

# Do computer generated ECG reports improve interpretation by accident and emergency senior house officers?

S Goodacre, A Webster, F Morris

## Abstract

**Objectives**—To determine whether access to a computer generated electrocardiogram (ECG) report can reduce errors of interpretation by senior house officers (SHOs) in an accident and emergency department.

**Methods**—Ten SHOs were asked to interpret 50 ECGs each: 25 with computer generated reports, 25 without. Their answers, and the computer generated reports, were compared with a “gold standard” produced by two experienced clinicians. The primary outcome measure was the proportion of major errors of interpretation. **Results**—The computer reading system made two major errors (4%, 95% confidence interval (CI) 1.1% to 13.5%) compared with the gold standard. Access to the computer report did not significantly reduce major errors among SHOs (46 (18.4%) with report *v* 56 (22.4%) without, odds ratio 0.64, 95% CI 0.36% to 1.14%, *p*=0.13) or improve the proportion completely correct (104 (41.6%) with report *v* 91 (36.4%) without, odds ratio 1.43, 95% CI 0.88 to 2.33, *p*=0.15).

**Conclusions**—SHOs have a high error rate when interpreting ECGs, which is not significantly reduced by access to a computer generated report. Junior doctors should continue to seek expert senior help when they have to interpret a difficult ECG.

(*Postgrad Med J* 2001;77:455–457)

Keywords: electrocardiogram; computer reading systems; accident and emergency

Senior house officers (SHOs) in accident and emergency medicine will have to analyse hundreds of electrocardiograms (ECGs) during a six month post despite reported error rates of interpretation as high as 20%.<sup>1,2</sup> Currently junior doctors rely on senior staff in the department for help in interpreting difficult ECGs, which may not be available 24 hours of the day. Although few errors translate into errors of clinical management, their potential impact may be considerable if they result in inadvertent discharge of those with serious pathology<sup>3</sup> or delayed recognition of myocardial infarction in which thrombolysis is indicated.<sup>4</sup>

Over the last 20 years computer programs have been developed to read and interpret

ECG recordings. Many modern ECG machines now routinely provide a computer generated ECG report. The accuracy of this interpretation has been reported to approach the level of a cardiologist.<sup>5,6</sup> Erroneous reports do occur, however, so there are concerns that physicians may be misled by computer generated reports.<sup>7</sup> The impact of access to a computer generated ECG report has been studied under experimental conditions with a variety of specialties of clinician with generally favourable results.<sup>5–7</sup> To our knowledge, no study has evaluated whether computer generated ECG reports improve the accuracy of interpretation of a group of clinicians who might find them most valuable: SHOs in an accident and emergency department.

We aimed to test the hypothesis that SHOs in an accident and emergency department will make fewer errors of ECG interpretation if they are given ECGs with a computer generated report, compared with interpretation without such a report.

## Subjects and methods

The study was carried out in the accident and emergency department at the Northern General Hospital, Sheffield. Over a one month period the Marquette M500 machine (Marquette Medical Systems Incorporated, Milwaukee, USA) was used to collect ECGs from the normal workload of the department. This machine has been shown to provide accurate ECG interpretation in a previous trial.<sup>6</sup> Fifty ECGs were chosen from those collected to present a varied and challenging test to the study group of junior doctors. These ECGs were randomly sorted and labelled from one to 50. All ECGs initially had a computer generated report. A copy of each ECG was then produced with the report removed.

Ten SHOs participated in the study. All had graduated two to three years earlier. Under strict examination conditions they were given one minute to analyse each of the 50 ECGs. They were instructed that they should treat each ECG as if a nurse had presented it to them in the accident and emergency department, and to write down any abnormality identified. For this process the SHOs were randomly allocated to two groups. The first group were provided computer reports only on even numbered ECGs in the series, the second group were provided computer reports only on odd numbered ECGs. Thus each of the 50 ECGs was analysed by 10 SHOs—five of whom were able to read the computer generated report, five of whom were not. This

Department of  
Accident & Emergency  
Medicine, Northern  
General Hospital,  
Herries Road,  
Sheffield S5 7AU, UK  
S Goodacre  
A Webster  
F Morris

Correspondence to:  
Dr Goodacre  
S.Goodacre@sheffield.ac.uk

Submitted 29 August 2000  
Accepted 23 January 2001

Table 1 Criteria for major or minor errors

Definition	Criteria
Minor error	First degree heart block, ectopic beats, sinus tachycardia <120 beats/min, sinus bradycardia >40 beats/min, non-specific T wave changes, ventricular hypertrophy, left or right axis deviation, right bundle branch block, Q waves
Major error	ST segment depression or elevation >1 mm, widespread/deep T wave changes, sinus tachycardia >120 beats/min, sinus bradycardia <40 beats/min, atrial fibrillation, narrow or broad complex tachycardias, second or third degree heart block

Table 2 Major errors reported by the computer reading system

Computer report	Correct interpretation by SHO	Minor error by SHO	Major error by SHO
Pacing spike missed leading to erroneous interpretation of paced rhythm	2	0	3
Atrial fibrillation misdiagnosed as sinus rhythm	3	1	1

gave a total of 500 SHO ECG interpretations: 250 made with access to a computer generated report, 250 without.

The "gold standard" for each ECG report was produced by two accredited accident and emergency clinicians with extensive experience in ECG reporting (SG and FM<sup>6</sup>). Both clinicians interpreted each of the 50 ECGs without referring to the computer report. They then conferred and agreed on a correct interpretation, to produce a gold standard answer. A list of potential errors of interpretation, compared with the gold standard, was drawn up and each potential error classified as major or minor (see table 1).

To determine the proportion of ECGs erroneously reported by the computer, the computer generated reports were compared with the gold standard answers and errors graded according to the table. To determine the error rate of the SHOs, with and without access to a computer generated report, each of the 500 interpretations produced by an SHO were compared independently by both senior clinicians to the gold standard, looking for additions or omissions, that were classified as minor or major errors in interpretation (see table 1). The assessors were blind to whether the SHO had a computer generated report. To measure agreement between the assessors, their decisions were compared and a  $\kappa$  score calculated. Any disagreements were then resolved by discussion.

The principal analysis was a comparison of the accuracy of the two groups of SHO ECG interpretations—those made with access to a computer generated ECG report versus those without. The primary outcome measure was the proportion of major errors missed in each group. The secondary outcome measure was the number of completely correct ECGs, that is without major or minor errors.

#### POWER CALCULATION

A similar previous study suggested SHOs would report approximately 50% of ECGs completely correctly and significant errors in approximately 20%.<sup>2</sup> A sample size of 500 was required to detect a change of 12.5% in the minor error rate or a change of 10% in the major error rate (alpha 0.05, beta 0.2).

#### STATISTICAL ANALYSIS

To determine whether availability of a computer generated report was an independent predictor of either major error or a completely correct report, logistic regression was performed using ECG, SHO, and computer report as cofactors. Analysis was carried out using SPSS for Windows, version 9 (SPSS Inc, Chicago, 1998).

#### Results

The two senior clinicians assessed 500 SHO ECG interpretations each and agreed with each other's assessment on 452 occasions ( $\kappa = 0.85$ , indicating excellent agreement). Consensus was reached on the remaining 48.

The computer reading system made two major errors in reporting the 50 ECGs (error rate 4%, 95% confidence interval (CI) 1.1 to 13.5%). The five SHOs who had access to the erroneous reports did not consistently reproduce these errors (see table 2).

Major errors were found in 46 out of the 250 ECG interpretations made by SHOs with access to the computer generated report (18.4%), compared with 56 out of the 250 interpretations made without a computer generated report (22.4%). Logistic regression showed no evidence of a relationship between use of a computer generated report and major errors of interpretation by the SHO. The adjusted odds ratio for major error with a computer generated report was 0.64 (95% CI 0.36 to 1.14,  $p=0.13$ ).

A total of 104 out of the 250 (41.6%) ECG interpretations made by SHOs with access to the computer generated report were completely correct, compared with 91 out of the 250 (36.4%) without. Logistic regression showed no evidence of a relationship between use of a computer generated report and completely correct interpretation by the SHO. The adjusted odds ratio for correct interpretation with a computer generated report was 1.43 (95% CI 0.88 to 2.33,  $p=0.15$ ).

#### Discussion

This study confirms previous findings,<sup>1,2</sup> that SHOs have a high rate of erroneous ECG reporting. By contrast, the Marquette M500 machine reported major errors in only 4% of ECGs. Despite this our study showed that there was no significant difference in the ECG error rate when junior doctors had assistance from computer generated reports. Before rejecting computer generated ECG reporting we must consider two possibilities: (1) did the study lack statistical power to detect a clinically significant difference? or (2) could SHO performance with the computer report have been improved if they had given more credence to the report?

The 95% CIs of our estimates include an odds ratio of 0.36 for the major error rate and 2.33 for completely correct reporting. However, by selecting challenging ECGs and setting a time limit, the study methodology was designed to maximise erroneous reporting. Previous studies have shown that apparently significant errors do not usually have a negative

impact upon clinical practice.<sup>1 2</sup> It is unlikely therefore that we have failed to detect a difference that would have been important in clinical practice.

The discrepancy between the SHO error rate (with access to the computer report) and the computer error rate suggests that SHOs tend to ignore its advice. Perhaps it would be worth repeating the study with the SHOs told to ignore the computer report only if they were sure the analysis was incorrect. Indeed it might be suggested that SHOs would perform better if they simply repeated the computer report verbatim. Such advice would be unfair and unwise. This study forced relatively inexperienced SHOs to make definitive judgments upon difficult ECGs. No allowance was made for prevarication or uncertainty. In reality, we would expect SHOs to seek senior advice when faced with this scenario.

However, the accuracy of computer reports compared to those of SHOs suggests that, while they may not replace a senior opinion, computer reports may have some value in alerting SHOs to potential errors. Practice might be improved by instructing SHOs to seek senior advice whenever they disagree with, or do not understand, a report.

In summary this study has confirmed that junior doctors have a high error rate in reporting

ECGs. Computer generated reports did not significantly improve this, even though the machine achieved a low major error rate compared with the SHOs. Computer generated reports may have a role in prompting SHOs to query their own ECG interpretation but should not replace experienced medical support.

Thanks to Karen Angelini and Jane Arnold for collecting the ECGs during the study period.

- 1 Morrison WG, Swann IJ. Electrographic interpretation by junior doctors. *Arch Emerg Med* 1990;7:108–10.
- 2 White T, Woodmansey P, Ferguson DG, *et al.* Improving the interpretation of electrocardiograms in an accident and emergency department. *Postgrad Med J* 1995;71:132–5.
- 3 Lee T, Rouan G, Weisberg M. Clinical characteristics and natural history of patients with acute myocardial infarction sent home from the emergency room. *Am J Cardiol* 1987;60:219–24.
- 4 Lee TH, Weisberg MC, Brand DA, *et al.* Candidates for thrombolysis among emergency room patients with acute chest pain. *Ann Intern Med* 1989;110:957–62.
- 5 Brailer DJ, Kroch E, Pauly MV. The impact of computer assisted test interpretation on physician decision making: the case of electrocardiograms. *Medical Decision Making* 1997;17:80–6.
- 6 Willems JL, Abreu-Lima C, Arnaud P, *et al.* The diagnostic performance of computer programs for the interpretation of electrocardiograms. *N Engl J Med* 1991;325:1767–73.
- 7 Hillson SD, Connelly DP, Liu Y. The effects of computer assisted electrocardiographic interpretation on physicians' diagnostic decisions. *Medical Decision Making* 1995;15:107–12.
- 8 Brady WJ, Morris F. Electrographic abnormalities encountered in acute myocardial infarction. *J Accid Emerg Med* 1999;17:40–5.