Audit of computer-aided diagnosis of abdominal pain in accident and emergency departments

P. A. STONEBRIDGE, P. FREELAND, J. B. RAINEY & D. A. D. MACLEOD
St John's Hospital, Livingston, West Lothian

INTRODUCTION

The initial overall diagnostic accuracy of computer-aided diagnosis of abdominal pain (CAD) was reported as over 94% (de Dombal et al., 1972). In our own hospital it produced a rise in diagnostic accuracy from 55 to 78%, and the immediate discharge rate from the A&E Department from 16 to 34%, plus a fall in the unnecessary laparotomy rate from 18 to 7% (Gunn, 1976). These results prompted a multicentre trial, which showed an improvement in diagnostic accuracy (45.6 to 65.3%) and a reduction in the negative laparotomy rate (25.2 to 10.4%) (Adams et al., 1986). A recent report has also suggested that the short term benefits have been maintained over a much longer period and that such systems should ‘be offered to all DGHs in the U.K. (McAdam et al., 1990). Despite this reported benefit there has not been a wide uptake of such systems (Sutton, 1989). Further the benefits from the use of computer-aided teaching packages as a decision support is a matter of current debate (Sutton, 1989; Kennedy et al., 1990).

The advantage of CAD lies in an improved diagnostic accuracy and its consequences, over the unaided clinician. The purpose of this report is to analyse the percentage diagnostic accuracy rates for CAD and the A&E junior staff over a 15-year period to see if CAD has maintained its early enthusiastic promise.

PATIENTS, METHODS AND RESULTS

CAD was set up in the West Lothian Surgical Service in 1975 and all patients presenting with acute abdominal pain (21324) have been entered to date. Thirty-three data points using a structured case record are entered by the junior A&E

Correspondence to: Mr P. A. Stonebridge, St John's Hospital, Livingston, West Lothian.
staff (senior house officer grade). A full description of the data points and methods of analysis and an analysis of the systems strengths and weaknesses has been previously published (Gammerman & Thatcher, 1991). From these data the computer, using Bayesian probabilistic analysis, places the patient into one of eight diagnostic groups (i.e. appendicitis, diverticulitis perforated peptic ulcer, non-specific abdominal pain, biliary disease, intestinal obstruction, pancreatitis, renal colic or dyspepsia) available to the computer program with the highest percentage probability. The diagnosis is recorded along with the A&E junior staff diagnosis before the data are entered into the CAD system. The patient's final diagnosis, based on clinical course, investigations, operative findings and pathology specimens is recorded on eventual discharge. There was no follow up of discharges other than review of emergency re-admissions. The accuracy of the initial computer and medical staff diagnosis then compared with the final diagnosis.

The diagnostic accuracy of CAD and the A&E junior staff over the 15-year period 1975–1989 is shown in Fig. 1.

**DISCUSSION**

This study has revealed a problem with CAD as previously commented on by Sutton (1989). The junior A&E staff diagnostic accuracy has remained fairly constant (range 60.4–70.0%), whilst there has been a gradual fall in the diagnostic accuracy of the CAD from an early high of 78.5% to a low of 55%. This gradual fall in the diagnostic accuracy of CAD is difficult to explain particularly as the computer hardware, program and datasheet have been constantly updated, though the data entered has remained the same. Logically, therefore, the cause must be the human part of the system. Part of this may be a certain lack of faith in CAD as it can be seen to be obviously wrong in individual clinical circumstances. As a result entering the data is often perceived as a chore rather than of actual use, especially as computers are no longer novel. An alternative reading is that the early years produced elevated accuracies through a 'study-effect' due to enthusiasm, novelty and interest, while more recent results represent the product of 'routine use'.

![Fig. 1. Diagnostic accuracy of acute abdominal pain of a computer-aided diagnosis system and A&E junior medical staff over a 15-year period (1975–1989).](image-url)
How do we investigate and correct this problem? A decision to withdraw CAD in its present form has been made, though the structured datasheets are being retained. This will allow the impact of CAD withdrawal on diagnostic accuracy and junior A&E staff training to be assessed. It will also allow the introduction of labour and time saving scanning ‘card-readers’ and then enable an assessment of its updated re-introduction.

CAD is not an outdated concept, but it does require audit of its performance and in our setting change. It also remains a unique, highly valuable and frequently used ongoing database, auditing aspects of and producing many additions to the literature on the subject of the diagnosis and management of acute abdominal pain.

REFERENCES