettes the variation between the probability assessments was large. The differences between maximum and minimum assessments (the range) for the individual vignettes varied between 48 and 100 percentage units. The assessments of the case vignettes with the least and the most divergent assessments are presented in Figure 1. The interquartile ranges (box sizes) for the vignettes with the least divergent assessments were between 10 and 19 percentage units, and for those with the most divergent assessments they were between 30 and 42 percentage units. The fact that the participants had converging opinions about a vignette did not mean, however, that their assessment was correct (for example, patients 20 and 26 in Figure 1, who were assessed as probable heart failure patients but who did not have that diagnosis).

The ten vignettes with the least divergent assessments were rated as having a rather high risk of heart failure (median heart failure probability assessments between 70 and 91 percentage units), and eight of them represented patients with heart failure. They had, on the average, five positive cues. The ten vignettes with the most divergent assessments were rated as having a medium risk of heart failure (median heart failure probability assessments between 37 and 81 percentage units), and five of them represented patients with heart failure. They had, on the average, 3.4 positive cues.

Table 3 shows some cue combinations. The "classic" heart failure findings (dyspnoea, oedema and rales) were found together in very few vignettes. Vignettes representing

Table 2: Characteristics of the case vignettes representing the least and the most divergent assessments

	Least divergent assessments n = 10	Most divergent assessments n = 10
Number of patients with heart failure	8	5
Age, mean and range	75 (61–92)	75 (56–84)
Sex, number of men	6	6
Number of patients with a history of myocardial infarction	3	5
Number of patients with dyspnoea	10	7
Number of patients with atrial fibrillation ^{***)}	8	1
Number of patients with leg oedema	3	4
Number of patients with rales	5	3
Number of patients with a systolic blood pressure ≤ 140 mmHg	6	4
Number of patients with signs of pulmonary stasis	5	4
Number of patients with cardiac enlargement	10	6
Number of patients with cardiac volume >700 ml/m ² ^{***)}	8	1
Relative cardiac volume (ml/m ²), mean and range ^{***)}	791(600–920)	533(370–900)
Echocardiogram: ejection fraction (%), mean and range (not presented as a cue)	40(25–58)	44.6(25–55)

^{***)} p < 0.01; ^{****)} p < 0.001

patients with and without heart failure differed regarding the combination of cardiac enlargement and a history of myocardial infarction, whereas vignettes representing the least and the most divergent assessments did not. For the rest of the cue combinations, no difference between the groups was found.

The regression line plot indicates that the case vignettes with the least divergent assessments were those with the highest and the lowest number of positive cues, and the case vignettes with the most divergent assessments were those with an intermediate number of positive cues (Fig. 2).

Discussion

In this study the aim was to describe patient characteristics that can make the identification of patients with systolic heart failure more or less difficult. Our hypothesis that the case vignettes with a large number of positive cues would be easier to diagnose correctly than those with an intermediate number was supported by the regression line plot (Figure 2). The small number of patients with few positive cues makes it more difficult, however, to draw conclusions about those patients (the left part of Figure 2).

The combination of dyspnoea, oedema and rales constitutes the "classic" clinical picture of heart failure. It was found in few of the patients and was therefore of little use for the judgements. All the vignettes represented patients treated with loop diuretics, which may have obscured some of the findings. Treatment with diuretics is, however, common in this group of patients, and information about the treatment was given in the vignettes to make it possible for the participants to evaluate its effect. Dysp-

noea alone was common in all the patients, and could therefore not discriminate between the groups.

A study on heart failure diagnostics in primary health care has shown that cardiac enlargement (measured by displaced apex beat) was the best single predictor of left ventricular systolic dysfunction, and that a combination of cardiac enlargement and a history of myocardial infarction had the best positive predictive value [8]. Apex displacement is a coarse method for measuring cardiac enlargement, and it is seldom used by Swedish doctors, who prefer to use X-ray for this purpose. In our study, the relative cardiac volume on X-ray was the single best predictor among the cues. The combination of cardiac enlargement and a history of myocardial infarction could discriminate between heart failure and non-heart failure patients, but it did not discriminate between the patients with the least and the most divergent assessments. However, the relative cardiac volume of the patients with the least divergent assessments were larger (920, 810 and 760 ml/m²) than those with the most divergent assessments (650 and 610 ml/m²), which could explain why it was easier for the doctors to reach consensus about them.

The ejection fraction is often considered a gold standard in heart failure diagnostics, and the fact that it was not included as a variable in the case vignettes might have made the judgement situation difficult for some of the participants. On the other hand, since the ejection fraction could also be a dominating variable, including it could have made it difficult to evaluate the influence of the other variables. B type natriuretic peptide has been shown to be useful for ruling out systolic heart failure in

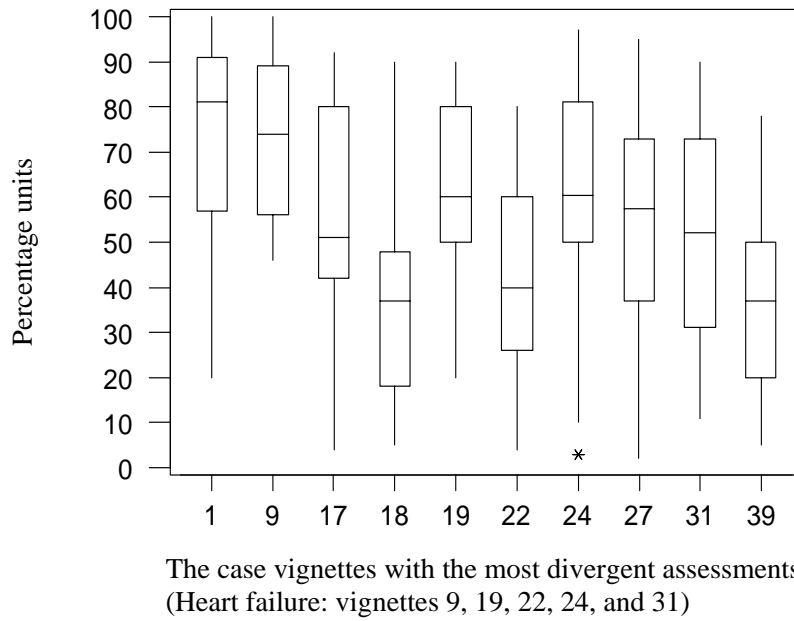
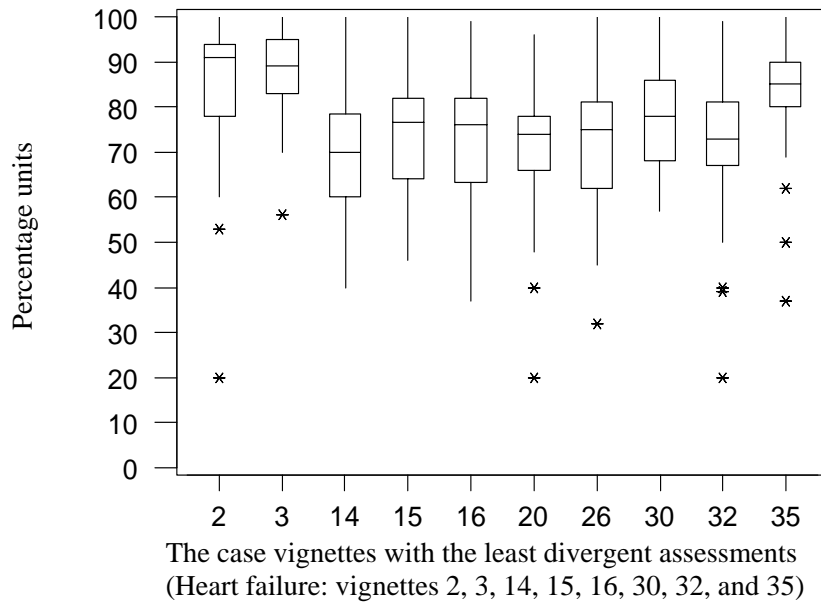


Figure 1

The case vignettes with the least and the most divergent assessments. The box size (= the interquartile range) reflects the participants' divergence in rating the probability of heart failure for each individual patient. The bottom of the box is at the first quartile (Q1), the top is at the third quartile (Q3), and the line across the box is at the median value. The "whiskers" (= the lines that extend from the top and bottom of the box) extend to the smallest and the largest observation (= participant) that is not considered an outlier. Outliers (*) are observations outside these limits.

Table 3: Some cue combinations.

	Least divergent assessments n = 10	Most divergent assessments n = 10	Heart failure n = 26	Non-heart failure n = 14
Cardiac enlargement and a history of myocardial infarction *)	3	2	9	0
Cardiac enlargement and dyspnoea	10	4	22	5
Cardiac enlargement and dyspnoea and atrial fibrillation	8	1	12	2
Dyspnoea and leg oedema and rales	1	1	2	1

*) $p < 0.05$ (for heart failure compared with non-heart failure patients)

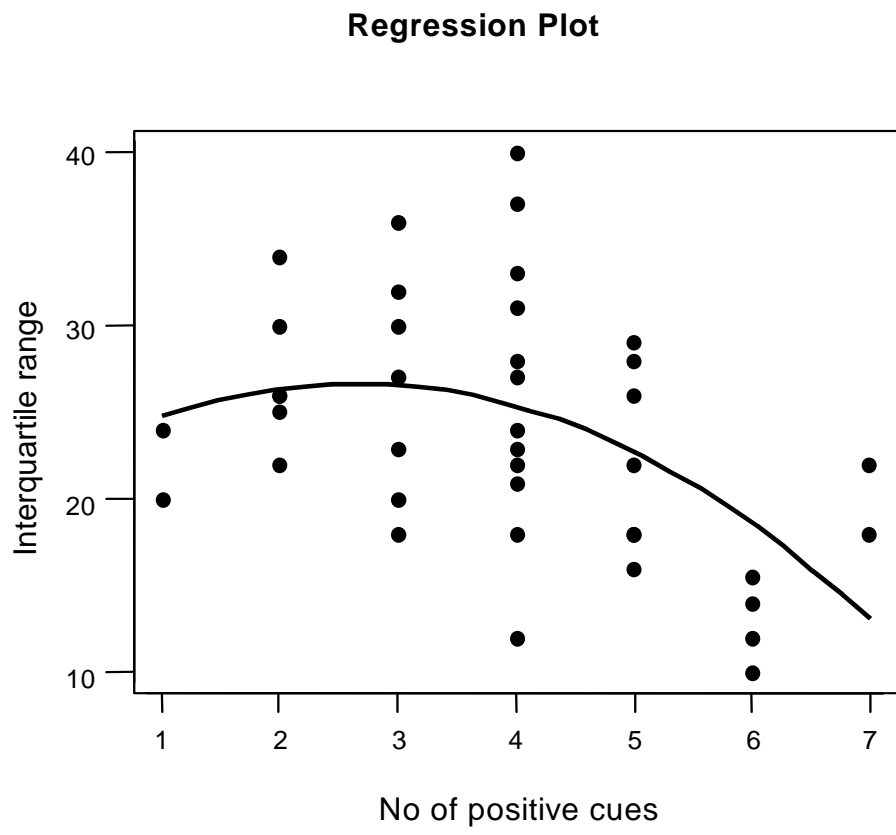


Figure 2
Association between assessment divergency and number of positive cues. A regression line plot representing the association between the interquartile range values (degree of assessment divergency) and the number of positive cues for the 40 case vignettes.

general practice [9]. As it is not used in routine clinical practice in Sweden, it could not be included as a variable in the vignettes.

The European Society of Cardiology has adopted guidelines for the diagnosis and treatment of chronic heart failure. The guidelines list conditions that are necessary or supporting for establishing the presence of heart failure, or oppose the diagnosis [10]. Our vignettes were modelled to represent an intermediate step in the diagnostic process, and therefore information that could help to rule out heart failure is missing (echocardiography, natriuretic peptides and electrocardiogram). Given the information in the vignettes, however, the following conclusions could be drawn: All the case vignettes with the least divergent assessments fulfil the symptom criterion in the guidelines. In five of them, the diagnosis is supported by information about both appropriate signs (rales) and chest X-ray (cardiac enlargement and/or pulmonary stasis), and in five of them by information about chest X-ray only. Nine of the case vignettes with the most divergent assessments also fulfil the symptom criterion. In one of them, the diagnosis is supported by information about both signs and chest X-ray, in five of them by information about chest X-ray only, in two of them by information about signs only, and in one of the vignettes, there is no further support to the diagnosis. The vignettes with the most divergent assessments thus had less support from the guidelines.

Relatively few studies on patients suspected of having heart failure have been performed in primary health care settings. Those that have been conducted often report over-diagnosis [4,5,11-14], while fewer studies report under-diagnosis [13,15-17]. Under-diagnosis is more difficult to study than over-diagnosis, as data not only must be collected from patients diagnosed as having heart failure, but also from a larger population. We have used a new method to look at heart failure diagnostics in a group of patients referred from primary health. Clinical Judgement Analysis studies using case vignettes with validated diagnoses can be another way of studying under-diagnosis.

With Clinical Judgement Analysis we can describe important aspects of the clinical judgement process (e.g. the influence of variable values of on the decisions, and diagnostic achievement). There are, however, other aspects that could be better studied with process oriented methods, e.g. successive decision-making, evaluation of information, and use of decision rules [18-21]. Findings with high sensitivity and low specificity, such as history and physical examination, are more important in an early stage of the diagnostic process, and findings with high specificity and low sensitivity, such as echocardiography, in a later stage, for validating the diagnosis. It would

therefore be valuable to supplement the perspective in this study with a study using a process perspective.

Conclusions

Identifying patients with mild heart failure is difficult, as these patients are not easy to characterise. In our study, a larger number of positive cues resulted in more diagnostic conformity among the participants, and the most important information was cardiac enlargement. The importance of more objective diagnostic methods in diagnosing suspected cases of heart failure should be emphasised.

Competing interests

None declared.

Authors' contributions

YS, JB and LES conceived of the study. YS carried out the data collection, performed the statistical analyses and drafted the manuscript. All authors participated in the design of the study, the interpretation of the results and the discussions of the drafts. All authors read and approved the final manuscript.

Acknowledgements

We thank all participating doctors and students. The study was supported by grants from the Stockholm County Council and the Swedish Heart Lung Foundation.

References

1. Mejhert M, Persson H, Edner M and Kahan T: **Epidemiology of heart failure in Sweden – a national survey.** *Eur J Heart Failure* 2001, **3**:97-103.
2. Nilsson G and Strender L-E: **Management of heart failure in primary health care. A retrospective study on electronic patient records in a registered population.** *Scand J Prim Health Care* 2002, **20**:161-5.
3. Watson RDS, Gibbs CR and Lip GYH: **ABC of heart failure: Clinical features and complications.** *BMJ* 2000, **320**:236-9.
4. Remes J, Miettinen H, Reunanen A and Pyörälä K: **Validity of clinical diagnosis of heart failure in primary health care.** *Eur Heart J* 1991, **12**:315-21.
5. Owen A and Cox S: **Diagnosis of heart failure in elderly patients in primary care.** *Eur J Heart Failure* 2001, **3**:79-81.
6. Skånér Y, Bring J, Ullman B and Strender L-E: **The use of clinical information in diagnosing chronic heart failure: A comparison between general practitioners, cardiologists and students.** *J Clin Epidemiol* 2000, **53**:1081-8.
7. Cooksey RV: *Judgment Analysis. Theory, Methods, and Applications* San Diego: Academic Press; 1996.
8. Davie A, Francis C, Caruana L, Sutherland G and McMurray J: **Assessing diagnosis in heart failure: which features are any use?** *QJM* 1997, **90**:335-9.
9. Smith H, Pickering RM, Struthers A, Simpson I and Mant D: **Biochemical diagnosis of ventricular dysfunction in elderly patients in general practice: observational study.** *BMJ* 2000, **320**:906-8.
10. Task Force for the Diagnosis and Treatment of Chronic Heart Failure of the European Society of Cardiology: **Guidelines for the diagnosis and Treatment of Chronic Heart Failure.** *Eur Heart J* 2001, **22**:1527-60.
11. Clarke KW, Gray D and Hampton JR: **Evidence of inadequate investigation and treatment of patients with heart failure.** *Br Heart J* 1994, **71**:584-7.

12. Mair FS, Crowley TS and Bundred PE: **Prevalence, aetiology and management of heart failure in general practice.** *Br J Gen Pract* 1996, **46**:55-61.
13. Nielsen LS, Svanegaard J, Wiggers P and Egeblad H: **The yield of a diagnostic hospital dyspnoea clinic for the primary health care section.** *J Intern Med* 2001, **250**:422-8.
14. Wheeldon NM, MacDonald TM, Flucker CJ, McEnderick AD, McDevitt DG and Struthers AD: **Echocardiography in chronic heart failure in the community.** *Q J Med* 1993, **86**:17-23.
15. Blankfield RP, Finkelhor RS, Alexander JJ, Flocke SA, Maiocco J and Goodwin M et al.: **Etiology and diagnosis of bilateral leg edema in primary care.** *Am J Med* 1998, **105**:192-7.
16. Hobbs F, Jones M, Allan T, Wilson S and Tobias R: **European survey of primary care physician perceptions on heart failure diagnosis and management (Euro-HF).** *Eur Heart J* 2000, **21**:1877-87.
17. Morgan S, Smith H, Simpson I, Liddiard GS, Raphael H and Pickering RM et al.: **Prevalence and clinical characteristics of left ventricular dysfunction among elderly patients in general practice setting: cross sectional survey.** *BMJ* 1999, **318**:368-72.
18. Boshuizen HPA and Claessen HFA: **Problems of research into medical problem solving: some remarks on theory and method.** *Med Educ* 1982, **16**:81-7.
19. Patel VL and Groen GJ: **The general and specific nature of medical expertise: A critical look.** In: *Toward a general theory of expertise: Prospects and limits* Edited by: Ericsson, KA, Smith J. New York: Cambridge University Press; 1991:93-125.
20. Patel VL and Groen GJ: **Knowledge Based Solution Strategies in Medical Reasoning.** *Cognitive Science* 1986, **10**:91-116.
21. Schmidt HG, Norman GR and Boshuizen HPA: **A Cognitive Perspective on Medical Expertise: Theory and Implications.** *Acad Med* 1990, **65**:611-21.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2296/4/12/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

