

GUEST EDITORIAL

Computers, diagnoses and patients with acute abdominal pain

F. T. DE DOMBAL

Clinical Information Science Unit, University of Leeds, Leeds

It is now almost a quarter of a century since the use of computers to assist with the management of patients suffering from acute abdominal pain was proposed (de Dombal *et al.*, 1972). During this time — despite repeated claims of benefit associated with the introduction of computer packages — their use has not become widespread. In the interim opinion has polarized, with unhelpful epithets such as 'old-fashioned luddite' and 'quack robodoc' taking the place of rational scientific debate (Sutton, 1989). None of this of course helps the practising A&E clinician. Perhaps it is time to take stock of the situation.

Two helpful recent developments have clarified considerably both the position and possible role of computers in assisting with the management of patients with acute abdominal pain. The first of these has been the publication over the last few years of several very large series with relatively consistent results. Thus in the early days of 'computer-aided diagnosis' results from small series in single centres could be (and were!) ascribed to the wild enthusiasm of the innovators. However, such criticisms are more difficult to sustain given the findings from three recent large scale series involving controlled studies in eight U.K. hospitals (Adams *et al.*, 1986), a 12-year study in a single DGH (McAdam *et al.*, 1990) and — most recently — a report from no less than 64 European institutions (de Dombal, 1991) all confirming broadly the earlier single-centre findings.

Taken together, the consistent findings of these studies are impressive. In a total of over 37 000 cases they confirm: (a) that unaided performance by inexperienced clinicians is not very good in this difficult area of medicine; (b) that computers can out-perform unaided clinicians — but that; (c) *aided* clinicians improve to match (and sometimes exceed) the enhanced computer performance. The combined final performance figures for doctors in these trials are impressive, and incidentally set important (and attainable) benchmarks for the future (combined initial diagnostic accuracy 66%, accuracy after investigation 76%, perforated appendix rate 12.5%, and negative appendicectomy rate 16%).

Overall, computer accuracy in these studies has been similar to the (enhanced) performance levels of (aided) doctors; and this raises a second important question — namely, just how are these improvements in clinical performance obtained?

Clearly it cannot be purely due to the computer (for the computer is no more accurate than the aided clinician). Clearly the computer package must contribute something (since when it is provided, whether for feedback, audit or teaching (de Dombal *et al.*, 1991), — performance improves — and when it is removed, performance regresses).

The exploration of this point has been the second important development in the field. Increasingly, opinion has hardened towards the concept that the computer is acting not as a surrogate doctor but as a catalyst, as a repository of 'good clinical practice' (determined by an extensive peer-group) and as a stimulus towards following appropriate clinical pathways (again as determined by peer-group). Increasingly, the concept of a clever computer has been replaced by the concept of 'objective medical decision-making' — using computers to encourage young doctors to follow the pathways and precedents of their experienced peers. The slogan 'If in doubt, ask the computer for advice' has been replaced by 'Do it right — get it right' (Lavelle, 1990) — and many find this latter concept much more to their liking.

In this context, the current study from West Lothian, (Stonebridge *et al.*, 1992) represents an important milestone. It too reports careful work over more than a decade — with impressive numbers of patients studied (over 20 000 in all). It too reports diagnostic accuracy levels for clinicians (range 60.4–70.0%, consistent over 15 years), of which the authors and hospital should be justifiably proud. It too suggests that the enhanced performance levels observed are the result of 'good clinical practice' rather than slavish adherence to the computer — and it too goes on to explore the role of the computer.

The study has much to offer computer aficionado, clinician, and clinical scientist alike. To the former group, the paper is a gold-mine. Consider for example the computer's performance — overall similar to the SHO's as regards accuracy. Simplistic detractors would conclude the computer has *precisely nothing to offer*. Enthusiasts would counter that the computer performance was roughly 20% better than most unaided A&E staff.

The present authors have fallen into neither of these traps. Indeed, their finding that the computer performance for the first few years exceeded that of even the best clinicians; but gradually deteriorated — until it was regularly exceeded by the clinicians performances — is both important and fascinating. Was this (as the authors suggest) because of gradual deterioration of data fed into the computer as clinicians tired of their new 'toy'? Was it *despite* regular updates in the computer's databanks — or was it perhaps *because* such updates took place with 'new' SHO's contributing less meticulous data? And if so, is this because the computer's 'diagnosis' was presented in the wrong way? Should the computer, instead of indicating '70% appendicitis' as its 'diagnosis', merely use this prediction as a starting point, and (as in the current Leeds system) then search 10 000 more files, identify (some hundreds of) patients similar to the present case, note what *actually* happened to these patients and display this information? (de Dombal *et al.*, 1992). Would this be more reliable? Would this be more helpful? Would this be more acceptable?

These are all fascinating questions which will need to be answered if the computer's role is to be expanded in future. For the clinician however, the message

from this study is very clear and complements that of other studies performed elsewhere. In this day and age, the performance levels described earlier *can* be attained over a period of decades and (since they seem mostly to depend upon young doctors 'doing it right') there would seem little or no excuse for failing to adopt the simple methods common to all the studies (such as insisting upon the use of structured data collection forms). It might not surprise the proverbial visitor from Mars that the routine use of computers for decision support provoked some anxiety — but the same visitor would regard refusal to adopt data-collection procedures with incredulity as they have shown time and again to be associated with improved performance.

Perhaps the greatest contribution of the work in West Lothian, however, has been to demonstrate how much careful analysis of meticulously collected data has to offer. In this field — as one suspects in many others affecting emergency care — optimal use of the technology available to us will only be obtained through the kind of careful studies described here. The present analysis is an object lesson to those 'wild-blue-yonder' enthusiasts or 'disgusted of Tunbridge Wells' detractors who between them have done much to retard the sensible implementation of computer-packages throughout medicine. Many of the questions raised by the present study are amenable to further study. For example, the authors plan to withdraw the direct use of the computer in its present form, rethink the *modus operandi*, watch what happens, and maybe proceed to an up-dated version's re-introduction. One can only applaud such ideas, wish them well, and look forward to seeing the results in the future. Such studies are surely entirely consistent with the development of a solid scientific foundation for emergency medicine.

We have taken stock. The message from the present state of the art is quite clear. There is, right now, improved performance to be had by getting young doctors to follow the precepts and practices of their peers. It would be absolutely astonishing if it were not so. In this process, which includes such practices as the use of structured data-collection forms, the computer may currently play an important but rather limited role, acting as a stimulus, giving self-audit to individual doctors, and possibly also teaching them. For a number of reasons (such as the increasing inability of all of us to keep up with our chosen field) one suspects extensions of computer's roles are inevitable. This process will be greatly aided by such careful studies as those of the West Lothian group.

REFERENCES

- Adams I. D., Chan M., Clifford P. C. *et al.* (1986) Computer-aided diagnosis of abdominal pain: A multi-centre study. *British Medical Journal* **293**, 800–804.
- de Dombal F. T. (1991) Objective Medical Decision Making, Acute Abdominal Pain. Final Report to COMAC-BME of the European Community. *EC Publication, Brussels*.
- de Dombal F. T., Barnes S., Dallos V. *et al.* (1992) How should computer aided decision support systems present their predictions to the practising surgeon? *Theoretical Surgery* (in press).
- de Dombal F. T., Dallos V. & McAdam W. A. F. (1991) Can computer-aided teaching packages improve clinical care in patients with acute abdominal pain. *British Medical Journal* **6791(302)**, 1495–1497.
- de Dombal F. T., Leaper D. J., Staniland J. T., Horrocks J. C. & McCann A. P. (1972) Computer-aided

- diagnosis of acute abdominal pain. *British Medical Journal* 2, 9–13.
- Lavelle S. M., Dawids S. G. & van Beneken J. E. W. (1990) The initiative on medical decision-making of the Concerted Action Committee on Biomedical Engineering (COMAC-BME) of the European Community. *Theoretical Surgery* 5, 107–111.
- McAdam W. A. F., Brock B. M., Armitage T., Armitage T., Davenport P. & de Dombal F. T. (1990) Twelve years experience of computer-aided diagnosis in a District General Hospital. *Annals of the Royal College of Surgeons* 72, 140–146.
- Stonebridge P. A., Freeland P., Rainey J. B. & Macleod D. A. D. (1992) Audit of computer aided diagnosis of abdominal pain in A&E. *Archives of Emergency Medicine* 9, 271–273.
- Sutton G. (1989) Cited in 'Physicians give 'penny arcade' robotoc the big thumbs down'. *Datalink*, Oct 30, pp. 3.