A. Hasman

Department of Medical Informatics,
University of Maastricht,
The Netherlands

Commentary

Computer-Aided Diagnosis of Abdominal Pain

Reflections on F.T. De Dombal et al.'s paper:
Computer-Aided Diagnosis of Acute Abdominal Pain

Introduction

This commentary is about one of a series of papers written by the late Tim de Dombal and colleagues, published in the British Medical Journal [1]. They report on a controlled prospective unselected real-time comparison of human and computer-aided diagnosis in a series of 304 patients suffering from abdominal pain of acute onset. The investigation was carried out in the professorial surgical unit of the General Infirmary in Leeds (UK). The computer-aided diagnostic system and some of the problems inherent to its implementation were described in another article [2].

According to the authors, the choice of the "acute abdomen" was a deliberate one, since such an area of diagnosis has several advantages. It is a common clinical dilemma (304 cases presented in less than one year). The number of possible diagnoses is relatively small, as evidenced by the low percentage of "unclassifiable" patients, the clinical diagnosis is usually made on the basis of a patient's symptoms and physical signs rather than on biochemical tests; and the final diagnosis is usually made at surgery.

The abdominal pain system is fed with medical history data and data about the physical examination of the patient. Then by using Bayes' rule it determines the probabilities of a set of seven diseases, including nonspecific abdominal pain, that most often cause abdominal pain.

The probabilities needed for applying Bayes' rule were derived from real patient data. Several researchers at that time advocated the use of subjective probability estimates instead of probabilities derived from real patient data. The former would make large-scale surveys to obtain enough patient data superfluous. To determine whether this was a good approach, clinicians' estimates of the probabilities were collected and the performance of the system based on real patient data was compared with the performance of the system using the clinicians' estimates. It was observed that the performance of the system reduced significantly when using clinicians' estimates. It was, therefore, concluded that real-life data from large-scale surveys were still necessary [3].

The abdominal pain system was not the first decision-support system to be implemented. Many decision-support systems that used Bayes' rule for calculating posterior probabilities of diagnoses had already been reported in the literature. Also other methods were investigated such as decision trees, branched sequential testing using entropy measures, and linear discriminant analysis. At that time it became apparent that the differences in performance were not so much due to the methodology used but more due to the relevance of the data used in the analysis. These studies usually investigated the feasibility of computer-aided diagnosis and were not extensively used for supporting clinical practice. The abdominal pain system was one of the first to be used in practice and the first to be evaluated thoroughly.

The HELP system of the Latter Day Saints hospital showed that decision support techniques can also be used in a broader context; in this case the decision-support system was integrated with a HIS that provided the necessary data.

What was Achieved?

A very important aspect of the work of de Dombal and colleagues is the thoroughness with which they developed and evaluated the system. Attention was paid to problems that could occur when capturing data, the necessary probabilities were determined from a database of patient data and the output of the system was validated in a controlled prospective trial. All items that possibly could be used by the system were checked for observer variation. It was found that certain items were interpreted quite different

by different clinicians. Such items were not included and items that were included were accurately defined. In this way data capturing errors were minimized.

A structured form was developed on which the data needed by the system could be documented by the clinician. The use of the form appeared to increase the performance of the clinician, probably because when using it more relevant data were captured than without it. Initially, these data were entered into the system by a research assistant and not by the clinician. In those days the system interface was not yet very user friendly. The system was validated in a controlled prospective trial in which the diagnostic performance of the unaided clinician was compared with that of the system. Because of the choice of the selected clinical conditions a gold standard was available, which made the evaluation of a system easier to perform. It appeared that the system performed better than the clinicians, even than the most senior ones. From the studies it was inferred that the use of the system would have caused fewer unnecessary laparotomies to be been performed and patients would have stayed in the hospital for a shorter period of time. In a later study, Adams et al. [4] showed (in 16,737 patients attending UK hospitals) actual improvements in diagnostic accuracy (from 45% to 65%), decision-making performance (negative laparotomy rates halved) and resource utilization (saving of over 5 million British pounds), associated with the use of the abdominal pain system, in the mean time implemented on desktop computers. These results were replicated in a number of other UK studies as well as in many of the hospitals, taking part in a European Concerted Action involving 19 countries and in the worldwide studies conducted by the Research Committee of the World Organization of Gastroenterology [5].

When a system is performing well in one institution it is not apparent that the system will also be accepted in other institutions. Many decision-support systems never left the place where they were developed. The abdominal pain program could be used successfully in other locations. However, problems were encountered during transferring the system. The quality of the advice of the system is, among others, dependent on the referral policy. Different referral strategies may result in different prior probabilities of the diseases. Since the acute abdominal pain program only deals with a limited set of diagnoses, its use also depends on, the judgment of a clinician whether to present the patient data to the system or not. It was also apparent that the interpretation of terms was different at different locations. This may not only lead to different conditional probabilities; the use of the system may also become dangerous under such circumstances. Therefore, much additional work had to be performed (e.g., training of clinicians) before the system could be optimally utilized at other locations.

Broadening the scope of the system beyond the small domain of acute abdominal pain (e.g., by including gynecological disorders, dyspepsia, etc.) appeared not to be successful.

Lessons Learned

The approach that de Dombal and colleagues adopted is still valid when designing and developing systems to be used in new and not yet totally understood domains. An example of such a domain is the area of electronic patient records. In this area a number of aspects is still not well understood and therefore deserves further study. In such situations, existing ideas about how to process the data should not be taken for granted but their value or truth should be investigated.

Nowadays, prototypes of systems can be quickly developed. The user can be involved and can indicate whether the interface is according to his needs. In new areas of investigation the underlying processes usually are not fully understood. Therefore, in such a case prototyping, having in mind that in this way the user requirements can be determined, is not enough. The prototype should also incorporate capabilities with which possible alternative hypotheses about the underlying processes can be tested [6]. This is exactly what de Dombal and colleagues have done. They did not develop a system only on the basis of perceived user needs without bothering about the quality of the components, but also investigated whether several of the factors on which computer-aided diagnosis was based were scientifically sound. For example, they did not simply accept the fact that the terminology used by the clinicians can be used without problems. They first performed observer variation studies and only those items with an acceptable range of variation in interpretation were included.

Although the abdominal pain system performed well, it was not used on a large scale. The problem probably was too specific, given the amount of time it took physicians to obtain a diagnostic prediction. The desktop computer version took the clinician five minutes, which is far too long when 15 to 20 patients are seen daily. So, even though de Dombal et al. could prove that the system performed at expert level, it was not used regularly.

As mentioned above, making the scope of the program broader reduced the performance of the program. De Dombal concluded elsewhere that knowledge-based systems should deal with a single area of clinical medicine and, where possible, with a single clinical problem. He also noted that integration into clinical practice is vital; the system should be easy to use and should not take much time.

Both constraints - each knowledge-based system should cover a single clinical problem and the system should be easy to use - can only be met when these knowledge-based systems are directly coupled to a computer-based patient record (CPR) system. If the CPR system can cope with data entry and terminology problems, a direct feedback from the coupled knowledge-based system is possible.

De Dombal was thinking in terms of separate systems that could be used by the physician to support him in areas in which he needs support. However, one may question whether a clinician will use such a system in a busy practice. Usually, clinicians are certain of their diagnosis or therapy and do not see why they should use such a system.

Vissers [7] described a protocol system to support treatment decisions in the A&E department. The protocol system contained treatment information for isolated fractures without concomitant lesions. The impact of the system on the protocol adherence of residents was investigated. In that study, the residents were first asked how certain they were about their own diagnosis and therapy (before consulting the system). It appeared that the residents were more than 90% confident, both about their diagnosis and treatment plan, although their diagnosis, for instance, was correct in less than 60% of the cases. Even when the protocol system is available, it will not be used for checking decisions that are considered to be almost certain.

From these and other experiences I conclude that decision-support systems such as the one presented by de Dombal and colleagues will only be used on a large scale in clinical practice when incorporated in a system in which patient data are also stored on a routine basis.

Conclusion

De Dombal and colleagues indicated already more than 25 years ago how decision-support systems should be developed. They had to use relatively primitive computer hardware and the use of the system was time consuming. Therefore, the practical use of the system was not extensive. Yet, their approach is prototypical for the way systems should be designed, that is, start with a system that is needed by physicians and do not develop systems that are meant to simulate physicians but, instead, that support them. As de Dombal once noted: support systems are to support. This means that - although duplication of human thought patterns may well be an interesting research task - systems designed for immediate practical use would do well to concentrate upon complementing human weaknesses, instead of competing with human strengths. Another message is to take nothing for granted when developing systems, for instance, by assuming that the terminology used by the physicians is interpreted in the same way by each of them, and evaluate the system thoroughly. Especially when using knowledge-based systems, the users need to be assured that the system has been tested in-depth and has shown to be of value. De Dombal and colleagues have shown at least three important things:

1. Computer systems can be designed that can be shown to perform at expert level. Use of these systems by inexperienced clinicians may improve the performance of these clinicians to the point where it matches that of their senior colleagues.
2. Knowledge is increasing at such a rapid pace that decision-support systems are needed by clinicians, whether they like it or not.
3. Such systems will only be used in practice if they fit perfectly within that practice.

It is sad that Tim de Dombal died too young. The approach to system development, as advocated by him and his colleagues, however, will still be of value in the next century.

References

[1] de Dombal FT, Leaper DJ, Staniland JR, McCann AP, Horrocks JC. Computer-aided diagnosis of acute abdominal pain. Brit Med J 1972;2:9-13.

[2] Horrocks JC, McCann AP, Staniland JR, Leaper DJ, De Dombal FT. Computer-aided diagnosis: description of an adaptable system, and operational experience with 2,034 cases. Brit Med J 1972; 2:5-9.

[3] Leaper DJ, Horrocks JC, Staniland JR, De Dombal FT. Computer-assisted diagnosis of abdominal pain using estimates provided by clinicians. Brit Med J 1972;4:350-4.

[4] Adams ID, et al. Computer-aided diagnosis of abdominal pain: a multi-centre study. Brit Med J 1986;293:800-4.

[5] Myren J, Bouchier IAD, Watkison G, De Dombal FT. Report of the World Organisation of Gastroenterology. Scand J Gastroenterol Suppl 1988;23:144.

[6] Hasman A, Tange H, Vissers M. Combining a scientific approach and prototyping in the design of EHCR systems. IEEE T Inform Technol Biomed 1998;2:117-23.

[7] Vissers M, Biert J, Van der Linden CJ, Hasman A. Effects of a supportive protocol processing system (ProtoVIEW) on clinical behaviour of residents in the accident and emergency department. Comp Meth Progr Bio 1996, 49:177-84.

Address of the author:
A. Hasman
Department of Medical Informatics,
University of Maastricht,
PO Box 616,
6200 MD Maastricht
The Netherlands
E-mail: hasman@mi.unimaas.nl