Pergamon

Information Processing & Management, Vol. 34, No. 1, pp. 133–134, 1998
Published by Elsevier Science Ltd
Printed in Great Britain
0306-4573/98 $19.00+0.00

PII: S0306-4573(97)00056-3

BOOK REVIEW

Rationalizing Medical Work: Decision-Support Techniques and Medical Practices. M. BERG. MIT Press, Cambridge, MA (1997). x+238 pp., ISBN 0-262-02417-9, $30.

Making a medical diagnosis and precribing treatment is a complex intellectual task. What makes medical diagnosis and treatment difficult is the volume of factual knowledge that could potentially go into a decision, the inherent uncertainty of medical observations, our limited understanding of the cellular and molecular operations of the human body, and the cultural shadings surrounding human health and disease.

These phenomena explain some of the problems facing health care, which include (a) tremendous variation in medical practices across regions, communities, and individual physicians and (b) growing documentation of errors of omission and commission in medical practice. Variations in practice have been well-documented by Wennberg and colleagues at Dartmouth Medical School, who published the *Atlas of Health Care* (American Hospital Association, 1995). This publication documents, for example, that a person in Boston, Massachusetts, is twice as likely to receive coronary artery bypass surgery for the same level of symptoms as someone in Portland, oregon, though half as likely to receive surgery for an equivalent amount of back pain. These differences in practice cannot be explained by severity of illness or other known factors.

Errors of omission and commission are also well documented. The former is exemplified by showing the lack of adherence to scientifically validated treaments in such common diseases as diabetes (Weiner *et al.*, 1995) and myocardial infarction (Antman *et al.*, 1992). Errors of commission are demonstrated by the numerous studies that document excessive prescription of antibiotics by primary care physicians (Bernstein *et al.*, 1982).

Many solutions have been advocated for these problems, a number of which rely on methods of formalizing medical decision making. Most of these techniques involve the use of computer tools. The development of these tools began nearly three decades ago. Some have been written up in medical journals. Certainly the general public has a fascination with the notion of technological wonders that diagnose disease and dispense advice. But the reality is that the impact of these systems in clinical care has been negligible. Few physicians use them in the care of patients and, as the author of this book notes, there is little data to support their use.

If medical care is prone to inconsistencies and errors, and computer programs can be written to formalize medical thinking, why have these tools failed to have an impact? This is the topic of this fascinating new book by Marc Berg, a physician and Ph.D. sociologist who studied this topic during a postdoctoral fellowship at the University of Maastricht. Whereas most other reports about these system have focused on their acceptance and/or the fitness of their algorithms. Berg observed their use in actual settings, including some of the mundane but revealing aspects that are not reported in academic papers.

As Berg notes, there has been a long-standing fascination with the use of automated means to improve medical practice. Much early work focused on the application of artificial intelligence (AI) and decision technologies. The history of AI in medicine parallels the history of AI in general. In fact, medicine has always been a ripe domain for AI applications. The first backward-chaining rule-based expert system was the dissertation of Edward Shortliffe, an M.D./Ph.D. student from Stanford University in the early 1970s. Shortliffe's dissertation resulted in a system called MYCIN that generated advice on the diagnosis and treatment of blood infections.

By the 1980s, research in medical AI blossomed (Shortliffe, 1987). New techniques were explored. Studies of cases on the system showed AI applications to have an equal and even superior performance to human decision makers. At the end of the decade, some systems were installed in clinical settings to see if the success from controlled evaluations could be generalized to the real clinical setting. As Berg's book points out, successes in the real world proved elusive, often due to factors that have little to do with the technology.

Berg is well qualified to write a volume devoted to the sociocultural side of medical decision-support technologies. The book is well written and should be understandable by those without a medical background, though those who have had medical training will probably have a deeper understanding of the

clinical realities that have hindered these systems. The book is amply documented with footnotes and citations.

The book is not an attempt to provide an exhaustive history of decision support systems. Readers from the Western hemisphere might have preferred to read about systems that were developed on this side of the Atlantic (e.g. MYCIN, INTERNIST-1), though Berg presens ample details on how the described systems work.

What does one learn from Berg's work? First, complex tasks like medical diagnosis and treatment are not easily reproduced on a computer. The imprecision and uncertainty of much medical knowledge makes it difficult to create models that work well with the precision of computer algorithms. But medical knowledge is also broad, and no systems have been able to capture it all, hence they have to restrict their domains. Humans with real illnesses, of course, do not restrict their domains! Thus, systems must do what human thinkers are often (though not always) capable of, which is recognizing when they do not know the answer.

A second lesson from Berg's book is that systems cannot be developed without detailed attention to human factors. Furthermore, it must be realized that systems cannot be grafted on to clinical workplaces but must be integrated into them. This has, of course, been long recognized in the human–computer interface (HCI) literature, but not often implemented by medical computing developers, even if they are (as in the case of the creators of the systems described in this book) real practitioners.

Though identifying all of these problems, Berg does not conclude that research into medical expert systems has been a failure, nor does he believe there is no role for these systems. He recognizes that like all computer technology, systems must be matched to the human tasks they best complement. They must also be integrated into the clinical workflow, with allowances for variation in local clinical cultures. The bottom line is that complex systems like these must be tailored to the local situation, and evaluations of them cannot factor out the human surroundings. System-oriented research is, of course, necessary to establish the validity of computer systems, but ultimately all aspects of computer systems, even including the algorithms, must enhance human tasks.

Although focused on decision support systems, this book has implications for information retrieval (IR). In fact, many of the same debates, especially with regards to evaluation, have ben going on in our field. Many writers, including this author (Hersh, 1994), have questioned the excessive reliance on batch-oriented recall-precision studies. They are criticized not because they should not be done at all, but rather because often the conclusions are not justified.

Berg's book is easy to read and a fascinating glimpse into an attempt to apply computers to significant and complex real world problems. He provides an objective, middle-of-the-road assessment of medical expert systems and where future research in their development and evaluation should lead.

REFERENCES

Antman, E., & Lau, J. *et al.* (1992). A comparison of results of meta-analyses of randomized controlled trials and recommendations of clinical experts: treatments for myocardial infarction. *Journal of the American Medical Association, 268*, 240–248.

Bernstein, L., & Barriere, S. *et al.* (1982). Utilization of antibiotics: analysis of appropriateness of use. *Annals of Emergency Medicine, 11*, 21–24.

Hersh, W. (1994). Relevance and retrieval evaluation: perspectives from medicine. *Journal of the American Society for Information Science, 45*, 201–206.

Shortliffe, E. (1987). Computer programs to support clinical decision-making. *Journal of the American Medical Association*, 258, 61–66.

Weiner, J., & Parente, S. *et al.* (1995). Variation in office-based quality—a claims-based profile of care provided to Medicare patients with diabetes. *Journal of the American Medical Association, 273*, 1504–1508.

Oregan Health Sciences University
Portland, Oregan

WILLIAM HERSH