Open Access: free full text available online
For the full versions of these articles see bmj.com

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

Jonathan Mant,[1] David A Fitzmaurice,[1] F D Richard Hobbs,[1] Sue Jowett,[2] Ellen T Murray,[1] Roger Holder,[1] Michael Davies,[3] Gregory Y H Lip[4]

EDITORIAL by Van Weert
RESEARCH p383

[1]Department of Primary Care and General Practice, University of Birmingham, Birmingham B15 2TT
[2]Health Economics Facility, Health Services Management Centre, University of Birmingham, Birmingham B15 2RT
[3]Department of Cardiology, Queen Elizabeth Hospital, Edgbaston, Birmingham B15 2TH
[4]University Department of Medicine, City Hospital, Birmingham B18 7QH
Correspondence to: D A Fitzmaurice
d.a.fitzmaurice@bham.ac.uk

BMJ 2007;335:380-2
doi:10.1136/bmj.39227.551713.AE

This article is an abridged version of a paper that was published on bmj.com on 29 June 2007.
Cite this version as: *BMJ* 29 June 2007, doi: 10.1136/bmj.39227.551713.AE (abridged text, in print: *BMJ* 2007;335:380-2)

ABSTRACT

Objective To assess the accuracy of general practitioners, practice nurses, and interpretative software in the use of different types of electrocardiogram to diagnose atrial fibrillation.

Design Prospective comparison with reference standard of assessment of electrocardiograms by two independent specialists.

Setting 49 general practices in central England.

Participants 2595 patients aged 65 or over screened for atrial fibrillation as part of the screening for atrial fibrillation in the elderly (SAFE) study; 49 general practitioners and 49 practice nurses.

Interventions All electrocardiograms were read with the Biolog interpretative software, and a random sample of 12 lead, limb lead, and single lead thoracic placement electrocardiograms were assessed by general practitioners and practice nurses independently of each other and of the Biolog assessment.

Main outcome measures Sensitivity, specificity, and positive and negative predictive values.

Results General practitioners detected 79 out of 99 cases of atrial fibrillation on a 12 lead electrocardiogram (sensitivity 80%, 95% confidence interval 71% to 87%) and misinterpreted 114 out of 1355 cases of sinus rhythm as atrial fibrillation (specificity 92%, 90% to 93%). Practice nurses detected a similar proportion of cases of atrial fibrillation (sensitivity 77%, 67% to 85%), but had a lower specificity (85%, 83% to 87%). The interpretative software was significantly more accurate, with a specificity of 99%, but missed 36 of 215 cases of atrial fibrillation (sensitivity 83%). Combining general practitioners' interpretation with the interpretative software led to a sensitivity of 92% and a specificity of 91%. Use of limb lead or single lead thoracic placement electrocardiograms resulted in some loss of specificity.

Conclusions Many primary care professionals cannot accurately detect atrial fibrillation on an electrocardiogram, and interpretative software is not sufficiently accurate to circumvent this problem, even when combined with interpretation by a general practitioner. Diagnosis of atrial fibrillation in the community needs to factor in the reading of electrocardiograms by appropriately trained people.

INTRODUCTION

Electrocardiograms for the diagnosis of atrial fibrillation are now generated and read in primary care, whether by a general practitioner, a practice nurse, or interpretative software. However, little research has been done into the extent to which the type of reader affects the accuracy with which atrial fibrillation is detected by electrocardiography. A systematic review of studies of interpretation of electrocardiograms found that physicians of all specialties made frequent errors when interpreting electrocardiograms.[1] An evaluation of a computer software algorithm found that it could detect atrial fibrillation with a sensitivity of 91% and a specificity of 99%.[2] A subsidiary question to the reliability of interpretation is whether a full 12 lead electrocardiogram is needed to diagnose atrial fibrillation. The aim of this study was to assess the accuracy with which general practitioners, practice nurses, and interpretative software diagnose atrial fibrillation on 12 lead electrocardiograms, single lead thoracic placement electrocardiograms, and limb lead recordings.

METHODS

We did the study as a prospective sub-study within the screening for atrial fibrillation in the elderly (SAFE) randomised controlled trial of different methods of screening for atrial fibrillation in primary care.[3] This involved 50 practices in central England, which were randomly allocated as 25 intervention practices, in which a screening programme was initiated, and 25 control practices. One general practitioner and one practice nurse from each practice was involved in the study. The training of practitioners in the intervention practices included a one hour session on how to interpret an electrocardiogram. Practitioners in the control practices received no training.

Generation of electrocardiograms—We obtained 2595 12 lead electrocardiograms from people aged 65 or over from the 25 SAFE "intervention" practices (figure).

We assembled at random 25 batches of 100 electrocardiograms, comprising a third each of 12 lead, single thoracic placement, and limb lead electrocardiograms. We distributed these to the 49 practices, where they were read by one general practitioner and one practice nurse.

Reading of electrocardiograms—We interpreted all the electrocardiograms with interpretative software. We asked each participant to indicate whether or not atrial fibrillation or atrial flutter was present in each case. Practitioners were blinded to patients' identities, the diagnoses made by the specialists, and the diagnoses generated by the interpretative software.

Reference standard—Two consultant cardiologists, blinded to the software interpretation and that of the primary care practitioners, read all the 12 lead electrocardiograms independently of each other.

Statistical methods—We calculated sensitivity, specificity, and positive and negative predictive values. We used regression to examine variation in sensitivity and specificity with the type of electrocardiogram and between control and intervention practices. For comparison of sensitivity and specificity between types of electrocardiogram and between general practitioners and nurses, we did both matched and unmatched analyses.

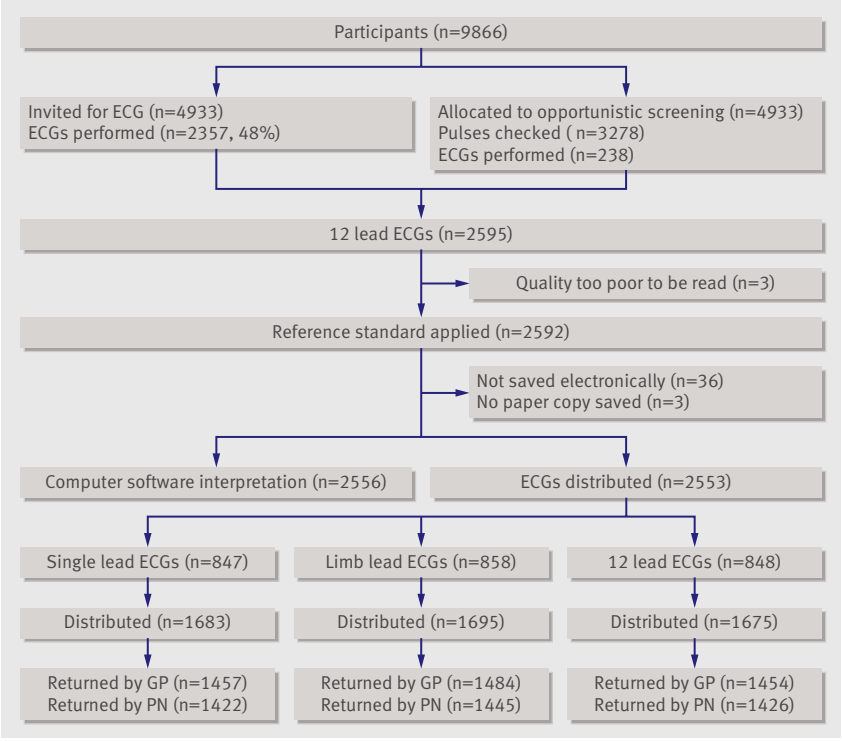
RESULTS

Forty two (86%) primary care physicians and 41 (84%) nurses returned the results of their interpretation. The prevalence of atrial fibrillation was 8.4%. Interpretative software was the most accurate method of reading electrocardiograms but did not give a rhythm diagnosis in 109 (4.3%) and missed 36 (17%) cases of atrial fibrillation. Ten per cent of computer diagnoses of atrial fibrillation were incorrect. The combined sensitivity of general practitioner and interpretative software was 92%, and the specificity was 91% (table).

General practitioners and practice nurses detected similar proportions of cases of atrial fibrillation (80% *v* 77% on 12 lead electrocardiogram), but diagnosis by general practitioners was more specific. Nevertheless, a diagnosis of atrial fibrillation by a general practitioner was still more likely to be wrong than right (positive predictive value 40.9%). Use of 12 lead, limb lead, or single thoracic placement electrocardiograms made little difference to the ability of primary care practitioners to correctly diagnose the presence of atrial fibrillation (P=0.52, matched (0.52, unmatched) for general practitioners; P=0.08 (0.35) for nurses). A significant difference occurred when diagnosing the absence of atrial fibrillation, however; 12 lead electrocardiograms gave a better outcome for general practitioners (P<0.001 (<0.001); P=0.12 (0.23) for nurses).

No significant difference existed between the performance of general practitioners from intervention and control practices. Control general practitioners showed a sensitivity of 84.0% and a specificity of 88.1%, and intervention general practitioners showed a sensitivity of 81.3% and a specificity of 88.9% (P=0.57 for sensitivity; P=0.19 for specificity). However, practice nurses from intervention practices interpreted electrocardiograms more accurately than did those from control practices: sensitivity 76.5% versus 68.9%; specificity 88.9% versus 78.9% (P=0.11 for sensitivity; P<0.001 for specificity).

The ability of individual general practitioners and practice nurses to diagnose atrial fibrillation accurately on an electrocardiogram varied widely. The sensitivity of individual doctors varied from 50% to 100% and that of practice nurses from 0% to 100%; the standard deviations of individual sensitivities were 31% for general practitioners and 37% for nurses. The false positive rate of general practitioners varied from 0% to 44% and that of practice nurses from 0% to 61%, and the respective standard deviations were 13% and 17%.



Flowchart of generation and reading of electrocardiograms. ECG=electrocardiogram; GP=general practitioner; PN=practice nurse. *Each ECG was sent to two practices, except for one batch that was sent to only one practice

DISCUSSION

In this study, general practitioners were unable to diagnose atrial fibrillation accurately on an electrocardiogram. Twenty per cent of cases of atrial fibrillation were missed, and the probability that a positive diagnosis was correct was only 41%. Changing from the 12 lead to simpler electrocardiograms resulted in further loss of specificity.

Our results are substantially different from those reported by Somerville and colleagues, who found that a general practitioner could detect atrial fibrillation accurately.[4] Our results suggest that such performance is atypical among general practitioners. Practice nurses were less accurate than the doctors, and interpretative software was more accurate than the doctors.

The generally lower P values on the matched analyses reflect the variation in performance between the raters. This variation was greater for practice nurses than it was for general practitioners, reflecting a greater range of ability.

Summary statistics of accuracy of interpretation of electrocardiograms by reader and type

Reader and electrocardiogram type	Sensitivity (95% CI)	Specificity (95% CI)
Interpretative software:		
12 lead	83.3 (78.3 to 88.2)	99.1 (98.7 to 99.5)
General practitioner:		
12 lead	79.8 (70.5 to 87.2)	91.6 (90.1 to 93.1)
Limb lead	82.5 (74.8 to 88.7)	88.5 (86.9 to 90.2)
Chest lead	84.8 (78.7 to 91.0)	86.4 (84.6 to 88.3)
Practice nurse:		
12 lead	77.1 (67.4 to 85.0)	85.1 (83.0 to 86.9)
Limb lead	72.0 (63.9 to 80.1)	83.4 (81.4 to 85.4)
Chest lead	68.7 (60.1 to 76.4)	82.8 (80.7 to 84.8)
General practitioner and interpretative software*:		
12 lead	91.9 (86.6 to 97.3)	91.1 (89.6 to 92.6)

For the purposes of calculations for this table, "uncertain" diagnoses are counted as missed diagnoses for sensitivity and as not atrial fibrillation for specificity.
*Defined as positive result if either (or both) is positive.

Strengths and weaknesses of study

The electrocardiograms being read for this study were generated as part of a screening programme and so reflect the sort of electrocardiograms that primary care practitioners would need to read if screening for atrial fibrillation. A strength of this study is the large number of practitioners and electrocardiograms involved. Previous studies have tended to use a single cardiologist as a "reference standard" for detecting atrial fibrillation on an electrocardiogram.[4,5] In this study, we used two consultant cardiologists, with a third arbitrating as necessary. The agreement between the cardiologists was very high (over 99%), confirming that diagnosis of atrial fibrillation can be made reliably through the reading of an electrocardiogram by a physician with relevant training and experience.

The response rate was reasonably high; 84-86% of participants returned their electrocardiograms. Non-respondents may have been less accurate at interpreting the electrocardiograms.

A primary care practitioner would not normally be sent 100 electrocardiograms to read, so it may be that the electrocardiograms were not read as carefully as they would have been in a clinical situation. The circumstances were also artificial in that the primary care practitioners did not have access to other clinical information when interpreting the electrocardiogram. However, symptoms and signs are of limited use in the diagnosis of atrial fibrillation.

Combining the results of the interpretative software with interpretation by a general practitioner led to some improvement in sensitivity, but at a cost of lower specificity. Therefore, although the lack of access of the study practitioners to the character of the pulse and the software interpretation might have led to an underestimate of sensitivity of interpretation of the electrocardiogram in the real world setting, we may have overestimated specificity.

Implications

If more than half of diagnoses of atrial fibrillation in primary care are incorrect, this might lead to a lot of unnecessary referrals. Conversely, if a screening programme relied on reading electrocardiograms in primary care, about 20% of cases of atrial fibrillation would be missed. Computer software performed much better, but still had an error rate sufficiently high to mean that decisions on treatment cannot be based on diagnosis by computer alone, even when combined with interpretation by a general practitioner. Therefore, strategies to identify atrial fibrillation in the community need to take into account how and by whom the electrocardiogram will be interpreted. If primary care practitioners are to detect atrial fibrillation on an electrocardiogram reliably, they need appropriate training. Alternatively, electrocardiograms generated in primary care will need to be sent to a specialist for accurate interpretation.

The recent guideline for atrial fibrillation from the National Institute for Health and Clinical Excellence recommends that electrocardiography is used to diagnose atrial fibrillation.[6] This study suggests that quality control of interpretation of electrocardiograms is an important part of diagnosing atrial fibrillation in primary care.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Electrocardiography is recognised as the standard investigation for diagnosing atrial fibrillation

Little evidence exists as to whether primary care practitioners can reliably diagnose atrial fibrillation on an electrocardiogram

WHAT THIS STUDY ADDS

Although a few primary care practitioners could diagnose the presence or absence of atrial fibrillation accurately on an electrocardiogram, most could not

The accurate identification of atrial fibrillation in the community requires that electrocardiograms are read by appropriately trained people

We are grateful for the support of the practices that participated in the SAFE study and for the support provided by the Midlands Research Practices Consortium.

Contributors: See bmj.com.
Funding: The work was funded by the Health Technology Assessment Programme. The authors are independent from the funders of the research. The views expressed in this publication are those of the authors and not necessarily those of the funders or the Department of Health.
Competing interests: None declared.
Ethical approval: West Midlands Multi-centre Research Ethics Committee approved the SAFE study.
Provenance and peer review: Non-commissioned; externally peer reviewed.

1 Salerno SM, Alguire PC, Waxman HS. Competency in interpretation of 12-lead electrocardiograms: summary and appraisal of published evidence. *Ann Intern Med* 2003;138:751-60.
2 Poon K, Okin P, Kligfield P. Diagnostic performance of a computer based ECG rhythm algorithm. *J Electrocardiol* 2005;38:235-8.
3 Hobbs FDR, Fitzmaurice DA, Mant J, Murray E, Jowett S, Bryan S, et al. A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over: the SAFE study. *Health Technol Assess* 2005;9:40.
4 Somerville S, Somerville J, Croft P, Lewis M. Atrial fibrillation: a comparison of methods to identify cases in general practice. *Br J Gen Pract* 2000;50:727-9.
5 Wheeldon NM, Tayler DI, Anagnostou E, Cook D, Wales C, Oakley GDG. Screening for atrial fibrillation in primary care. *Heart* 1998;79:50-5.
6 National Collaborating Centre for Chronic Conditions. *Atrial fibrillation: national clinical guideline for management in primary and secondary care*. London: Royal College of Physicians, 2006.

Accepted: 21 May 2007