Maastricht University

Rationalizing medical work : decision support techniques and medical practices

Citation for published version (APA):

Berg, M. (1995). *Rationalizing medical work : decision support techniques and medical practices*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg. https://doi.org/10.26481/dis.19951012mb

Document status and date:
Published: 01/01/1995

DOI:
10.26481/dis.19951012mb

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 13 Mar. 2024

Rationalizing Medical Work

Decision Support Techniques and Medical Practices

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus, Prof. Mr. M. J. Cohen,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen
op donderdag 12 oktober 1995 om 16.00 uur

door

Marc Berg

geboren te Leiden

Promotores
(Thesis advisors)

Prof. Dr. G. H. de Vries
Prof. Dr. G. H. Blijham, Universiteit Utrecht

Co-promotor
(Thesis co-advisor)

Dr. A. Mol

Beoordelingscommissie
(Examining board)

Prof. Dr. W. E. Bijker (voorzitter)
Prof. Dr. A. Hasman
Prof. Dr. A. C. Nieuwenhuijzen Kruseman
Prof. Dr. S. L. Star, University of Illinois (USA)
Prof. Dr. G. A. M. Widdershoven

Contents

Preface

This book is about decision support tools (such as protocols, expert systems and clinical decision analysis) and medical practices. It is a study of how medical practices are transformed by these tools and vice versa; of what "rationalizing medical work" does and does not look like. As such, it is the outcome of the coalescing of a sociological interest in what it is that medical personnel do and a fascination for the way (formal) technologies develop and function in concrete work practices.

Although it would be gratifying to locate the drive for these fascinations within oneself, it would also be gratuitous. I have always considered myself lucky to have as supervisors the two persons who have most thoroughly shaped (and sharpened) my theoretical views: Gerard de Vries and Annemarie Mol. As a guardian for intellectual quality and consistency, Gerard is the thesis supervisor *pur sang*. I owe most of the hard times I had in the creation of this book to him, but I would not have liked to have missed his critical insights, sharp analyses - and the didactical floggings. In her own unique way, Annemarie has been indispensable: her in-depth, critical, wide ranging, yet always wholly supportive mode of supervising is everything a PhD student can hope for. I consider myself lucky, again, that we have gradually become "colleagues" and "friends" rather than supervisor and supervised.

Geert Blijham, my third supervisor, was vital in getting access to the field of oncology: both practically and intellectually. He has read and commented on several versions of this manuscript, and helped to ensure that I got the message down in clear terms. The way he could summarize this thesis in two sentences always formed an important counterbalance to the tendency to take one's work too seriously.

Harry Schouten has also been vital to my meanderings through oncology. I owe a great deal to our many discussions, the corrections he made whenever I got the "medical stuff" wrong, and his cheerful enduring of my nosy presence. The doctors, nurses and patients who I got to know during my fieldwork, and the tool builders and/or physicians I interviewed: without them, there would not have been a book.

Of the many people who have read, discussed, commented upon or critiqued earlier portions and versions of this book, there are several that I especially want to mention. Of my colleague PhD students, Antoinette de Bont, Monica Casper, Ruud Hendriks, Jessica Mesman, Susan Newman, Irma van der Ploeg, Stefan Timmermans and Paul Wouters have been indispensable as both friends and critical readers. I have also benefited greatly from the support and helpful comments of Ed Berg, Wiebe Bijker, Geof Bowker, Bruce Buchanan, Adele Clarke, Diana Forsythe, Harry Marks, Randi Markussen, Bernike Pasveer, Leigh Star, and Lucy Suchman. Each in his/her own way has left a definite imprint on me and this manuscript. In

the last phases of this project, Hellen Heuts, Arjen Stoop and Angel Waajen have lent many a helpful hand, for which I am also grateful.

My daily site of work, the department of Health Ethics and Philosophy at the University of Maastricht, has developed into a congenial, supportive and warm place to "come home to" - not in the least due to the many colleagues/ friends who are in a similar phase in their careers and lives. Intellectually, I also owe much to the Maastricht-based research group on Technology and Society (BOTS), chaired by Wiebe Bijker. The Netherlands Graduate School of Science and Technology Studies (formerly LOOWTOK, now the Research School on Science, Technology and Modern Culture), albeit in a constant flux, has provided me with high quality and highly enjoyable workshops and seminars to sharpen my understanding of what the field of Science and Technology Studies was about in the first place.

Saskia van der Lyke, finally, has endured the constant flux that I was in during these past years. Above all, she sharpened my understanding of what life is about - and what more could one ever want.

Introduction

In the oncology ward's nurses' office, Irene is looking for Mr. Field's medical record. Mr. Boottle, one of the patients she is taking care of, is suffering from the same type of throat-cancer as Field. He will be treated with the same type of treatment - and Irene wants the protocol that describes it. She enters the medication room of the ward, where lists of antibiotic-solutions are taped to the wall, and little piles of scribbled chemotherapy recipes occupy the larger part of one of the desks. The record she looks for is not there. Her colleague Matthew is, however, and he accompanies Irene to where he has last seen the record: in the room of the ward's residents. When they cross the hallway, they can hear the somewhat demented Ms. Laury yelling from her room, and a pager beeping in the nurses' office. The usual afternoon sounds.

The messiness of the residents' room is equally familiar. Patient records and chemotherapy prescriptions are scattered about on the tables - there is even a prescription lying in the little basket with sugar and milk bags. They find Field's record, and routinely flip through it to where they know the protocol's scheme is summarized. Scanning the list, Matthew says: "You should change the Mannitol into Lasix, and do the Zofran intravenously on the first day. From then on, you only have to do it intravenously if necessary. And I guess you can just copy the timings they used here. That doesn't really matter much."

***

De Dombal on the making of a list of data items to be used as input for his computer-based decision support tool (dealing with acute abdominal pain complaints):
Some items we've never had money to investigate. That is why the X ray of the abdomen is not part of our system: we've never had resources to find out whether it would make a difference. We've only had money to investigate the historical and physical examination data - and the system now only deals with those.

And when we had made the list, and the physicians really missed something they did not want to miss, we simply put it back in. So that's why [hospital X's] form has room for filling in the temperature, the pulse, and the respiration rate - although in 1920 somebody already said that these data were useless in acute appendicitis. But they want it, so we do it. They're only on the form, however: the computer doesn't use them...

...

And when we began we ignored all kinds of mathematical niceties to get the thing going. There was not much around to hold on to. It was not a question of doing it by the book - we were making the book up as we went along. I wouldn't be trying to justify it. We couldn't go anywhere for help. There was Lee Lusted in the States - but nobody could get to the States in those days. Somebody told us about Lusted's book on Bayes. We read it, and we thought, well, yeah, that sounds reasonable. If somebody would have given us a different tip, we might have gone into a completely different direction.

...

But don't get me wrong: we now have a system which has been running for years. In [hospital X] the system has achieved a false negative operation rate which is much lower than elsewhere. In addition, the system's efficiency is such that acute abdominal pain patients spend about half the number of days in the hospital here compared to elsewhere in the UK! (Interview de Dombal, July 18, 1993)[1]

***

In the last two decades, editorials and articles in medical journals have expressed worries about the current state of medical practice. At the start of an invited series of articles for the *Journal of the American Medical Association*, Eddy argues that medical practice finds itself confronted with a profound challenge:

> The plain fact is that many decisions made by physicians appear to be arbitrary - highly variable, with no obvious explanation. The very disturbing implication is that this arbitrariness represents, for at least some patients, suboptimal or even harmful care. (Eddy 1990b)

The numbers of operations between regions, for example, show variations which are in the order of three- to twentyfold. Physicians vary in what they do and how they reason: for example, when having to decide on the appropriateness of indications for coronary angiography or coronary artery bypass graft, UK physicians judge considerably different than US doctors (Wennberg 1984; Brook et al. 1988).

According to these authors, the troubles are not surprising. The "ingredients needed for accurate decisions are simply missing for many medical practices." Often, there is no evidence available to determine one preferred action of choice. As a consequence, physicians "turn to their own experiences" - which are "notoriously misleading" (Eddy 1990b). Moreover, many authors argue that the problems modern doctors face are often too complex for ordinary humans to grasp comprehensively. Medical decision making is hampered by "one of the most frustrating failings of the human mind ... : its Lilliputian capacity for storing and retrieving important but infrequently used information."[2] Eddy agrees: "even if good evidence were available, it is unrealistic, even unfair, to expect people to be able to sort through it all in their heads" (ibid.).

How to address this challenge? Many authors enthusiastically refer to so-called decision support techniques as solutions: tools which can aid physicians in rationally practicing medicine. Different tools are mentioned: computer-based systems (such as the tool de Dombal was working on), protocols (or "practice policies"), and decision analytic techniques ("clinical decision analysis").[3] These tools, which directly intervene in the medical decision making process, hold promise for an optimally scientific medical practice. Together with the medical practices they seek to rationalize, with the workplaces of medical personnel like Matthew and Irene, they are the main characters of this book.

Schematically, decision support tools can be said to consist of two components, which may or may not be separated: a knowledge base and an inference mechanism. The former contains the "knowledge" needed to solve medical problems within a certain domain. Given a patient's description, the inference mechanism draws upon this knowledge base and generates an advice tailored to this individual case (Figure 1).

In *computer-based decision support systems*, the inference mechanism can be embedded in symbolic, so-called "conditional rules," or in statistical formulas. These tools are said to improve medical decisions by "efficiently placing current experts' knowledge at the doctors' disposal" and by helping physicians to "avoid errors caused by the imperfection of the human

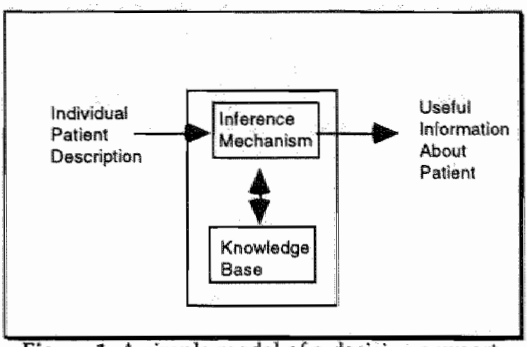


Figure 1. A simple model of a decision support system. From: Reggia and Tuhrim (1985, 7).

intellect" (Potthoff et al. 1988). Through such help, Schwartz argued in 1970, the "computer as an intellectual tool" can fundamentally alter the practice of medicine:

> [It can] help free the physician to concentrate on the tasks that are uniquely human such as the application of bedside skills, the management of the emotional aspects of disease, and the exercise of good judgment in the nonquantifiable areas of clinical care. (1970)

Protocols, subsequently, offer predefined, step-wise and optimal paths through complex or troublesome medical situations. At each step, decision criteria can be built in to determine whether the next step can be taken: a protocol can be seen as a chain of (more or less simple) conditional rules. Well-defined protocols should be able to "reduce inappropriate variation in services, improve the quality of care, and produce better health outcomes" (Field and Lohr 1990). By analyzing "decisions before the fact," protocols prevent the impossibility of having to rationally decide every time again from scratch (Eddy 1990c). These tools, Margolis proudly acclaims, help us to "define, not for others, but for ourselves, the practice of rational medicine" (1983).

Finally, *clinical decision analysis* brings statistical techniques to bear on problems concerning the management of individual patients. The technique should allow the physician to choose the management strategy with the highest expected utility (see further). According to its advocates, it is the most "flexible, practical and yet rigorously logical approach" of all decision making aids. "It is a logical, explicit, reproducible and objective process" that enables comparing alternatives and dealing with uncertainties; you can "examine in exquisite detail the assumptions and data on which the recommendations are based" (Weinstein et al. 1980; Detsky 1987). Its advocates address it as a veritable "new basic science in medicine," which is vital given the proliferation of medical tests and technologies and the troubles physicians have interpreting statistical properties of data.[4]

First and foremost, decision support tools would thus help to "transform the 'art' of medical decision making into a 'science'" (Komaroff 1982). They would make medical practice more rational, more uniform, and more efficient. Secondly, and closely related, the tools would help re-

lieve the physicians' ever increasing intellectual load. They would support the physician where (s)he is weakest. Finally, the tools would achieve all this in a relatively unobtrusive and transparent way: simply by doing better what the physician had been doing all along. Expert systems, their advocates argue, reason like physicians do - they truly "understand" medical practice. Likewise, clinical decision analysis just makes the uncertainty, the risks and benefits inherent in each and every decision explicit.[5] Decision support techniques, it is said, fit medical work because they are structured like medical work itself. The main difference is that the tool's "predetermined, formal, and explicit scheme of logic" allows us to *improve* upon this work. The tools, in other words,

> can help us to demystify the practice of medicine, and to demonstrate that much of what we call the "art" of medicine is really a scientific process, a science which is waiting to be articulated. (Komaroff 1982)

Not everyone is enthusiastic about these tools, however. Several authors challenge both their feasibility and desirability, contesting many of the advocates' claims. Dreyfus' influential *What Computers Can't Do* (reprinted in 1992 as *What Computers Still Can't Do*), takes on a line of argumentation which characterizes much criticism.[6] In his "critique of artificial reason," Dreyfus criticized computer scientists' attempts to achieve 'intelligent' systems - but the criticism is readily extended to all decision tools "using prespecified rules and principles" (Dreyfus and Dreyfus 1986, 185). In a phrase which captures the critics' central tenet best, Dreyfus and Dreyfus state that such tools all "suffer from the impossibility of replacing involved knowing how with detached knowing that" (ibid., 184).

In most areas of human life, Dreyfus argues, rule-based systems cannot achieve an adequate level of expertise. They can function in *formal* or *formalizable* domains, separated from the messy real world, such as tic-tac-toe, mathematical theories, or highly delimited areas such as the evaluation of a few clear-cut laboratory data. In these domains, behavior may be represented as strictly rule-governed. All relevant in- and output is fixed; the mechanical processing of a set of rules can deal with all possible events. Outside these domains, the competence of rule-based systems is limited *in principle*. Here, rules or principles (if existing at all) always contain a *ceteris paribus* condition. "They apply 'everything else being equal', and what 'everything else' and 'equal' means in any specific situation can never be fully spelt out without a regress" (1992, 56-7).

To illustrate this argument, imagine a system which would classify an emergency department patient with acute abdominal pain into one of seven possible diagnoses according to a pre-set list of rules. The system (including the selection of the diagnostic categories) would be based on past experiences with several hundreds of such patients (*cf.* de Dombal's system, Chapter 2). It would then not be unreasonable to state that in general, "everything else being equal," an emergency department patient with acute, sharp pain in his belly would have one of these seven diagnoses. Likewise, "everything else being equal," this patient's symptoms would lead the tool to classify the patient adequately. On the other hand, everything else might *not* be equal whenever events would occur that were not

foreseen when the tool's rules or principles were encoded. If the patient is a pregnant woman, or if the patient looks psychotic, or if a neighbourhood's water supply has been poisoned, the system's inferences will probably be hopelessly off. Trying to program all these possibilities into the tool is impossible, Dreyfus argues: the number of rules stating all the exceptions will soon explode. Moreover, how can we ever begin to try to tell the computer just when the basic rules apply, and when they might not? A computer program could not even begin to fulfil these tasks, according to Dreyfus: the domain of acute abdominal pain is just too unrestrained to tackle. The thing is, he says, that we usually *do not* behave according to any clear-cut rules. As Wittgenstein has argued, we often cannot even begin to make explicit *why* we say something is "equal" or not: we shrug our shoulders and simply say "we do that this way" (1958, §185-242). In the end, the ceteris paribus condition points to the background of activities which is *prior* to all rulelike activity. This "lifeform," which makes rulelike activity possible in the first place, can never completely be made explicit.

The conclusion of this argument, shared by many authors, is that human expert performance cannot be achieved with formal systems. Such tools grope for something which is fundamentally impossible. "The everyday world, it appears, cannot be described by ... any finite number of any kind of statements," Blois states in the *New England Journal of Medicine* (1980). Criticizing computer-based decision tools, he argues that the selection of facts relevant to a given situation is the physician's unique competence. Similarly, critics argue that the complexity of "the real context of clinical decision making" is such that decision analytic techniques will have a hard time achieving "clinical relevance" (Ingelfinger 1975; Cummins 1990). A proper analysis of a decision problem requires *all* possible outcomes, and *all* possible concerns and valuations about these outcomes, to be explicated in advance. Bursztajn et al., however, state that for both doctor and patient it is

> difficult if not impossible ... to include in the decision analysis all the various and sundry considerations that ultimately guided their actions, some of which they were not conscious of, some of which they would not have been comfortable dealing with explicitly and publicly, and some of which they never could have predicted would come up. (1981, 146)

Decision analytic techniques, they state, are but poor representations of the complexities that go into real-time decision making. One cannot separate the decision from its context:

> When decision analysis separates facts from values, one cause from others operating along with it, an individual's welfare from that of the family or community, an illness from a person's life history, one decision from the context of other decisions that define its significance, it turns decision-making into something rather artificial and academic. (ibid., 166-7; *cf.* Dreyfus and Dreyfus 1986, 180-4)

Protocols, finally, have been under similar attack. Such rigid, predetermined schemes are said to threaten the physician's "art": they dehumanise the practice of medicine, and reduce the physician "to a mindless cook."[7] By determining the path of action, the protocol renders the physicians' skills superfluous. Physicians then merely have to be able to understand

directions, and to be willing to do what they are told: they are certainly *not* expected to think for themselves (Ingelfinger 1973). Moreover, such tools open the way for increased and uninformed control by "outsiders":

> There will be those in academia, in public health, in the insurance fields, in health maintenance organisations, and most surely and most terribly in the fields of law and government who will desire and will move to require strict conformity of practice to the presumed ideal. (May 1985)[8]

Such coercion to a perceived "best way," May continues, stifles the life of medicine: it can become a form of "tyrannical domination" against which individual choices of physicians should be protected (ibid.).

All in all, these critics argue, the tools' impoverished, codified versions of physicians' know-how do not do justice to the intricate, highly skillful nature of medical work. The idea to create formal tools which make medical decisions is utterly mistaken. Every attempt to take the practical control of the decision process from the physician's hands is doomed to fail - and is dangerous. Discussing medical expert systems, Lipscombe summarizes this issue as follows:

> the systems lack the flexibility of human interpretative processes, [thus] their use requires situationally specific human support. The current "control" design, however, provides no room for such support. (Lipscombe 1989)[9]

***

The arguments of the advocates seem to be diametrically opposed to those of the critics: whereas the former stress the universal reach of the tool and the smooth fit between tool and practice, the latter stress the tool's strangeness and impossibility. This book takes a different stance: it takes the positions raised by advocates and critics as points of investigation, not as a priori assertions. Drawing on ideas and methodologies from science and technology studies[10], I empirically study what "rationalizing medical practices" implies. The book is about the work - in a broad sense - needed to obtain and maintain a place for decision support tools in medical practices.

To underscore the multidimensionality of these processes, the book switches between three interrelated levels of analysis. First, focusing on discourses in *medical literature,* Chapters 1 and 2 investigate the way views of medical practice changed so that the tools came to be seen as the solution for the perceived problems. Second, Chapters 3 and 4 focus on the work of *system builders* to construct and implement some specific decision support techniques. Drawing mainly on interviews and participant observation, analysis centers on the negotiations these construction processes entail, and on how individual tools and practices are transformed in ways beyond anybody's or anything's grasp. Third, Chapter 5 focuses on *medical personnel* at work with the tools. Following them in their daily environments, I look at how the tools figure in and transform the work of medical personnel.

The book discusses the struggles of decision tool builders to get their tools to function, and the problems and challenges that medical personnel

encounter in their work with the tools. Doing so, it emphasizes the inevitable divergences between the claimed goals of the toolmakers and their end products. It thus touches on issues of standardization, of "universality," and localization; on what medical work consists of, and on what happens when such practices meet the tools discussed here. It points out how these tools are never what they seem - without drawing the conclusion that they are useless and should be discarded. Rather, throughout this book I draw attention to what these tools are and bring about; how medical practices are transformed by the tools and vice versa. In this way, some insight may be acquired into what is gained and what is lost when decision support techniques are put to work.

As a necessary corollary, the book does not contain the grand conclusions that may be found elsewhere. Decision support techniques will not be depicted as necessarily deskilling, for example - nor as always relieving medical personnel's task load. I would argue that one should be sceptical of overall utopian or dystopian accounts (Kling 1991): the worlds we are dealing with, and the changes that we will be witnessing are far too complex to be reducible to such one-word evaluations. This does not mean that this book is not evaluative in any way. On the contrary: it addresses the assumptions of both advocates and critics, and takes a stance vis-à-vis them. It also touches on their normative claims - including recurrent ideas on how decision support tools should be evaluated.

The differences and similarities between the techniques discussed are manifold: several are unraveled and explored in the course of this book. However, they are all *formal tools*: they utilize a process of mechanized inference (a formula, or a set of rules) to convert input to output (*cf.* Figure 1).[11] In this way, they attempt to yield more or less specified and more or less binding advice in the diagnosis/treatment of individual patients. Much more can and will be said about these carriers of "commodified expertise" (*cf.* Abbott 1988, 146-50). This description, however, serves as a way to delineate a class of tools from the much larger group of (social) technologies which can be said to pursue the "rationalization of medical practice": ranging from educational initiatives to the creation of electronic medical records and electronic textbooks. The latter developments, however influential and important, fall outside the scope of this study.

Notes

1. The account is edited somewhat to enhance readability.

2. Haynes et al. (1986), "How to keep up with the medical literature: how to store and retrieve articles worth keeping." *Annals of Internal Medicine*, 105: 978-84, quoted in Wyatt (1991b).

3. Many alternative names have been given to the techniques I discuss: as a group, they have been called "formal methods for decision making," "decision-aids," "intellectual technologies," and so forth. Sometimes the word "algorithms" is used as a general category, but this term is more often used for a type of protocol - although here as well, conceptual variation abounds.

4. See Silverstein (1988) and Balla et al. (1989). The remark on "basic science" is from Politser (1981).

5. On clinical decision analysis, see e.g. Balla et al. (1989); on expert systems see Reggia and Tuhrim (1985) and Shortliffe et al. (1979). For a similar argument on the "natural" place of protocols in medical practice see e.g. Margolis (1983) and Eddy (1990c). Forsythe (1993a; 1993b) and Kaplan (1987) analyse these views with regard to computer-based decision tools.

6. This group of "critics" encompasses a diverse (and not necessarily unanimous) set of authors. All in all, the computer-based tools have had most attention. For a philosophical type of critique (often closely resembling Dreyfus' line of reasoning) see e.g. Winograd and Flores (1986), Searle (1984), and Collins (1990). For an application of Collins' ideas in the medical domain, see Lipscombe (1989) and Hartland (1993b). For an anthropological perspective yielding similar conclusions see e.g. Nyce and Graves III (1990), Graves III and Nyce (1992), and Forsythe (1992; 1993b). Some tool builders have become critics as well: see e.g. Engle (1992) or the cautious editorials of early "advocates" Barnett (1982) and Schwartz (1987) (see further Chapter 3 and 6). Finally, within the new field of *Computer Supported Cooperative Work* similar points are made: see e.g. Button (1993), Greenbaum and Kyng (1991) and Star (1989b).

7. P. Cutler (1979). *Problem solving in clinical medicine.* Baltimore: Williams and Wilkins, p. 53. Cited in Anonymous (1982).

8. May discusses the phenomenon of consensus statements. See Chapter 2.

9. Many computer-based tool builders *say* that they do not want to "supplant" physicians; that they only want to support, to aid physicians in their diagnostic or therapeutic work. However, Lipscombe argues, many systems built by these authors still offer *solutions* to medical problems instead of *advice*; they still grant the *situated control* of the decision making process to the computer. See Chapter 5 for a further discussion of this issue.

10. See for recent overviews and appraisals e.g. Bijker and Law (1992), Latour (1987), Pickering (1992), Clarke and Fujimura (1992), Lynch (1993).

11. For an introductory discussion of the notion of "formal systems" in the context of Artificial Intelligence see Haugeland (1985).

1

The Withering Flower of Our Civilization: Reconceptualizing Postwar Medical Practice

In 1982, Komaroff started an editorial in the *American Journal of Public Health* with the following lines:

> Over the past 30 years, there have been increasing attempts to transform the "art" of medical decision making into a "science," to supplement a spontaneous, informal, and implicit set of judgments with the conclusions of a predetermined, formal, and explicit scheme of logic. The driving force behind this effort has been the perception that clinicians make medical decisions in an idiosyncratic manner, sometimes compromising the quality of care or wasting medical resources.

Komaroff speaks of a clear-cut project driven by an equally clear-cut need: a gradual and unswerving process of turning an art into a science, propelled by an awareness of the poor decision making capabilities of physicians. This rationalization project is achieved through implementing "explicit schemes of logic": decision support techniques.

This depiction, recurring in many other articles and editorials, misses several crucial points. It takes for granted that the problem has simply been around for all those years, waiting for a solution which slowly crystallized in the form of the current decision support techniques. It takes for granted, likewise, that postwar medical practice was an "art" waiting to be demystified; that it was "really a scientific process" waiting to be explicated (ibid.). Through a focus on the medical profession's discussion of these issues in postwar medical editorials and textbooks, this chapter presents a different view. It demonstrates that there was no single, unilinear process in which a previously "unscientific" practice became "scientific." Likewise, there was no single, stable realization of "a problem" which for thirty years guided the search for a solution. The "medical practices," "problems" and "solutions" discussed by editorials and textbooks were conceptualized *differently* within different discourses. What "scientific" medical practice *is* according to medical authors, and, concurrently, what medical practice's problems *are*, took on different forms in different times and contexts.

Focusing on the United States in the period 1945-90, this chapter mainly draws upon editorials from two leading American general medical journals.[1] In addition, some medical textbooks' introductions were traced

through different editions. To get a fuller grasp on the emergence of those discourses in which decision techniques became prominent, more extensive searches were made into their backgrounds.

Using sources such as editorials and textbooks obviously prohibits making claims about the way medical practices did actually change, or were actually structured, or how the discussions described affected the day-to-day work of practitioners. Editorials serve distinctive, rhetorical purposes in the creation and maintenance of "the profession" and of the medical societies which the journals in question represent.[2] In addition, my focus on the postwar literature leads to a relative neglect of earlier attempts to redefine medical practice as a scientific activity.[3] While these considerations are of utmost importance for a more explicitly *historical* approach to the topics discussed, they do not hinder the project set out here: grasping the transformations in postwar conceptualizations of "medical practice" in medical literature.

As used here, the term "discourse" indicates an internally coherent way of depicting and describing medical practice, including its problems, the solutions for these problems, and whether or not (and in what sense) the practice of medicine is a scientific activity. Although the discourses have their origins and moments of widespread acceptance in different periods, there was no clear superseding of one discourse by the other. The different discourses sometimes argued with each other, or blended - or sometimes simply co-existed.[4] My sequential discussion should thus not be understood as reflecting a gradual "enlightenment": the discourses were not steps towards a single view but a set of disparate configurations. Moreover, each of these discourses can be seen to have operated in a circumscribed social and political time and space, indicating the historically contingent nature of their emergence and fluctuation.

I first characterize the discourse on medical practice and its problems which dominated the journals during the early postwar years. Subsequently, I zoom in on some of the divergent conceptualizations of "scientific" medical practice encountered in the material studied and its concurrent, divergent shortcomings.

Early Postwar Medical Practice: Its Nature and Problems

There are men and classes of men that stand above the common herd:
the soldier, the sailor, and the shepherd not unfrequently;
the artist rarely; rarelier still, the clergyman;
the physician almost as a rule.
He is the flower of our civilization,
and when that stage of man is done with,
only to be marveled at in history,
he will be thought to have shared but little
in the defects of the period
and to have most notably exhibited the virtues of the race.
(Robert Louis Stevenson)[5]

The Science and the Art of Medicine

One often encounters passages similar to the epigraph of this section in early postwar editorials of medical journals. The tone is one of pride in the profession's accomplishments in providing medical care. Emerging "from a confusion of superstition and ignorance in its earliest days," editorials emphasize, medical practice has found a firm footing in the sciences of medicine (Anon. 1962). The achievements of these sciences in the first half of the twentieth century have been enormous:

> Never in the recorded medical history of the world have there been so many inspiring discoveries the importance of which has startled at times entire nations ... Some diseases have almost been eradicated; at least they do not constitute serious health hazards. Other diseases are being brought under control at such frequent intervals that many persons have not grasped the enormity of such medical efforts. (Anon. 1950b)

Nevertheless, editors acknowledge that much remains to be done. We may have escaped the age of traditionalism, one editor argues, but our present age is "still stoutly laced with superstition and ignorance." Inspired by the military and medical success of the coordinated research effort during World War II, editors state that "all-over coordinated scientific investigation would accelerate the finding of needed truths that seem just beyond the horizon." With more high-quality research, the hope of seeing "alabaster cities gleam, undimmed by human tears" might become true.[6]

Notwithstanding this jubilant confidence in the benefits that science would bring, medical *practice* is generally not regarded as a science. "The medical profession has three responsibilities," an editorial in the *New England Journal of Medicine* argues in 1952: "to cure the sick, to prevent disease and to advance knowledge." Of these three responsibilities, the editor addresses only the advancement of knowledge as a scientific activity. This separation between medical *practice* and *science* is pervasive in the early postwar years. "Science is nothing more than a method of reasoning equally applicable to the laboratory or the clinic," a distinguished physician states. In mentioning the "clinic," however, he is not talking about medical practice but discussing the need for clinical research. Medical care is only aided by science; it is not a science itself. The practice of medicine consists of *applying* scientific medical knowledge to individual patients

with unique symptoms and complaints. This application is an art requiring "medical ingenuity, experience, skill and individual attention." The relationship of the individual patient with his/her individual physician, inspired with the "spirit of dedication," is its pivotal core. So, investigators have found that physical exertion is better for the cardiac patient than "the time-honored regimen of rest," an editor writes in 1959. But how do they apply this new scientific fact?

> How do physicians judge whether or not a cardiac patient can safely undertake a particular job? The answer is the same as it is to many questions in medicine ... Clinical judgment, common sense, and courage comprise the key to management.[7]

In the 1950s, when editors proudly spoke of "scientific medicine," they were proud of the *availability* of "scientifically" generated medical knowledge and technology.[8] The phrase "scientific medical practice" was seldom used, and when it was it rarely meant anything other than the presence of scientific knowledge. Moreover, when medical practice was addressed in terms of its scientific character, it was often in a critical tone. Medical care has paradoxically been jeopardized by the same advances of medical science that have helped it, authors state. The increase of "scientific paraphernalia" threatens to turn physicians into mindless technicians: the overabundance of recently discovered laboratory tests or "so-called miracle drugs" can lead to a "'pushbutton' type of decerebrate medicine." An editor makes this point poetically:

> For the mechanization of an art, a skill or a culture cannot adequately take the place of the personal sensitivity of its followers, which had made a living thing of it, nor can the interpretation of a physical phenomenon replace the sympathetic understanding of a total human problem. The air brush is not a complete substitute for a bundle of camel's hair, and the drawing of a horizon with a rule cannot convey exactly what the artist sees.[9]

The relationship between the science and practice of postwar medicine was thus fundamentally ambiguous. On the one hand, editors argue that "a generation or two ago the physician had no choice but to rely largely on his skillful exploitation of the art of medicine" (Talbott 1961). The art of medicine, to be successful, requires a science to be applied. On the other hand, medical practice should never let this science impinge too close - lest the ability to artfully apply this science might be lost. Nothing is more important than "the authentic 'feel' of the case - the appreciation of the degree of the patient's pallor, the odor of her breath, the precise color of the blood-tinged sputum, the dryness of the parched tongue, the querulous reaction to her illness and the psychosomatic background for this reaction" (Anon. 1951a).[10]

The Problems of Medical Practice

Early postwar editorials did distinguish many problems. The costs of medical care were rising so much that more and more people had trouble paying for their medical bills. Also, the number of physicians was felt to be

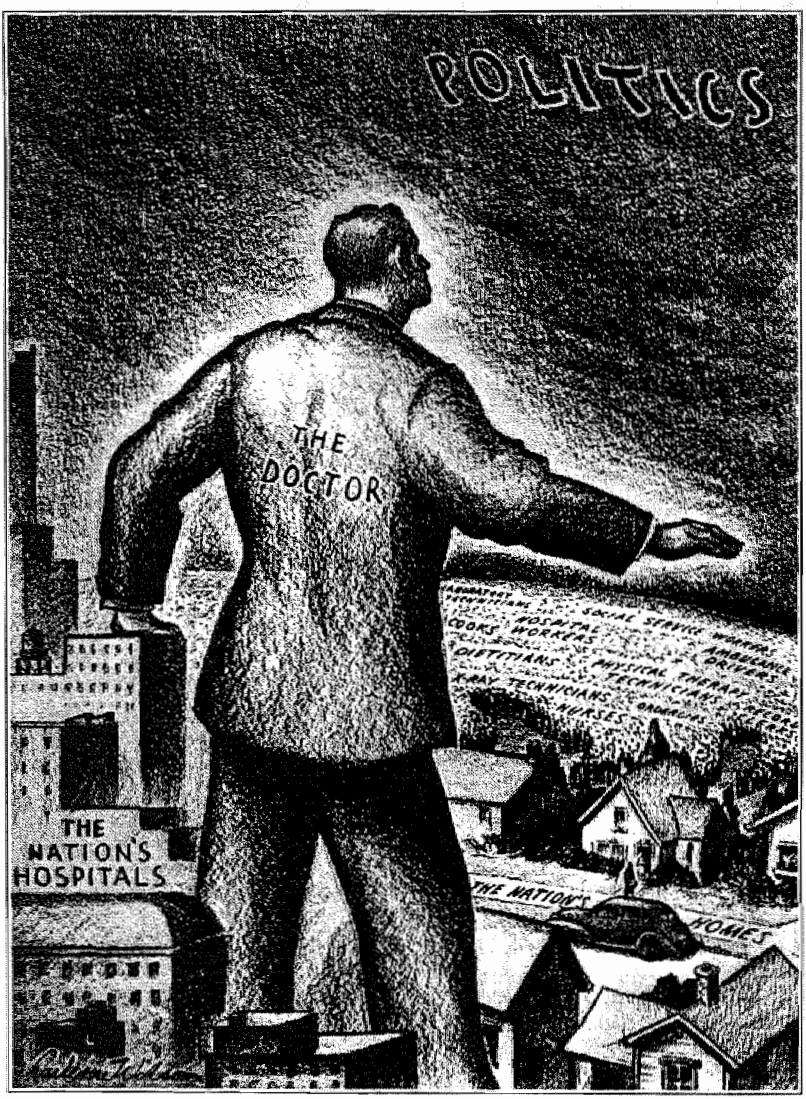


Figure 1. A cartoon showing the proper role of the physician according to an editorial in the *JAMA* (Anon. 1938). The stormy clouds of politics, the growing new health professions, the rising power of hospital administrations: all should be kept in check to improve the physician's protective role of the nation's homes.

inadequate and unevenly spread across the nation. One editor warns in 1948:

> from many sections of the United States complaints have come lately that persons who have called physicians late at night have been unable to secure attendance ... In one western community a fire chief gave to the press information to the effect that he had called twenty-four doctors and had been unable to secure attendance by any of them. (Anon. 1948d)

Finally, when physicians are available, they often make "errors in judgment and errors caused by ignorance or misinterpretation of the patient's condition" - leading, for example, to the performance of "unnecessary operations" (Anon. 1948a). While not willing to underwrite the criticism from "outside" the profession, another editor cautiously confirms that "medical service is far from perfect." "Is there a shortage of doctors? Is 'socialized medicine' the solution for our medical problems? Is there too much specialization? Why is it hard to get a doctor in an emergency? How can the cost of medical care be reduced?" (Anon. 1948c).

What lies at the root of these problems? A similar answer is often given. "In the practice of medicine ..., technic has outrun the present capacity of society to place scientific progress at the service of mankind. At this point while the spirit may be considered as willing, the social and economic flesh presents its weaknesses" (ibid.). These external restraints, editors argue, become more and more salient now that medicine's scientific achievements continue at unprecedented speed. "The success of the medical profession in increasing the life span has resulted in a whole new set of socioeconomic problems." This older age group necessitates "more detailed care at greater cost to the hospital" (Oughterson 1955).

The "errors" and "unnecessary operations" were also commonly explained in this vein. The conditions under which general practitioners in poor urban areas have to work "permit only a superficial and unsatisfactory approach to the problems of diagnosis and therapy," Collings and Clark argue (1953). In addition, "ineffective or obsolete methods of practice established by tradition," like "the fixed idea that some form of treatment must be given," can persist since they are "forced on the profession by the majority of its patients" (Anon. 1950c).[11] It is the social and economic context of medical practice which is to blame.

One of these external obstacles to optimal medical practice is the perceived increase in governmental involvement and regulation. In the eyes of the medical profession, "political control or domination of medical practice" is an evil to be avoided at all costs (Anon. 1945):

> When the Government finances and operates medical services, three things that are essential to good care are inevitably compromised: First of all, the wholly voluntary relation between the doctor and the patient... [Second,] complete privacy, with no intermediary between the patient and the physician. [Third,] maximum incentives for the doctor to do his best work and constantly to improve himself. (Judd 1960)

Against the backdrop of Communist Russia and the atrocities of Nazi-science, the image of the free, autonomous professional was a powerful one (see Figure 1).[12] Increasing governmental control "would seriously handicap, if not abolish, many existing enterprises that are gradually ac-

announcing

A NEW PUBLIC RELATIONS AID

. . . to boost your PR rating

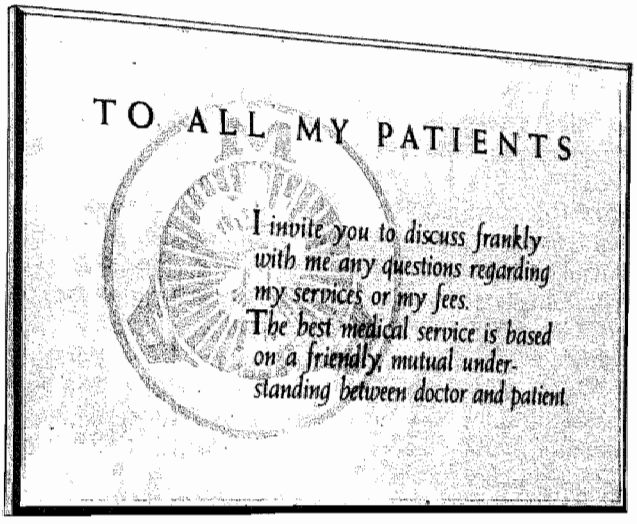


Figure 2. A solution to medical practice's problems in the postwar years. This "attractive office plaque" should be "given a prominent place on physicians' desks or waiting room walls. [It will] show that America's doctors are sincerely interested in providing the best of medical service." (Announcement by the American Medical Association, JAMA, Vol. 147 (1951), No. 11, adv. page 47)

complishing the very things that contribute to the improvement of medical care" (Anon. 1946). It will inevitably lead to "the assembly line type of practice so greatly deplored elsewhere."[13]

The causes of medical practice's problems were thus peripheral to the physician's art. Moreover, editors argued, the problems themselves were marginal as well. "The general health of the population of the United States is constantly improving," an editorial argues in 1950. "No one can deny this without resorting to falsification. Those who claim that a health crisis exists in this country cannot prove it." Another editorial insists that "the great bulk of medical care is handled efficiently and economically." "To the credit of the medical profession," practices where this is not the case are "exceptional." Often, it is argued, the so-called "problems" are in fact merely patients' misconceptions. "The vast majority of patients' grievances against their doctors stem from one thing - misunderstanding ... Patients must be encouraged to talk over with their doctor any questions they might have regarding his services or his fees" (see Figure 2).[14]

In this early postwar discourse, then, medical practice was seen as the artful application of medical science. To address its problems, editorials did not call for a more "scientific" medical practice, as they would some decades later. At most, authors stressed the necessity of a "total view," of "diligence and sound judgment" and of a "careful and judicious" evaluation of the patient (Doyle 1952; Goldwater et al. 1952). Solving "the present-day inadequacies of medical care" required addressing "various socioeconomic factors with which the majority of physicians have had little experience and over which they have little control" (Anon. 1946). The origins of the problems of medical practice were seen to lie in the existence of restraints external to the practice of medicine itself: the fight was against corruption seeping in at the *margins*. Once these restraints would be removed, once governmental control was diminished and more high quality physicians were trained, the practitioners' art would thrive undisturbed.

The one decision support-like instrument referred to in editorials in these years were guidelines for the evaluation of disability (Anon. 1958). These guidelines would help physicians find their way through a labyrinth of complex demands of different institutions. Doctors "are often confronted with a maze of zigs and zags formed by the differing and fragmentary concepts of diverse agencies, such as the Veterans Administration, Social Security Administration ... Now, at last, the zigs and zags are being pieced together into straighter lines." In agreement with the discourse described here, these guidelines are intended to overcome troubles generated by agencies *external* to medical practice. It is again the weakness of the socioeconomic flesh which is at stake. The physician's position as the flower of our civilization remains untouched. A 1959 editorial proudly quotes President Eisenhower's tribute to the medical profession. "We shall one day establish a world community of peace-loving nations," the President says addressing the American Medical Association. "In bringing about this happy result," he continues, "no one can or will do more than the doctors of medicine" (Anon. 1959).

The Multiple Sciences of Medical Practice

In these same years, however, other typifications of medical practice started to emerge. Editorials from the 1950s and 1960s resonated with an awareness that medicine was living through its "Golden Years." The intertwinement of professional power and scientific development seemed to hold the promise for indefinite accomplishment and strength.[15]

Against this background, "scientific" became a positive (albeit as yet rather vague) denotation for medical practice as well. Science "permeates" medical practice, it was argued - the boundaries between the sciences that generate and the physicians who apply medical knowledge should not be drawn so rigidly (*cf.* Anon. 1947a). More and more, authors and editors spoke about "scientific medical practice" in a way which added connotations that went beyond the mere application of scientific knowledge. What these connotations would be, however, was not evident. "Scientific" is not the unequivocal attribute it would seem to be. What "scientific medical

practice" actually meant could be and was answered in more than one way; what specific problems would come into focus differed concurrently.[16] In this section, I describe some subsequent (but not necessarily superseding) discourses on medical practice and its problems. The different depictions of the "scientific" nature of medical practice, it is argued, throw their distinctive shadows - and contain their distinctive solutions.

Situating Science as the Foundation of Practice

First, a discourse can be discerned in which "scientific medicine" denoted a configuration which was essentially as described above: a foundation of scientific medical knowledge, authored by the basic (and, increasingly, the clinical) sciences, skillfully applied in unique patients. Here, however, a set of problems was centralized which was absent in the discourse discussed above. Elaborating themes dating back to the turn of the century,[17] a deficiency *within* medical practice was pointed at as the cause of these problems. Attention was drawn to recurrent, specific problems of communication between physicians and researchers.

For one, editorials stress that due to the rise of clinical and epidemiological research, clinical observations have become "irreplaceable scientific data." Therefore, "scientific precision" is crucial in their gathering and classification. But, it is argued, medical practice lacks such precision: it is characterized by terminological chaos and the "loosely catalogued, cross-referenced and integrated" state of clinical records (1950d; Anon. 1965). Due to "prejudiced, incomplete, or noncomparable reports" clinical research often fails "to answer a single question" - thus hampering "the progress of medicine" (ibid.; Howard 1961). In international studies the situation is even worse. A comparison in 1960 between the "cardiopulmonary semantics" of Britain and the United States reveals "striking ... discrepancies in terminology." For example, the common term "chronic bronchitis" refers to a benign, low grade inflammatory process in the United States, while it indicates a potentially grave and progressive pulmonary disease for "the British." How can international research efforts come to anything when this "Tower of Babel situation" persists? (Meneely et al. 1960).[18]

As a possible solution to these problems, standards are created which set the minimum requirements of medical-records systems (Anon. 1961b). In addition, the new, "scientifically designed and arranged" *Standard Nomenclature of Disease* is praised.[19] It describes "the main characteristic about every disease," with a "clarity that far surpasses words and with unambiguity by means of a specific code number for each disease - a code number in which each digit conveys a factor without confusion." With this tool, "the comparison of clinical experience between institutions becomes sound, is easily encouraged, and is attainable as uniformity and accuracy of clinical expression are prescriptive and customary" (Anon. 1957b).

Besides researchers, "health care providers, planners, government agencies, and epidemiologists" would also benefit from these developments: all these groups depend on "meaningful statistics" (Anon. 1954b; Froom 1975). For all these groups, in addition, it was argued that "precise

definitions are particularly important now that computers offer the promise of significant assistance in data retrieval" (Anon. 1967).[20]

Finally, medical practice would benefit as well. Standardization of terminology ensures that "hospital records are easily understood by all who read them, and physicians can understand what others are talking about when diseases and operations are mentioned" (Anon. 1954b). Also, improved communications would help to fight the diagnostic errors and observer variation which keep being reported (*cf.* Garland 1959). Many treatments are applied without "proper data" to support the effectiveness of the therapy, an editorial argues. The "procedures of diagnosis and treatment performed by the physician should have at least as much validity as is required of a drug company before a new drug may be offered on the market" (Anon. 1963b). Standardizing terminology and records would enhance the ability to apply the results such research would generate.[21]

What kind of "scientific medical practice" do we see here? The authors and editors quoted did not move away much from the image of medical practice described in the previous section. For them, the scientific status of medical practice still derives primarily from the foundation of scientific knowledge it rests upon. What was added is a set of new problems, due to strained communication between researchers and physicians and among physicians themselves. These troubles were seen to be caused by a weakness of medical practice *itself*: the practice of medicine lacks uniform records and terminology. Through standardizing nomenclatures and record-keeping procedures, medical practice could improve both the production of data needed for medical research and the application of the results of this research. Such measures would align individual efforts and increase efficient and precise communication; they would fit medical practice more smoothly into the rapidly expanding postwar research enterprise (*cf.* Marks 1988). In this way, medical practice could become "scientific" (now not a negative term): providing the optimal, artful application of scientific knowledge to the needy patient.

Situating Science in the Structure of Medical Action

In the view of scientific medical practice just described, medical action plays a secondary role. It produces the data upon which much medical science feeds and applies the scientific knowledge thus generated. A different rendering of medical practice emerged in the late 1960s and 1970s, with the often quoted works of Feinstein and Weed. Here, medical practice was not primarily the application of a science located elsewhere: the practice of medicine *itself* was depicted as a scientific activity. "Clinicians do not usually regard ordinary patient care as a type of experiment," Feinstein argues in his classic work, *Clinical Judgment*. Nevertheless, "every aspect of clinical management can be designed, executed, and appraised with intellectual procedures identical to those used in any experimental situation." At the bedside, "exactly the same principles of scientific method" apply as to "any other experiment" (1967, 21-2, 46). Likewise, Weed states in 1968 that "the practice of medicine is a research activity." For these authors, medical action *has the same structure* as scientific action:

> The scientist defines a problem clearly, separates multifarious problems into their individual components, and clarifies their relationships to each other. He records data in a communicative and standard form and ultimately accepts an audit from objective peers by seeking publication in a journal. (Weed 1971, 4)

According to these authors, each and every step occurs in the clinical setting as well. Although the physician's actions in medical work might differ in practical shape, a doctor defines, separates, clarifies, records and audits just as much.[22]

With this depiction of medical work as a "scientific" activity in itself, these authors introduced a general, explicitly normative framework. With this framework, medical practice could now be scrutinized and judged. When, as above, "scientific medical practice" means the optimal usage of scientific knowledge, one can scrutinize medical practice for places where this knowledge is improperly used or not used at all. Standardization is then a way to guarantee optimal flow of information so that the benefits of science reach those who need it. By explicitly labeling the distinctive steps in the clinical process as elements of the scientific method, however, Weed, Feinstein and others made medical practice analyzable in a new and thorough way. The individual steps of "the experiment," the definition of the starting point, the planning of the intervention, the observation of the outcome, could now be discerned *and* judged. It was a *redescription* in terms of a new *normative* framework - with its concurrent new problems and new solutions.

The deficiencies were now not merely "problems of communication," but inadequacies which affect the heart of medical practice. Since medical action is often not recognized as *consisting of* clinical experiments, these authors argue, medical practice largely lacks "the scientific qualities of valid evidence, logical analyses, and demonstrable proofs" (Feinstein 1967, 23). The explosion of diagnostic tests and therapeutic possibilities, combined with the increasing specialization and (thus) cooperation of physicians, has drastically changed the daily work of the clinician. Nowadays, these authors say, medical activities concerning a specific patient's problem are often thoroughly spread out in time and space. At the same time, physicians are wielding more and more powerful therapeutics, with a greater capacity to do harm (ibid., 28; Hurst 1971). It is clear, Feinstein, Weed and others agree, that the need for more standardized terminology and procedures is most urgent. Without it, there is no way that medical practice will ever be able to properly adhere to the rules of the scientific method.

Several interrelated strategies were proposed to bridge the gap between the state of actual practice and its scientific potential. Weed and his followers argued that a basic step towards a solution lay in a new, standardized type of medical file. In the "problem-oriented record," all data, action plans and progress notes of all personnel involved are organized around the problem(s) of the patient. In this way, physicians can *act scientifically*. Through the problem-oriented record, the doctor "is able to organize the problems of each patient in a way that enables him to deal with them systematically" (Weed 1968; Hurst 1971).[23]

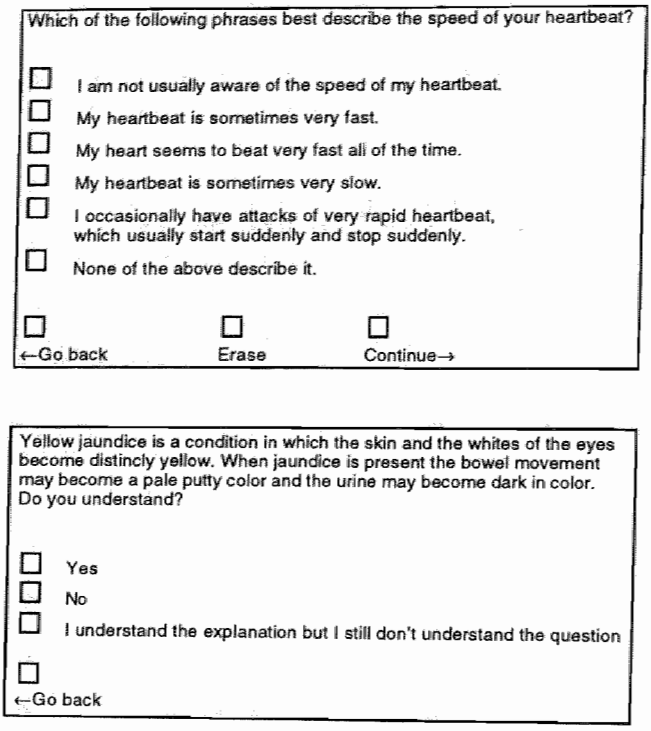


Figure 3. Screens showing one of the over 200 questions of an automated medical history and a "supplemental explanation." From Mayne et al. (1968).

Others were attempting to automate the medical history: a computer would ask a long list of standardized questions to the patient (somewhat individualized through branchings in the list), and would subsequently present the answers to the physician in a summarized form (see Figure 3; *cf.* Barnett 1968; Mayne et al. 1968).

Weed's dream was to put all these intertwining efforts together with an automated version of the problem-oriented record:

> all narrative data presently in the medical record can be structured, and in the future all narrative data may be entered through series of displays, guaranteeing a thoroughness, retrievability, efficiency and economy important to the scientific analysis of a type of datum that has hitherto been handled in a very unrigorous manner. (Weed 1968)[24]

For Feinstein, the basic problem was the unstandardized language physicians use, and the unstandardized procedures through which they acquire their data. Physicians will erroneously write "dyspnea," for example, when they refer "to the fallacious 'shortness of breath' described by a patient whose only symptom is inspiratory chest pain."[25] A main issue was to oppose the usage of the "gaggle of eponyms, synonyms, and terms, which, although declared officially obsolete 30 years ago, still bob up with

amazing persistence" (Anon. 1954a). A thorough uniformity in terminology, and "objective preparation," "delineated precision," and "standardized interpretation" in the acquisition of clinical data is essential to fulfil medical practice's scientific potential (Feinstein 1967).[26]

Given these basics, algorithms (detailed, graphic protocols) could subsequently structure the diagnostic and therapeutic processes themselves. With the conceptual apparatus in place, the protocol could improve the sloppy design and execution of these scientific experiments called patient care. Feinstein explains their role eloquently:

> Until recently, a clinician who wanted to retain the traditional "art" of diagnostic reasoning could not avoid its concomitant scientific aphasia. Having no symbols, no structures, and no tactics with which to demonstrate his patterns of thought, he could not attempt to express his reasoning with any of the traditional oral, written, or graphic patterns of scientific communication. A chemist could use chemical formulas, drawings, and arrows to show the path of an enzymatic transformation; a physicist could use photographs to show the path of an electron's movement; but a clinician had no substance or method that could show the path of a rational sequence. ... [With the recently developed techniques of] algorithms, flow-charts and decision tables ... a clinician can now, at long last, specify the flow of logic in his reasoning. [With these], diagnostic reasoning can begin to achieve the reproducibility and standardization required for science. (1974)

In the influential work of Feinstein, Weed and others we thus find another discourse on medical practice and its problems. "Scientific medical practice" here denotes more than the scientific base of that practice, the existence of scientific knowledge. Medical practice is itself declared a scientific activity, since its actions are structured like scientific action itself. The way the different steps of the scientific method are executed, however, leaves much to be desired. The increasing specialization, and the subsequent increase in actors dealing with one and the same patient, will only aggravate the problem.

As in the previous discourse, these authors argued for standardization measures. Earlier, however, standardization of terminology and records had been required to optimize the application of available scientific and technological knowledge. Here both the goal and the scope of standardization have changed. Standardization is now seen as a fundamental prerequisite to medical practice as a scientific activity in itself: it is a *sine qua non* for the fullblown development of this new science. What is being standardized changes accordingly: creating new, standardized record-keeping procedures and a uniform terminology is only the first step. Protocols can subsequently structure the ongoing process of medical work. With this depiction of the nature of medical practice, a niche is created for a tool which supports the core of the physician's task: the scientific work of diagnosis and treatment.[27]

A Turn to the Mind: New Sciences of Medical Practice

For the two discourses just described, the notion of a "scientific medical practice" was closely intertwined with the idea of standardization. These close links, however, evoked mixed feelings. Standardization, it was felt by

some authors, would also benefit hospital administrations, insurance companies and government agencies - and thus play into the hands of those who were most threatening to the fervently defended professional sovereignty. The notion of such impending regulatory measures conflicted with the image of the autonomous, free professional. When, in 1977, standard protocols for the collection of laboratory data were being offered to practicing physicians, an editor retorted:

> It is a small step from "offer" to "provide." It is another small step from "provide" to "require." And all of this would finally be enforced with the clout of Medicare. Be alert. The first sproutings of the bureaucratic liana may be no farther away than your own clinical laboratory. (Crosby 1977)[28]

While these polemics went on, other notions of what would constitute a scientific medical practice came to the fore. In the 1970s and 1980s, new discourses became prominent in which the scientific character of medical practice became a thoroughly *individualized* notion. Rooted in the booming field of cognitive psychology, these discourses contained an image of medical practice which perfectly fitted the profession's vision of the autonomous physician. Here, medical practice's scientific character was not a feature of a social, collective practice. Shifting away from the alignment of individual efforts, the focus on standardization faded. In these new vocabularies, the scientific status of medical practice was redefined as *a feature of the physician's mind.*

The "cognitive revolution" in psychology had re-established the human mind's active, constitutive role in human behavior - a role which had been out of the spotlight when behaviorism had been dominant.[29] In their well-known *Human Problem Solving* (1972), Newell and Simon developed the argument that human minds can be seen as Information Processing Systems. Humans, these authors state, form mental, symbolic representations of problems that are presented to them. These problems are then solved by logically processing the symbols - that is, by moving around and modifying chunks of information.

In 1978, Elstein and his co-workers brought this approach to the attention of the medical profession. Drawing heavily upon Newell and Simon's work, they called their influential book *Medical Problem Solving*.[30] Like Newell and Simon, Elstein and his colleagues regard problem solving as the symbolic processing of information. Contrary to the former authors, however, they state from the outset that medical problem solving is a "hypothetical-deductive" process. The physician, these authors argue, mentally *tests hypotheses*. Generating possible hypotheses (i.e. diseases or disease-categories) from a few symptoms and signs, the physician subsequently chooses between the hypotheses by testing them against other symptoms and signs obtained from the patient.

Interestingly, Elstein's team thus introduced a terminology strongly affiliated with notions of the "process of scientific reasoning." Problems are solved by the logical processing of symbols, Newell and Simon had argued; "scientific reasoning" is not called upon in their work. In medical education, however, the field of Elstein and his co-workers, the "hypothetico-deductive method" was a familiar notion. Since the mid-

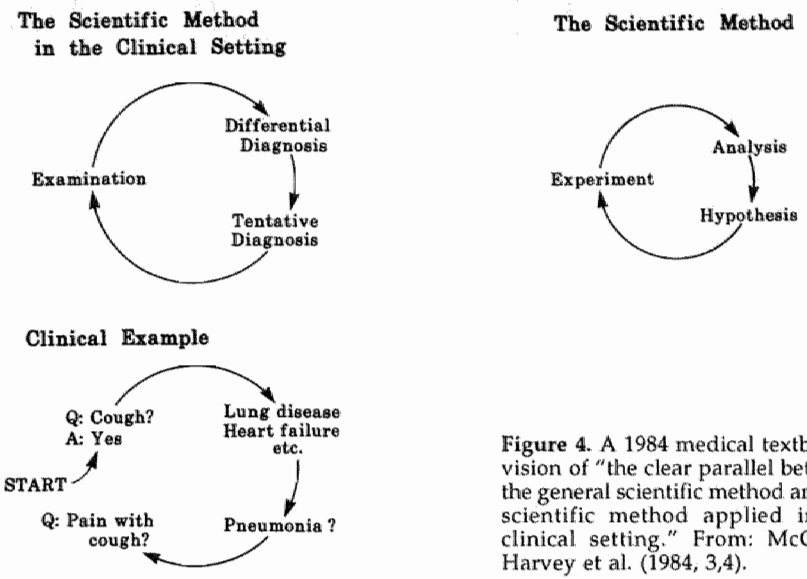


Figure 4. A 1984 medical textbook's vision of "the clear parallel between the general scientific method and the scientific method applied in the clinical setting." From: McGehee Harvey et al. (1984, 3,4).

1950s, it had been seen as a useful technique for teaching students clinical reasoning.[31] In the process of incorporating Newell and Simon's work, Elstein's group turned what used to be a teaching technique into an explicit and detailed *description* of what medical problem solving is.

In this way, *medical action was equated with scientific action, which, in its turn, came to denote a specific thinking process.* "Medical problem solving is a progressive building process of hypothesis generation, testing and verification - devoid of intuitive, inspirational leaps" (Kassirer and Kopelman 1988).[32] It is not hard to see how the attractiveness of this conception lay in its perfect fit with both the notion of the autonomous physician *and* the prevalent scientific rhetoric addressing medicine as a whole (see Figure 4). Compared to the early postwar discourses, the opposition between "practice" and "science" has dissolved completely - but both have changed in the process. Medical practice, here, is the re-enactment of "the incremental process ... [of] scientific discovery" (ibid.), and the scientific character of medical practice has become a mental category. "Training the mind" becomes crucial; "just as the researcher draws conclusions from experimental data, the practicing physician must base his decisions on the analysis of the medical history, physical findings, and laboratory tests" (Greenberg 1978).

The notion of the physician as a processor of symbols, however, is only one variant of the cognitivist discourses on medical practice. Another main variant also locates the scientific status of medical practice in the mind. Like the information-processing approach, it turns medical practice

into a science while concurrently underwriting the profession's image of the individual, autonomous physician. Here, however, medical diagnosis is seen as an exercise in *statistics*.[33] Medical problem solving, authors argue, necessarily involves calculation - whether explicitly or not. "Since 'chance' or 'probabilities' enter into 'medical knowledge', then chance, or probabilities, enter into the diagnosis itself" Ledley and Lusted had argued in a seminal article in *Science* (1959):

> The reasoning foundations of medical diagnosis and treatment can be most precisely investigated and described in terms of certain mathematical techniques. [...These] are inherent in any medical diagnostic procedure, even when the diagnostician utilizes them subconsciously, or on an "intuitive" level. (ibid.; *cf.* McDermott 1971)

Often, notions from economics are invoked as well. These authors thus link up to the spread of economists' terminology into the editorials of the leading American medical journals in this period: notions like "cost-effectiveness" and "marginal utility" become more and more in vogue.[34] According to these authors,

> the essence of clinical judgment resides in the ability of the physician to weigh the advantages and disadvantages of a diagnostic or therapeutic procedure and to choose a course of action for a particular patient based on estimates of such costs and benefits. (Schwartz et al. 1973)[35]

Here, the (potential) scientific status of medical action is found in its exact, quantitative, calculable nature.

The two approaches often fight for their hegemony of the terrain of medical judgment. Authors within the statistical approach often accuse the symbolic information-processing theorists of remaining too vague. Talk of "chunks of information" lacks quantitative precision:

> Only statistics seem able to handle the high degree of variability and the complex pattern of correlated data. With so many variables, practically nothing comes clear and distinct; there are thousands of combinations possible. So calculating the odds of this and that, given such and such, seems the only way to have a reasonable basis for decisions. (Clouser 1985)

In complex situations, the mind can only reach a solution with statistical techniques - whether it does so consciously or not. "The need for accurate probability assessments cannot ... be avoided" (Borak and Veilleux 1982).

In their turn, many authors within the information-processing approach deny that the physician is an "intuitive statistician" or that physicians "grow decision trees in their heads." These are ridiculous assumptions, they argue. In their view, statistical models should not be understood as a *description* of physicians' problem solving or decision making. All they do is "predict" or "black-box like simulate" this process, which, at heart, consists of non-quantitative, symbolic reasoning. "Mathematical psychologists" are wrong to assume "that the kinds of mathematics that work in physics and chemistry will fit the requirements of psychology," they counter. "It is assumed that quantification is important for representing psychological phenomena. It is quite possible, however, to get mathematical precision without quantification."[36]

Within this cognitivist reconfiguration of the shape of medical practice, we see two competing discourses. One approach views the physician as mentally manipulating non-quantitative symbols, according to complex processing rules. The physician follows the hypothetico-deductive method in solving a patient's problem. Alternatively, the other approach argues that physicians (intuitively) calculate their way to a decision. Here, the scientific status of medical practice resides in its exact, mathematical character.

Although the differences between these discourses are substantial (I come back to these later), they also have much in common. With the approaches described here, the "scientific nature" of medical practice has become a *mental* category. What makes medical practice a science is not the fact that the shape of medical work follows positivist rules of scientific method, as in the discourse exemplified by Feinstein and Weed. Here, medical practice is a science since the physician's individual thought processes (consciously or not) themselves exemplify scientific reasoning. What medical work is, then, is again reshaped. Medical work is not seen as a social activity, but as an individual, cognitive process.

We have come back to the notion of the individual, autonomous professional we saw in the early postwar discourse - but the role of this individual has changed radically. No longer is physicians' work depicted as the authentic art of a Man of Character. The individual's merit has shifted from "character" to "method": since the autonomous physician mentally follows the rules of science, (s)he has become the quintessence of the *science* called "medical practice."[37]

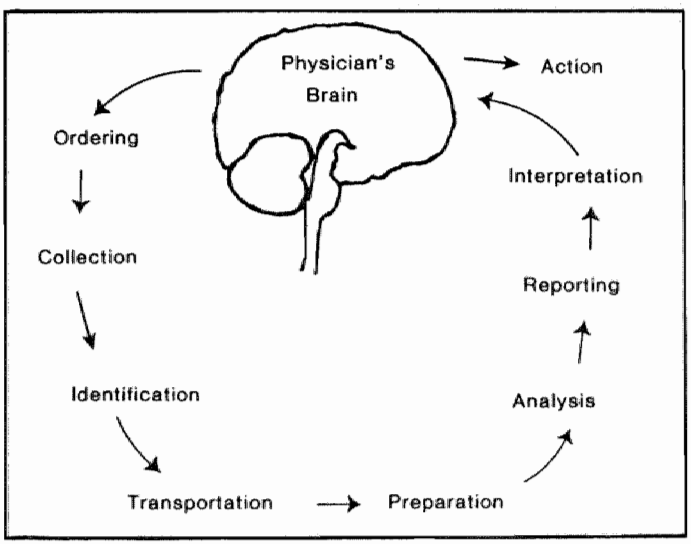


Figure 5. The physician's brain as the core of medical action. From Lundberg (1981).

> In the past few days, I have been feeling as if my calvarium and that of my fellow clinicians has been raised a little. A lot of people have been peering in to the interstices of the doctor's not always efficient cerebral cortex. This collaborative study of the interesting routes we take in making a clinical diagnosis promises to be fruitful, a trifle revealing and, I hope, not too embarrassing.
> (Pellegrino 1964)

In the cognitivist discourses, medical practice is again described as a scientific activity - albeit in a new way, generating a new picture of the nature of medical practice. Again, the models used to describe medical problem solving function at the same time as a normative framework. New problems are highlighted, and new deficiencies spotted. Actual physicians' behavior is compared to and judged by the model, and, subsequently, the model explains the flaws found (*cf.* Gigerenzer and Murray 1987, 162). Seen from this angle, these new discourses on medical practice turn out to be something of a Trojan Horse. They open up medical practice to a new type of scrutiny, undermining the stamp of scientificness they had just established. The physician's mind, the new locus of the scientific character of medical practice, appears to be fundamentally deficient - whether seen from the statistical or the information-processing perspective.

It is noted, for example, that physicians are sloppy in their adherence to the scientific process. It is very difficult for a physician to "avoid wishful thinking or introduction of fantasy," since "we are inclined to form opinion before we have all the facts" (Danilevicius 1975). "Typically, physicians form tentative hypotheses (diagnoses) from a small amount of data. As new information is acquired, they sometimes ignore new data that contradict hypotheses currently remembered" (Politser 1981).[38]

Authors within the statistical discourse pass a similar judgment. Physicians, they argue, "fail to meet commonly accepted and well justified standards for probability interpretation and revision" (Hershey and Baron 1987). Often, it is argued, they

> equate the probability of a test result, given illness, with the probability of illness given that result. Such misinterpretations are extremely worrisome when it is considered that they may be used as the 'scientific' basis of some major clinical decisions. (Politser 1981)[39]

In a sweeping indictment, Eddy agrees: it is clear that "physicians do not manage uncertainty very well, and that many physicians make major errors in probabilistic reasoning, and that these errors threaten the quality of medical care" (1982; 1990b).

It is thus no surprise that physicians can be shown to behave thoroughly idiosyncratic, and to treat similar types of patients in astonishingly varying ways. "When people do what comes naturally, they are likely to be as inconsistent as can be" (Lusted 1968, 149).[40] Whether seen from the statistical or the information-processing standpoint, physicians are only too human - and humans are "limited in their capacity to think, and biased by cognitive processes that interfere with rational decision making" (Ebbesen and Konecni 1980). Cognitive theories of mental processes can explain why these flaws must occur. Physicians just lack "the ability to

manipulate all these data at once" - the human mind can only hold so many chunks in its short term memory (Leaper et al. 1972). "There are limits to man's capabilities as an information processor," McDonald explains. These "assure the occurrence of random errors in his activities" due, for example, to "sensory overload" (1976).

A redrawing of what medical practice is thus involves a concurrent redescription of what its problems are. Medical practice's "scientific" nature is redescribed and physicians are subsequently judged *in terms of* that description. Physicians are seen to "intuitively do statistics," or generate and test hypotheses - and concurrently it is found that they do so suboptimally (Clouser 1985, 42). The description of medical action as a scientific activity functions as an explicit yardstick with which it becomes possible to evaluate (and criticize) what physicians do.

With these new yardsticks in place, new solutions are also called for. Standardizing clinical terminology or introducing problem-oriented records is no longer sufficient. Now that medical work is an activity of individualized minds, the problems of medical practice and their solutions are transformed accordingly. No longer are the lack of awareness of recent scientific developments or the absence of standardized vocabulary or procedures central matters of concern. These issues are now secondary. For what do they really matter when the physician is just not *able* to make rational decisions?[41] What help is a standardized vocabulary when the complex environment physicians have to deal with simply surpasses their cognitive powers? Different solutions are necessary. "We are trying to solve in our head problems that far exceed the capacity of the unaided human mind," Eddy argues in a series of columns in the *Journal of the American Medical Association*. "There are tools ... to help us ... improve the capacity of physicians to make better decisions" (1990b).

So, computer-based decision tools are applauded as the solution to these problems. The physicians' technical tasks, like diagnosis, determining treatment, and prognosis, authors argue, can be supported by such tools. "The main reason computers are able to conduct these tasks better than physicians is because the ideal performance of these tasks requires three critical features - memory, objectivity, and probability - qualities which machines display far better than humans" (Maxmen 1987).[42]

Likewise, decision analytic techniques are called for. These can help the physician since they explicitly take into account those matters which are "implicit in every medical decision": "the likelihood of the outcomes of actions, the risks and benefits associated with these outcomes, and value judgments on how the patient's interests are best to be served" (Balla et al. 1989; *cf.* Silverstein 1988). Since the combination and adequate integration of all this information is the physician's weak spot, the power of clinical decision analysis to find the optimal path of action makes it the perfect tool for the job.

Finally, Eddy and others make a case for the protocol (*cf.* Komaroff 1982). In their usage, the protocol fulfils a different role than in the discourse exemplified by the works of Feinstein and Weed. There, the protocol served as a tool to standardize activities over time and place so that the (cooperative) task of caring for patients can become more efficient - more "scientific." Here, however, the protocol is portrayed as the ideal and natu-

ral tool for compensating the physician's cognitive incapacities. As Eddy eloquently argues, "they are an integral part of our individual and collective psychology." Protocols prevent "mental paralysis or chaos": they

> present a powerful solution to the complexity of medical decisions. They free practitioners from the burden of having to estimate and weigh the pros and cons of each decision. ... [They] provide an intellectual vehicle through which the profession can distill the lessons of research and clinical experiences and pool the knowledge and preferences of many people into conclusions about appropriate practices. (1990c)

In keeping with the cognitivist discourses, the tool is here called upon to support the individual physician's thinking process.

New Practices, New Problems, New Solutions

The 1982 editorial by Komaroff, quoted above, said that the "increasing attempts to transform the 'art' of medical decision making into a 'science'" were driven by the "perception that clinicians make medical decisions in an idiosyncratic manner, sometimes compromising the quality of care or wasting medical resources." As we have seen, however, different discourses contain different configurations of what medical practice is, what "scientific" means, what the central problems are and how these have to be dealt with. There has been no gradual transformation of a practice into a science; rather, we have seen a set of discontinuous images in which the notions of "science" and "medical practice" *themselves* changed shape.

In a similar fashion, Komaroff's remark that this process was driven by a perception of inept decision making cannot be maintained. In the early postwar discourse, I argued, a frontal attack such as Komaroff's on the noble carrier of medicine's art was unheard of. The problems of medical practice were seen to be marginal and to be caused by socioeconomic imbalances which were not the physician's responsibility. A second discourse on medical practice, while framing the relation between the science and the practice of medicine much as the early postwar discourse, pointed at "interface" problems between science and practice. Sloppy record-keeping and terminological inconsistencies obstructed the progress and utilization of clinical research.

That the work of physicians' themselves was seen as "inept," however, that medical practice was seen to be internally deficient, only came with the discourses that described medical action itself as a science. Only through a redescription in terms of an explicit, normative framework did the practice itself open up for scrutiny. For Feinstein, Weed and others, the performance of this "research activity" called medical practice was flawed since physicians often do not adhere to the rules of the scientific method. Since medical record-keeping and/or the medical vocabulary are so disorganized, these authors argue, there is no way physicians *can* precisely plan, execute and observe an intervention. Without these basic prerequisites, the scientific method has no leg to stand on.

It is not until this discourse, then, that tools to "support" medical action are spoken of. Only inside this discourse does a clear-cut niche for the "algorithm" or "protocol" come into being. Now that the distinctive com-

ponents of the scientific task of diagnosis and treatment are laid bare, and the sloppiness of the actual execution of these components is revealed, the protocol becomes the natural route to an adequate performance of the science of medical practice.

Finally, Komaroff's observations that "medical decision making" consists of a "spontaneous, informal, and implicit set of judgments," and that "clinicians make medical decisions in an idiosyncratic manner," may be seen as typical of the way medical practice is rephrased in the cognitivist discourses discussed last. In these discourses, medical practice *is* managing mathematical uncertainty, knowing your prior probabilities, or testing hypotheses. Medical practice has become an enterprise of reasoning or calculating minds. In the same vein, the major flaws are the limitations of individual cognitive abilities. The new yardsticks applied here, the models of symbolic reasoning or mathematical calculation, reveal a veritable "non-Freudian cognitive psychopathology."[43] We have come a long way from the postwar focus on the external, social causes of the problems of medical practice. With the cognitive redefinition of medical practice, the latter had become incapacitated at its very core: the physician's mind.[44]

As said above, however, the different discourses have not simply replaced each other: all depictions still coexist. Medical practice, apparently, has many faces - and so does its scientific nature. These often mingle, merge, or appear seemingly peacefully side by side. In an introductory chapter to a 1979 textbook, for example, Feinstein criticizes the view that therapeutic action is but an art. Denying that the "purely clinical work of patient care does not present any scientific challenges," he restates his views as portrayed above (1979). The next chapter of this textbook, however, discusses the relation between the art and the science of medicine - and redraws their relations as the early postwar editorials did. The science of medical practice, this author argues, resides in the foundation it provides for the "art of medicine - the skillful application of medical knowledge in the optimal care of the patient" (Wyngaarden 1979).

Notwithstanding these mixtures, however, the cognitivist discourses on medical practice have become more and more prevalent in medical textbooks and editorials. In its 1988 edition, the textbook just quoted contains several introductory chapters which stress that medical practice is a scientific activity. "The scientific basis of medical practice is well established, but only recently has the time-honored *art of medicine* come under scientific scrutiny," one of them argues. Moreover, "science" is a matter of "intellectual abilities": "collecting information and synthesizing it into integrated concepts compatible with known diseases" (Morgan Jr. 1988, italics in the original).

With this growing prevalence of cognitivist discourses, the flawed physician's mind is increasingly seen as a central origin of a broad array of problems of medical practice. More and more, the escalating costs of medical care, the public's dissatisfaction with medical practice and the suboptimal quality of care are being transformed into problems which result from the individual physician's mental incapacities.

This transformation affects both the nature of the problem and its explanation. When Komaroff says that the "idiosyncratic manner" of physicians' decision making is "sometimes compromising the quality of care or

wasting medical resources," he is not merely attributing a new cause to an already existing problem. In being drawn into the cognitivist discourse, what "quality of care" consists of is transformed as well: it increasingly becomes a measure of the accuracy of individual decision making. Increasing the quality of care, then, becomes equivalent to "optimizing human cognitive functioning" (Potthoff et al. 1988; *cf.* Eddy 1990b). Likewise, the problem of the increasing costs of medical care now becomes a problem of "wastage of resources" - due to suboptimal decision making of the physician (Eisenberg 1986, 9). Last but not least, the problems of diagnostic error and "unnecessary surgery" were in the early postwar discourse often ascribed to unfavourable circumstances in which physicians were forced to work. Now, however, they are attributed to the occurrence of "pathology" in the physician's decision making.

In line with these changes, decision support techniques are more and more called upon to aid the physicians' floundering mind; not to amend the structure of medical action. It is to these tools, and the differences between them, that I now turn.

Notes

1. These journals are the *New England Journal of Medicine* - henceforth *NEJM* - and the *Journal of the American Medical Association* - henceforth *JAMA*. The focus was not exclusively on literature from the United States, but the journals scanned were American. The details of the developments, therefore, should probably be seen as specific American events. However, the larger outline of the picture sketched here is one which can be seen in the Dutch and English literature as well (for a discussion of the Dutch literature, see Stoop and Berg 1994). American developments were somewhat later often taken over in these countries.

2. Like the notion of "scientific medical practice," "the medical profession" is of course no unitary whole. For the purposes of this chapter, however, I feel that this abstraction is warranted (*cf.* Abbott 1988).

3. As Harry Marks pointed out (personal communication), "the intellectual and social programs of the 1970s might be seen to represent a fulfilment of the (failed) intellectual program of the Progressive era reformers in American medicine."

4. For the notion of "co-existence" see Mol and Berg (1994). The discourses I describe should be seen as ideal-types: they are distilled from many different texts. Some articles sometimes form subtle mixes; in others the typical elements of a discourse can only be partially found. Within some discourses, moreover, subsequent subdivisions can be made - as will be seen in the subsequent chapter.

5. Quoted without reference in Koontz (1959) and Talbott (1961).

6. To avoid cluttering the text with references, I will when pertinent give the references (per paragraph) in a footnote. The phrase "escape from traditionalism" is from Anon. (1956a), "laced with superstition" is from Anon. (1962), "cooordinated scientific investigation" is from Anon. (1948b), and the "alabaster cities" were mentioned by Anon. (1956c). On the (effects of the) success of medical research in World War II, see Rothman (1991) and Marks (1988).

7. The 1952 editorial is Anon. (1952b), the distinguished physician is Rutstein (1962), the description of "medical ingenuity" is from Neal (1951), and the "spirit of dedication" is again taken from Rutstein (ibid.). The last quotation is from Goldwater (1959).

8. The promise of Science in postwar medicine should not be understood as resulting primarily from technically derived advances in medical care. The importance of the image of "science" for the strengthening of the professional identity and autonomy cannot be un-

derestimated. In fact, this scientific promise was already shaping the medical profession's identity when it was as yet very unclear what a "science of medicine" actually was or could be - let alone that it had resulted in major clinical applications. See for this latter point Rosenberg (1987, 150), and Vogel and Rosenberg (1979). Both books deal with 19th and early 20th century American medicine. In her thesis, Pasveer focuses on Britain and the Netherlands (1992). For the situation in postwar medicine, see Starr (1982). For a general review of these issues, see Warner (1985).

9. The phrase "scientific paraphernalia" is from Beecher (1953), "miracle drugs" is from Neal (1951), and the "pushbutton" analogy was coined by Alvarez (1953). The last quotation is from Anon. (1952a).

10. The "Art - Science" distinction has always had its rhetorical usage, of which the distribution of responsibilities and capacities stands most prominent. To describe medical practice as an art, for example, emphasizes the importance of experience, of skillfull mastery. This effectively ensures that e.g. the teaching of medicine remains under the direct control of the members of the medical profession. Likewise, the depiction of medicine as a science had its rhetorical function - *cf.* note 8; see also Abbott (1988). Anderson's fascinating story of the attempts to introduce computer diagnosis into the wards of an Australian hospital is a case in point (1992). It would be erroneous, however, to reduce the usage of the terms "Art" and "Science" to this rhetorical function alone. That would, for example, bypass the fact that the *meaning* of these terms changes thoroughly over time. The changing meaning of "science" with regard to medical practice is the topic here. But it would, for example, be interesting to follow up on the close resemblance between the notion of "the Art of medical practice" in the 1950s and the principle of specificity Warner discusses as typical for medical practice up to the 19th century (1986).

11. Also, the lure of financial gain sometimes may corrupt the physician. "Fee-for-service plans places a premium on keeping the patient sick rather than well" Oughterson argues (1955).

12. See Starr (1982, 338-41) and Rothman (1991, 51-69).

13. Anon. (1952c). This editor refers to the British National Health Service, which was a frightening example for the American medical profession. "By being socialized, an art may become a trade ...; by socialization, a calling may become a business" (Lister 1957).

14. The 1950 editorial is Anon. (1950a), the "great bulk" is from Anon. (1961b), and the editor "crediting" the medical profession is Anon. (1947b). On the misunderstanding, see Anon. (1951b). This misunderstanding, it is said, is frequently due to the many lay reports on medicine's problems showing "a regrettable failure to see both sides of the issue" (Anon. 1948c). So was the case with the fire chief mentioned above, of whom the press had said that he had called twenty-four doctors and found none willing to come. After "a subsequent investigation," he appeared to have "actually talked only to two osteopaths" (Anon. 1948d).

15. Starr (1982, 290-378), Rothman (1991, 51-69). *Cf.* note 8.

16. I do not want to imply that at the level of medical knowledge the meaning of the word "scientific" is obvious. But here the term "scientific" figured in an existing "language game" (*cf.* note 8), which was not transported unequivocally into the realm of medical practice. It is the creation of this *new* language game which concerns me here.

17. See Stevens (1989); see also note 19.

18. See Willems (1993) for an elaboration of these different usages.

19. "Standardization" was not a thing new to medical practice. Calls for "exactitude, clarity and precision" in medical writing can also be found in editorials from the first half of this century, and the development of standardized, pre-packaged drugs was applauded as a scientific achievement in the 1920s (e.g. Anon. 1922; Anon. 1926). See also Stevens' history of 20th century American hospitals (1989) on the standardization of medical education, of hospitals, and so forth; and see Reverby (1981) on early 20th century attempts to standardize record-keeping in order to optimize medicine's scientific efficiency. Also, international efforts to standardize diagnostic classifications go back at least to the 19th century (*cf.* Bowker and Star 1994). The explicit, early 20th century calls for standardization of medical practices, however, disappeared in the decades that followed. With the increasing resentment against and fears for government and hospital administration involvement with physicians' work, the "standardization-practices" of

the American Medical and the American Hospital Association were largely limited to setting standards on residency training, on the organization of the hospital medical staff - but not, for example, on medical procedures. The renewed, explicit attention to standardizing (aspects of) medical work, thus, were new developments (*cf.* Stevens 1989).

20. For a more recent, similar call see Jencks (1992). The computer was both a reason for standardization and a means to standardize - see the attempts, starting in the early 1960s, to devise an automated medical record (Anon. 1963a); on that this electronic record has proved to be more perplexing than expected see e.g. (Korpman and Lincoln 1988). See Kaplan (1995) for a brief history of medical computing.

21. The perception that "the mass of medical writing is expanding at an alarming rate" led to calls for standardization as well. A more standardized "classification of knowledge," Rutstein argued (1961), is a prerequisite for the physician to "maintain contact with medical knowledge." On the "information crisis" as a phenomenon of American science in the 1950s and 1960s, see Wouters (1992).

22. This was not entirely new: in their attempts to "sell" randomized clinical trials to the medical professions, proponents of *these* rationalizing technologies often argued that physicians should allow such controlled experiments since physicians themselves, with each new treatment, were always already experimenting (Marks forthcoming, MS 37-8); see note 3. Analogies such as these had been made much earlier as well. Flexner, for example, already stated that "the progress of science and the scientific or intelligent practice of medicine employ ... exactly the same technique" (1910, 55). And see also Bishop (1900).

23. The problem-oriented record developed by Weed was much discussed and experimented with through the 1970s. In the 1980s, these discussions quieted down somewhat, at least partly because of the developments described below. *Cf.* Aring (1970) and Neellon (1974).

24. To attempt to achieve this, Weed later developed the *Problem Knowledge Computer* (*cf.* Weed 1971) . On this project (and its small impact) see Weaver (1991).

25. Feinstein used this example in a critique of Weed's plans (1973c). Dyspnea, properly used, indicates a shortness of breath of pulmonological origin.

26. *Cf.* Bleich (1971), LoGerfo (1977) and Feinstein (1987b).

27. In this discourse, the "art" of medicine is transformed as well. When medical practice itself is addressed as a scientific activity, the art of medicine is transformed into a supplementary skill of adequately handling the personal aspects of care. This changing position is reflected in the fact that it was being addressed in a different, less solemn way. "The 'art of medicine' has been for long - perhaps too long - a rich source of 'corn' for the gristmills of medical orations. Yet it never has been clearly defined," Vaisrub argued in 1971. "Perhaps," he continues, "it should be sought in the poetry of physicians." This relocation effectively puts the art in a marginal position - as an activity which only reflects upon medical practice.

28. *Cf.* Ingelfinger (1973). The usages of the term "autonomy" has just as many aspects as the usages of the terms Art and Science (see note 8). It is as much a rhetorical device for defending acquired professional privileges as a true concern of physicians regarding their service to their patients. For a detailed and powerful account of the medical profession's struggle for power and autonomy in the late 19th and 20th century, see Starr (1982); see also Abbott (1989).

29. On the cognitive revolution in psychology, see e.g. Baars (1986) and Boon (1982).

30. See for the sources of the title: Elstein, Shulman and Sprafka (1990).

31. Groen and Patel (1985). The problem of the increase of scientific data (see note 21) affected medical education as well. In the 1960s and 1970s, educational reformers pleaded for an educational approach which was less oriented to the amassing of facts. Medical students should learn "proper discipline in approaching medical problems." What was needed was a new educational philosophy, emphasizing the "primacy of process over content" (Elstein et al. 1978, 1-2). Weed was one of the main spokesmen for these reformers. See Lave (1988) on the fundamental link between cognitive psychology and a positivist epistemology.

32. See also e.g. Wulff (1981), Smith (1988) and the American Board of Internal Medicine (1979).

33. This is itself a simplification: different versions can be discerned within this variant as well, some focusing mainly on the physician as a statistical diagnostician, some focusing mainly on the maximization of utility. When looked at in more detail, what now appears as "one discourse" opens up in several, different approaches. Conversely, if I would have strictly remained at the level of medical journal's editorials, I would probably not have differentiated between the two discourses described in this subsection (see also Chapter 2). All quotes used here were explicitly descriptive.

34. Physicians are told to learn more of this "dismal science": "patients deserve it, payers seek it, society needs it, and common sense dictates it" (Anon. 1968; Riesenberg 1989).

35. *Cf.* Lusted (1971; 1975), Silverstein (1988) and Eddy (1990a).

36. The remark on decision trees is from Moskowitz, Kuipers and Kassirer(1988), "mathematical precision" is from Baars (1986, 182). The other quotes can be found in Elstein, Schulman and Sprafka (1978, 11, 32, 40-5, 98). Still other authors mix the two discourses by saying, for example, that the physician may fit each description, depending on the nature of the problem at hand (*cf.* Szolovits and Pauker 1978).

37. The shift in cultural legitimation for the profession's jurisdictional claims from "character" to more technically accounted means is described by Abbott (1988, 190-2).

Of course, instances can be found where authors talked about "scientific reasoning," or even "hypothetic-deductive reasoning" long before the cognitivist discourse came into being (see note 22). Notwithstanding these isolated instances, however, it is only within this discourse that the notion of medical scientific action as a *mental* process becomes paramount.

38. *Cf.* Elstein, Schulman and Sprafka (1978, 252-72) and Knaus (1986).

39. Physicians perform poorly in maximizing utilities as well. In a classic article, Elstein et al. found that even when physicians' estimated probabilities are reasonably correct, they still did not reason consistently with them. Unaided clinical judgment, they argue, follows the principle of minimizing the most important risk, regardless of its probability (Elstein et al. 1986).

Politser summarizes findings from both statistical and information processing approaches. The question just how far the statistical and information processing approaches can be merged in this kind of research is much debated. Kahneman, Slovic and Tversky (1982), who are among the central authors in this cognitive psychologists' debate, argue that the two approaches should and can be merged. With their work, they state, they can both demonstrate that human judgments deviate from statistical theories like Bayes Theorem and explain these deviations in terms of information processing terminology. Gigerenzer and Murray vehemently deny that they succeed in doing so. In their view, the information processing terms Kahneman and his co-authors introduce (like the "heuristics" which explain certain deviations found) are in fact nothing but statistical categories (1987, 150-162).

40. *Cf.* Lusted (1975), and Kassirer and Pauker (1978). On medical practice variations, see e.g. Wennberg (1984).

41. *Cf.* McDonald (1976) and Kanouse and Jacoby (1988). It is not that Feinstein is not critical of the physician. Especially in his *Clinical Judgment* (1967), he is - and very much so. "If the clinician seems knowledgeable and authoritative, and if his reputation and results seem good, he can be condoned the most flagrant imprecisions, vagueness, and inconsistency in his conduct of therapy. The clinician does not even use a scientific name for his method of designing, executing and appraising [treatments]. He calls it *clinical judgment*" (26-27). Feinstein is not so much blaming the *individual* physician, however. He does not localize the cause of medical practice's problems in the physician's mind (although he wrote a series of articles addressing the reasoning processes of physicians: 1973a; 1973b; 1974). Rather, Feinstein blames the *medical profession* for its laziness; *for not having generated a standardized and validated language*. Given these conditions, there is not much an individual physician is to blame for.

42. Maxmen voices these ideas in rather strong terms. Other computer-based system-defenders often stress that such systems should not and cannot replace the physician. This does not affect the argument here, however. *Cf.* Barnett (1968), de Dombal et al. (1972) and Southgate (1975). See also Chapter 2.

43. Citation is from A. S. Elstein's "Comment" on Borak and Veilleux (1982).

44. The newly gained scientific status of medical practice did not help much to sustain the idea that medical practice was essentially in good shape. On a more general level, "science" had lost some of its indisputable status it had in the postwar years. In 1974, Moser still stated that "we *are* in a state of chaos, but it is simplistic nonsense to blame the mess on the clean, flourishing, nonmoral, logical and technical disciplines of science". Other authors were not so sure anymore. What had medical research meant for those concerned? Some research appeared to be only motivated by "the zeal to produce for the sake of recognition ... and the urge to present a paper at a scientific meeting in order to qualify for an expense-free holiday" (Anon. 1960). Medical science may be nonmoral, but it does have the capacity to visit "incomparable mischief upon mankind" (Ruby and Morganroth 1970). "It has become fashionable in academic circles to raise doubts about the effectiveness of all medical care," an editorial acknowledges in 1973. But "this is an understandable and needed reaction against the therapeutic enthusiasm that has been rampant in the profession and the general population for the past 25 years" (Haggerty 1973). In the early 1970s, in Starr's terms, the medical profession was witnessing the end of its mandate (1982, 379-393).

2

Multiple Rationalities: The Different Voices of Decision Support Techniques

In the 1940s and 1950s, statistical methods quickly became of paramount importance in psychology (Gigerenzer and Murray 1987). By the mid-1950s, reporting on experiments without discussing significance levels or null-hypotheses had already become unthinkable. This quick reception, Gigerenzer and Murray argue, had everything to do with the desire to create a truly scientific psychology. Incorporating a universal, objective method to test hypotheses was seen as a means of bridging the gap between the foundering attempts of the social sciences and True Sciences like physics.

As Gigerenzer and Murray point out, the new statistical tools (joined somewhat later with the digital computer) were subsequently transformed into *metaphors for the working of the mind.* Through looking at the mind in this non-behavioristic way, new questions could be posed, and new investigational strategies became possible. In studies of perception and detection, for example, the mind was seen to set levels of significance, and to test null-hypotheses as to whether stimuli had been perceived. Both computers and statistics were scientists' *tools* which were subsequently transformed into scientific *theories* (ibid., 3).

Analogous transformations have occurred in the case of decision support techniques. In the previous chapter, I have demonstrated how the different characterizations of "scientific medical practice" simultaneously constituted different normative frameworks with which the redescribed practice of medicine was judged. By focusing more on the publications of the tools' designers, this chapter zooms in further on the relations between the discourses and the tools.[1] It will become clear that the normative frameworks, and the distinctive shadows they cast on the physician's performance, have emerged *together* with the development of the decision support techniques.

In addition, redirecting the perspective more explicitly to the evolving tools and their builders allows me to elaborate on the theme of the differences between discourses. Through analyzing designers' depictions of their tools and their mutual criticisms, the ideal-typed views of medical practice and "rationality" *inscribed in* the tools are elucidated.[2] In zooming in, further differences within previously seemingly "unified" discourses appear. What seemed to be a single discourse on the level of mainstream medical journals becomes a multitude of fiercely debated, subtly diverging

views. With each type of tool, it is demonstrated, come distinct definitions of what medical practice is, what a rational practice looks like, and how a rational practice may be achieved.

I have divided the different decision support techniques into three categories. The statistical tools (including the diagnostic tools and clinical decision analysis) are discussed first, followed by the protocol and, finally, the expert system.

The Quest for Objective Inference: The Statistical Tools

Like the psychologists, the early postwar medical profession had also felt that the science of medicine, the footing of medical practice, could benefit from statistical methods. Especially where the design of clinical experiments was concerned, "the clinician should cooperate with the statistician from the start," editors argued (Anon. 1952b). Without a sound, statistical methodology, research work of clinicians would result in incomparable and falsely drawn conclusions. The promise of an objective, universal means to infer the truth of a hypothesis from given data was powerful.[3] In 1965 the *Journal of the American Medical Association* started a series called "The Mystic Statistic" to "eliminate the aura of mystery, the suggestion of arcanum, from biostatistics." "Modern medicine, the hard taskmaster of us all" simply demands of us to "suffer statistics," editors sighed (1957; Anon. 1965). More and more, statistics became the *sine qua non* of clinical research.

In medicine, then, a similar association between the "scientific" character of research and the application of inferential statistics came into being. As in psychology, and partly through the work of psychologists, statistical theories subsequently started to become models for depicting medical practice itself: physicians were seen as making decisions which were essentially *statistical* in nature. Here as well, this phenomenon was closely linked with coinciding developments in the field of computers. The explorations of what exactly the computer could do for medicine was inseparable from newly developing views of what medical practice was and how it could be improved. It was felt that "mathematics and mathematical devices can play [a very important role] in biological investigation and in clinical medicine" (Hubbard 1964). The search for this role, however, reshaped the nature of the game. Builders of statistical tools often closely cooperated with investigators probing the workings of the physician's mind, and they phrased their descriptions of medical practice in the same way (*cf.* Jacquez 1964; 1972). The statistical techniques and the calculating computer converged as metaphors for the workings of the physician's mind. The intricate intertwinement of these developments is portrayed in the first sentences of Ledley and Lusted's influential article in *Science*:

> The purpose of this article is to analyze the complicated reasoning processes inherent in medical diagnosis. The importance of this problem has received recent emphasis by the increasing interest in the use of electronic computers as an aid to medical diagnostic processing. Before computers can be used effectively for such purposes, however, we need to know more about how the physician makes a medical diagnosis. (1959)

Abdominal Pain Chart

NAME | REG NUMBER

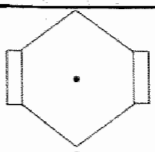
MALE FEMALE AGE | FORM FILLED BY

PRESENTATION (999, GP, etc) | DATE | TIME

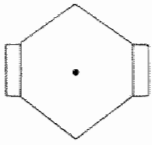
PAIN

SITE

ONSET



PRESENT



RADIATION

AGGRAVATING FACTORS
- movement
- coughing
- respiration
- food
- other
- none

RELIEVING FACTORS
- lying still
- vomiting
- antacids
- food
- other
- none

PROGRESS
- better
- same
- worse

DURATION

TYPE
- intermittent
- steady
- colicky

SEVERITY
- moderate
- severe

HISTORY

NAUSEA yes no

VOMITING yes no

ANOREXIA yes no

PREV INDIGESTION yes no

JAUNDICE yes no

BOWELS
- normal
- constipation
- diarrhoea
- blood
- mucus

MICTURITION
- normal
- frequency
- dysuria
- dark
- haematuria

PREV SIMILAR PAIN yes no

PREV ABDO SURGERY yes no

DRUGS FOR ABDO PAIN yes no

♀ LMP

pregnant

Vag. discharge

dizzy/faint

EXAMINATION

MOOD
- normal
- distressed
- anxious

SHOCKED yes no

COLOUR
- normal
- pale
- flushed
- jaundiced
- cyanosed

TEMP PULSE

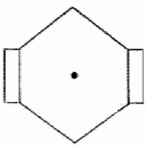
BP

ABDO MOVEMENT
- normal
- poor/nil
- peristalsis

SCAR yes no

DISTENSION yes no

TENDERNESS



REBOUND yes no

GUARDING yes no

RIGIDITY yes no

MASS yes no

MURPHY'S +ve −ve

BOWEL SOUNDS normal absent +++

RECTAL — VAGINAL TENDERNESS
- left
- right
- general
- mass
- none

History and examination of other systems on separate case notes

INITIAL DIAGNOSIS & PLAN

RESULTS
- amylase
- blood count (WBC)
- computer
- urine
- X-ray
- other

DIAG & PLAN AFTER INVEST

(time)

DISCHARGE DIAGNOSIS

Figure 1. Form used with de Dombal's tool (1991).

The answer to this last question, however, would then already be framed in computational, probabilistic terms: diagnosis had then already been modeled to the statistical tool Ledley and Lusted were developing.

The "tools-to-theory" phenomenon thus implies that it is only with the coming of these statistical tools that the process of decision making itself is constantly and matter-of-factly described in statistical terms. Moreover, with different tools, different images arise of what medical practice is and should be. I distinguish two types of statistical tools: *diagnostic tools* (often computer-based), and *clinical decision analysis*. They will be discussed in turn.

Statistical Diagnosis

In 1972, a group of physicians and computer scientists led by de Dombal publish a series of articles on a "computer-aided diagnosis" system, running in the University Department of Surgery in Leeds, UK. The system deals with the diagnosis of appendicitis: lacking highly specific tests, clinical diagnosis of appendicitis is a notorious problem in emergency medicine.

Their system works through filling in a form listing the required historical and examination data (Figure 1). Upon completion of the form, a computer utilizes a statistical formula (Bayes' Theorem) to calculate the probability of the patient in question having a certain diagnosis (Figure 2). The knowledge base of this system consists of a data base containing previously entered patient data; the diagnoses the computer can pick from represent the seven diseases most often causing acute abdominal pain (together covering over 95% of all cases).

According to the authors, the whole process of typing, processing and printing the results (see Figure 3) seldom lasts more than 20 minutes, and often takes no more than 5 minutes. The system performs well: de Dombal and his colleagues show that the computer has a diagnostic accuracy of more than 90% - while experienced physicians hardly ever reach 80%. Dramatically, the authors add that this implies that usage of the tool would significantly decrease the rate of "unnecessary" operations. Fewer healthy appendices would be removed, without inducing an increase in the number of ruptured appendices due to delaying surgery (de Dombal et al. 1972). Some years later their trials show that this is indeed the case (Adams et al. 1986; de Dombal 1989).

Many tools like de Dombal's computer-aided system were built in the 1960s and early 1970s: tools for diagnosis of congenital heart disease, thyroid disease, and so forth.[4] The builders of these tools describe medical work in specific ways. De Dombal and his team see physicians, like their tool, as making diagnostic decisions by "working from a set of ... probabilities" (Leaper et al. 1972). Medical diagnosis is a "problem in conditional probability"; the physician is a "diagnostic computer" which is "turning out possibilities on the basis of concurrence of several manifestations, together with such factors as age, sex, and the sequence of events" (Anon. 1961; McDermott 1971). Whether consciously, intuitively or otherwise, the physician comes to a diagnosis through a process of statistical inference.[5]

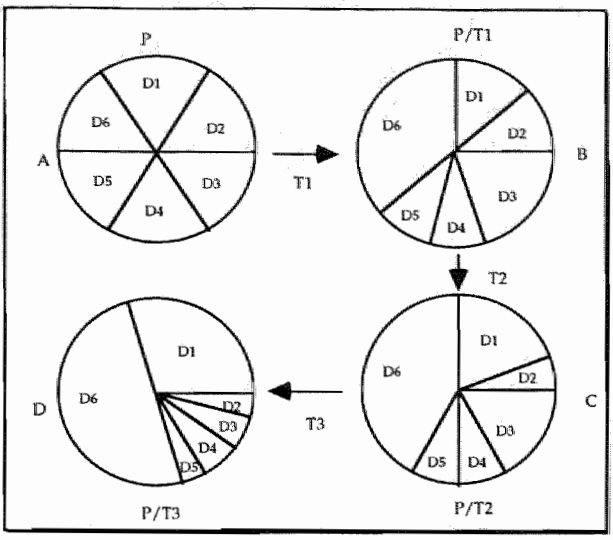


Figure 2. Bayes' Theorem states that one can calculate the probability of having a disease given a symptom by the following formula: P(D|S) = (P(S|D)*P(D))/P(S). In words: once you have the probabilities of exhibiting the disease (P(D)), of having the symptom (P(S)) and of having the symptom given the disease (P(S|D)), you can calculate the chance that a member of your population with symptoms S has disease D (P(D|S)). This illustration (Sultan et al. 1988, 223) shows how the sequential application of this formula results in a disease (D6) gaining prominence over five alternative diagnostic categories (T1-3 indicate the performance of three subsequent diagnostic tests).

```
POSSIBLE DIAGNOSIS
 APPEND DIVERT PERFDU NONSAP CHOLEC SMBOBT PANCRE
PROBABILITIES ARE
      0.0     0.0     2.7     0.0     0.9     3.1     93.2

CLINICIANS DIAGNOSIS
 PRIMARY     -CHOLEC
 SECONDARY   -SMBOBT

COMPUTERS DIAGNOSIS
 PRIMARY     -PANCRE  93.2
 SECONDARY   -SMBOBT  3.1

NEITHER OF YOUR DIAGNOSES SEEM LIKELY. PROBABILITIES
INDICATE PANCRE AS PRIME POSSIBILITY

++ SUGGEST CHECKING THE FOLLOWING.....
AMYLASE
TENDERNESS....
SITE PRESENT
```

Figure 3. Printout from de Dombal's system (from Horrocks et al. 1972). The computer has compared its diagnosis with the "clinician's diagnosis," which can be optionally typed in along with the clinical data. The seven mentioned diagnoses are: acute appendicitis, acute diverticular disease, perforated duodenal ulcer, nonspecific abdominal pain, acute cholecystitis, acute small bowel obstruction and acute pancreatitis.

As argued in the previous chapter, once the diagnostic task is seen as an exercise in statistics, physicians are seen to perform second-rate. In this view, physicians, like everyone else, "are not able to extract from data nearly as much certainty as is latent in the data" (Lusted 1968, xi; *cf.* Cebul 1988). The computer can do this for them. Since differential diagnosis is now a problem in conditional probability, it

> may be solved with a computer with an accuracy which depends only upon the accuracy of the statistical data regarding the incidence of symptoms in diseases and the accuracy of the data collected from the patient in question. (Anon. 1961)

Moreover, through continually updating its data-base, the computer can benefit better from its "experience" than the physician does (ibid.).

It is thus the tool itself which provides the normative yardstick against which the physician's performance is judged. Through supplementing the vague, inadequate diagnostic "judgment" of the physician by the computer's objective calculations, it is argued, medical practice would be better off.

What is the rationality inscribed in these techniques? Rational decision making here implies the existence of a formula which automatically reaches the best possible diagnosis given a set of data. *Context* is deemed irrelevant, since the essential elements of the decision are caught in the formula. All other context is bias. In addition, *content* would ideally be irrelevant too. According to de Dombal's team, the ultimate ambition would be that the computer, given a set of data, itself selected "from its files the most appropriate diagnosis from the whole spectrum of recognized clinical ailments" (Horrocks et al. 1972). Ideally, the statistical tool performs the trick, regardless of the specifics of the problem at hand.

Medical practice is made more rational by replacing the physician at the supreme moment, the diagnostic decision, by a process of mechanized statistical inference:

> The computer's conclusion will not be biased by irrelevant factors. For instance, the computer will not weigh most heavily its most recent experience, as a physician is prone to do. The computer performs just as well at 2 a.m. as at noon. (Anon. 1961)

Statistical tools, which had become the *sine qua non* of clinical research and had turned the clinician's experimenting into a truly scientific activity, now also became the epitome of rational medical practice (*cf.* Gigerenzer et al. 1989, 211-2).

Clinical Decision Analysis: Reshaping the Statistical Tool

The development of clinical decision analysis can be seen as an attempt, gaining momentum in the 1970s, to broaden the scope of the diagnostic tools. The method these authors championed deals with making optimally rational choices in situations of uncertainty. It grew out of mathematical approaches to decision problems which had gained much interest and impact through their usage by the military during and after World War II.[6] Decision analysis merges Bayesian statistics with economists' utility the-

ory: it opts to select the action with the highest expected utility out of a range of possible actions. It thus aligned with the "economization" of the profession's vocabulary mentioned in the previous chapter. Within this gradual but important change, the decision analyst's vocabulary, with its economic model of rational man, fitted perfectly.

The steps of clinical decision analysis were drawn directly from the original work of decision analysts like Savage (1972b, orig. 1954) and Raiffa (1968).[7] To find the action with the highest expected utility, the utilities of all the possible individual outcomes (the value of that outcome for the patient, expressed as a number) are multiplied by the probability that that outcome will occur. The action which yields the highest total expected utility is the action of choice (see Figure 4). As Figure 5 shows, when the possible outcomes and intermediary decisions increase, the tree gets very complex. Nevertheless, the method should in principle be able to generate the optimal decision for any medical decision problem.

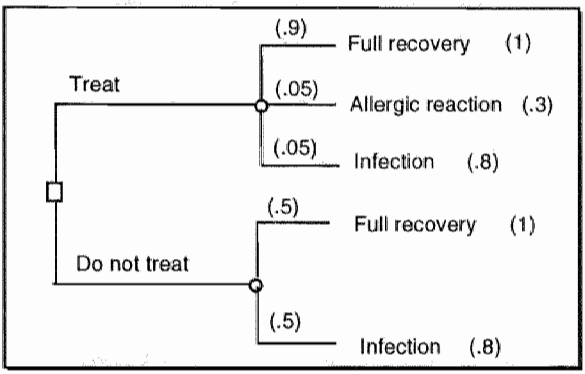


Figure 4. A simple decision tree for throat infection. A square node is a decision point, circular nodes indicate chance occurrences. The decision analysed is "to treat or not to treat." Possible outcomes are listed, and each possible outcome has a probability assigned to it (.9, .05 and .05 for the upper branch) and a value, between 0 and 1, stating its utility (1, .3 and .8 for the upper branch: an allergic reaction, which might be fatal, is worse than an infection, which is usually self limiting, but might lead to complications such as peritonsillar abscesses). The net score of "treat" is (.9 * 1) + (.05 * .3) + (.05 * .8) = 0.955, and the net score of "not treat" is (.5 * 1) + (.5 * .8) = 0.54. Since the expected maximal value of "treat" is higher than of "do not treat," the physician, given this situation, these probabilities and these utilities, should treat.

Both decision analysts and diagnostic tool makers see the physician's decisions as quantitative operations.[8] For decision analysts, however, focusing primarily on *diagnostic* decisions misses exactly those medical decisions which are important. In medical practice, they argue, important decisions are not so much those concerning what is the *case*, but those concerning what to *do*. In other words, what matters in medical practice is *action-selection* - including diagnostic actions. "Diagnosis" is just one little step in the whole process of "patient management," whereas "action-selection" is of continual importance (Lusted 1968, viii-ix; Ginsberg 1972). Moreover, decisions about which diagnostic test or therapeutic procedure

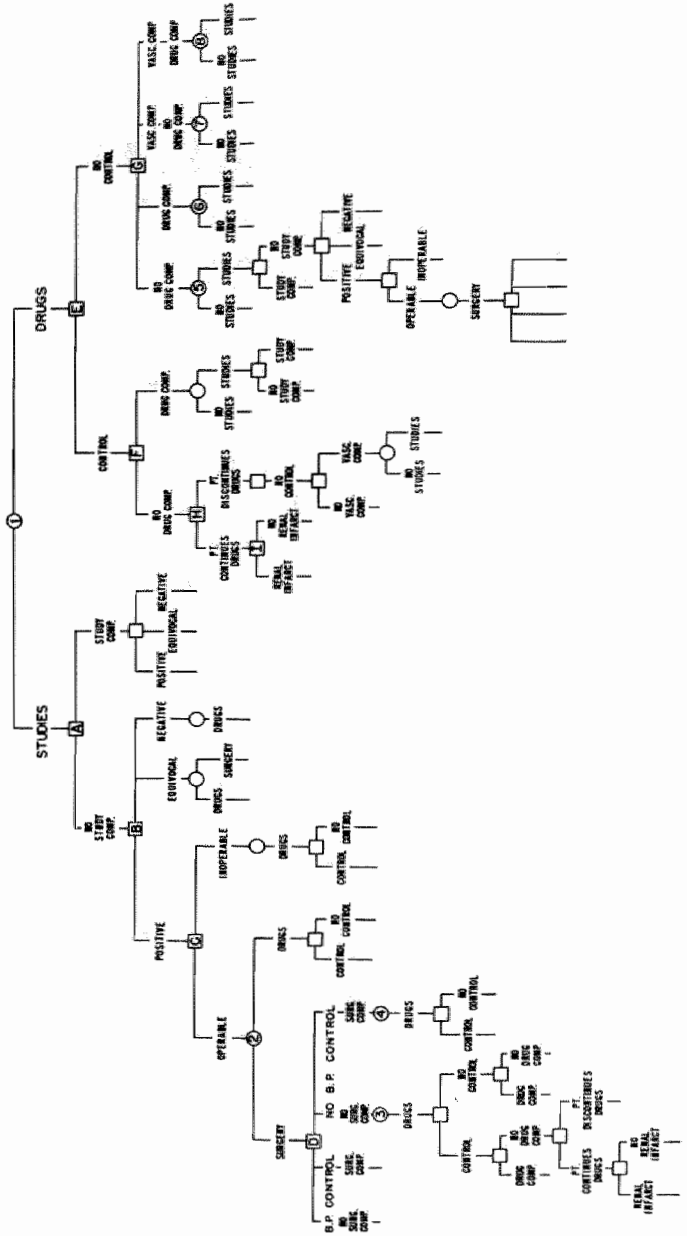


Figure 5. A more complex decision tree, describing possible actions and consequences in a patient thought to have either essential hypertension or renal artery stenosis. From Schwartz et al. (1973).

should be used often have to be made whether or not a diagnosis is already known. In fact, decision analysts argue, a "diagnostic decision" as such is hardly ever taken. Most decisions physicians face are like the decision trees portrayed: where treatment is required while the diagnosis is not yet known (Figure 4), or where a decision has to be made between a diagnostic or a therapeutic action (Figure 5).

Selecting an action in an uncertain situation, moreover, is always something more than just factually knowing what is the case. In all decisions, these authors argue, *value-judgments* have to be taken into account. A merely diagnostic approach

> has the serious deficiency that it is indifferent to the risks and pain involved in various tests and has no way of balancing the dangers and discomforts of a procedure against the value of the information to be gained. In this sense it lacks a key element that characterizes the practice of a good physician. (Gorry et al. 1973)

Choosing an *action* thus implies weighing risks and benefits - something the diagnostic tools could not do. Clinical decision analysis, according to these authors, offers an integrated, quantitative and objective approach to all physicians' decisions, combining the probabilistic nature of medical practice with the need to take the utility of outcomes into account.[9]

Another major difference is that decision analysts see the *patient's perspective* as an essential part of the rationalization of medical practice. Coinciding with a general trend in this period (*cf.* Armstrong 1983; Arney and Bergen 1984), clinical decision analysis discovered the "patient as subject". So far, Slack argues in 1972, the physician was expected to make the decisions for the patient, "all on a 'doctor-knows-best' basis." Medicine, however, should become "patient centered" as much as Rogers' client centered therapy, he argues:

> [There should be a kind of] partnership with the patient. In a step-by-step manner, advancing through the diagnostic process, the physician will communicate possible plans of action, and the patient will decide what should be done (1972).

For the diagnostic tools, the patient existed only as an object, from which clues had to be gathered to come to the right diagnosis. Here, however, the patient's subjective preferences (incorporated as "utilities") are a fundamental input of the decision analytic procedure. Adequate rationalization of medical practice, these authors hold, is impossible without full-blown and systematic attention to the patient's voice (Schwartz et al. 1973; Balla et al. 1989).

A final major modification introduced by the founders of decision analysis was the usage of *subjective* probabilities. The early diagnostic tools discussed in the previous paragraph (as de Dombal's tool) used objective probabilities: probabilities generated through analysis of large samples of patients. A major problem with objective probabilities, however, is that these data are often hard to obtain. Since clinical decision analysis' scope is so much broader, this problem is even more poignant:

> When seeking the data for test characteristics, therapeutic effectiveness, risk, and cost, the clinician hoping to use decision analysis is confronted with a bewildering array of conflicting definitions, methodologies, and means of reporting findings.

> Often, related research has not asked the decision-oriented questions that are of interest to the clinician, and it therefore does not even purport to supply the relevant data. (Cebul 1984)

When one cannot find the "massive amounts of clinical data" (Gorry et al. 1973) needed to build a decision tree in the medical literature, one can of course try to elicit the required probabilities from the files of one's own patients (as de Dombal's team did). The creation of a "sufficiently large and reliable database," however, is a "slow, tiring and troublesome job," which can easily take years (Bjerregaard et al. 1976).

To overcome this problem, decision analysts recommended using subjective probabilities, or "disciplined personal opinion," instead (Savage 1972a). To not let the rationalization of medical practice get stuck from the very outset, the physician is asked to supply what (s)he thinks are the probabilities, based upon his or her experience. Of course, decision analysts admit, these probabilities are often rather inaccurate. However,

> when there is a paucity of objective evidence at hand, we require a methodology that brings information, however vague and imprecise, into the analysis, rather than a method that suppresses information in the name of scientific objectivity. (Raiffa 1968, 155)

The power of statistical tools should not be left to wither in those tiny sections of the realm of medicine where objective probabilities do happen to be available.

Moreover, the absence of proper data is a problem for "our traditional, implicit or tacit clinical decision making" as well, decision analysts contend. "Because a decision must be made, the question is not whether probabilities will, in fact, be used, but what estimates are to be employed" (Schwartz 1979; Kassirer et al. 1987). Using the formal method of decision analysis, then, can only improve the quality of the decisions made.

Different Tools, Different Practices, Different Rationalities

The dissimilarities between clinical decision analysis and the diagnostic tools reflect different depictions of medical practice and different rationalities. Many diagnostic tool builders do not warmheartedly agree with the decision analysts' alternatives. They fear that the increase in applicability and scope decision analysis brings is bought by loosing the scientific, objective nature of their tools. The ideal of objective inference, in the eyes of some, is invalidated too much. For one, the use of subjective probabilities is intensely debated. Many statisticians do not even accept these probabilities as meaningful in the first place. Does it make any sense, these critics ask, to multiply a subjective probability with another? Or to add them (*cf.* Schoemaker 1982)? Moreover, cognitive psychology has shown that humans are rather poor in estimating probabilities. As Tversky and Kahnemann have argued, subjective probabilities show "large and systematic departures from proper calibration" and they change when the procedure by which the estimate is elicited is altered (1974). What, retort the diagnostic tool builders, should we make of decision trees construed

with such poor information? When de Dombal's tool was programmed with subjective probabilities obtained from physicians, it performed almost as poorly as the physicians themselves (Leaper 1972, interview de Dombal). Are we not creating "silk purses from sows' ears" (Cebul 1984)? What is left of the ideal of objective inference when all this subjectivity is let in through the back door?

Likewise, the emancipatory, humanistic ideal of incorporating patients' preferences is regarded with scepticism. The gathering of utilities is controversial and strewn with problems, decision analysts themselves admit. Elaborate and lengthy modes of questioning are needed to gather them properly, involving questions like "would you rather have a 10% chance of death but not loose your leg or have your leg amputated and not run a chance of dying?" (*cf.* Weinstein et al. 1980, 199). Since these methods are so awkward, authors argue, it might be that your measurements tell you more about the method used than about the patients' actual preferences (*cf.* Hershey and Baron 1987). Utilities appear to change when questions are framed differently and are highly unstable over time.[10] Gathering them is the "art and science of decision analysis" (Bell et al. 1988). Here again, it is argued, much vagueness and subjective judgment are let in.

The same, finally, can be said of the creation of the tree itself. The difficulty of this process is often commented upon. What outcomes to include? Should, in the example given in Figure 4, additional side effects of the antibiotics be included? Should the branch "do not treat" be expanded by including an option to wait one day to see whether the infection subsides - and if not, then treat? How detailed should you get? How important are omissions? These choices, decision analysts often state, are also part of the "art," the "subjective part" of decision analysis, where error lurks close. The "fidelity" of the trees is often poor, Kassirer and his co-workers remark; often, important outcomes are simply overlooked (1987; Schwartz 1979).[11] The usage of the tree itself, in other words, already violates the ideal of objective inference. The way it is structured greatly influences the outcome of the decision analysis, yet it is saturated with subjective guesses and non-quantified *content*.[12]

So we see two different tools, linked to two different views of what medical practice is, how it should be rationalized and what "being rationalized" *means*. They disagree about what the central "act" of medical practice is, and about the ideal of objective inference. While both groups of authors adhere to this ideal, clinical decision analysts feel that it should be interpreted somewhat liberally in order to get a meaningful tool in the first place. On the contrary, diagnostic tool builders point out that by condoning this polluting of the objective inference ideal, you just replace one type of messiness by another.[13]

Diving into the backgrounds of the "statistical discourse" we encountered in the previous chapter (one of the variants locating the scientific character of medical practice in the physician's mind), we found two types of tools which lay at its root. The differences between these tools are hard to discern from the perspective of the mainstream medical literature; they only come into view when one closes in on the specific tools themselves. Only then does it become clear how, in and through the development of these tools, medical practice has been redescribed to their likeness. Only

with the coming of the statistical diagnostic tools was the process of differential diagnosis seen as a problem of conditional probabilities. The same process occurred in the case of clinical decision analysis, which is often said to do nothing but "make the hidden decisions explicit" (Lusted 1975). This only became the case, however, once the desires of patients were seen as the expression of "personal utility structures," and once medical decisions were redescribed as "weighing risks and benefits" (Weinstein et al. 1980, 299) - transformations which only occurred with the coming of decision analysis.

Capturing the Clinic: The Protocol

A second category of tools attempting to rationalize the practice of medicine includes an array of techniques which go by a plethora of names: guidelines, algorithms, practice policies, standards, statements, protocols. The different labels are used in widely diverging ways: what Eddy calls "practice policies," the US Committee on Clinical Practice Guidelines calls "guidelines"; what some call "protocols," others refer to as "algorithms."[14] All of these tools, however, have in common that they are or can be read as a set of *instructions* telling medical personnel to do A in situation B. These instructions may be more or less elaborate, precise, or binding; they may be formatted in different ways, and may have been construed off-handedly or in a highly structured way - but they all share this common feature. They may be elaborately designed as a flow-chart containing great detail, or they may consist of a number of rather vague and general recommendations, but they all guide medical personnel through a sequence of steps. It is as an umbrella term indicating this feature, then, that I use the term "protocol."

Given this plethora of terms and techniques, it will come as no surprise that the background of the protocol in medical practice is difficult to circumscribe. The emergence of such tools in medical practice was much less of a traceable, concentrated activity than the development of the statistical tools. In its simplest form, a "protocol" is nothing but a written instruction; and, of course, the idea of regulating action through a recipe is an ancient one (*cf.* Goody 1977). Protocols figured in most of the discourses discussed in the previous chapter. We saw protocols which should overcome troubles generated by agencies external to medical practice (thus fitting the early postwar discourse on medical practice and its problems), and we encountered protocols which were to compensate for the physician's cognitive incapacities. In the latter case, the resemblance of the sequential and logical nature of the protocol to mental processes was emphasized.

It cannot be said, then, that "the protocol" was uniquely tied to one of the discourses on medical practice discussed. As such, the protocol has been incorporated in several of these discourses. On the other hand, I do argue for a special link of the protocol to the discourse Feinstein and Weed helped come into being. It was in this discourse that protocols, as a tool relevant for medical practice, came into full view - and in a highly specific way.

Feinstein and Weed modeled the practice of medicine to the steps of the scientific method, as taught by neo-positivist philosophy of science. Riding the wave of the standardization efforts engulfing the practice of medicine, they argued that standardization could do for medical practice what it had done for medical science. The success of medical science during and after World War II had shown the merits of strengthening the collective effort through the coordination and linking of individual actions. Protocols were part and parcel of this enterprise: in the booming field of clinical research, the protocol was essential to assure that the actions and interpretations of outcomes would be similar in all participating institutions (Howard 1961; *cf.* Marks 1988).[15]

If medical practice would follow, if here cooperative work would also be made more efficient and communication would be streamlined, this particular science could fully expand as well. Here, the protocol and "(scientific) medical practice" were brought together in a way which transformed the meaning of both these notions. Medical practice became defined as the logical and sequential (i.e. protocol-like) execution of the steps of the scientific method; the standardizing protocol, concurrently, became a tool to structure and coordinate this scientific work of diagnosing and treating patients.

The circumscribed view of medical practice and rationality connected to this protocol and the way it is positioned vis-à-vis the other types of tools are discussed in the remainder of this section. I first discuss Feinstein's eloquent and harsh criticism of the statistical tools. Next I describe a typical example of the type of tool which is linked to this kind of criticism: a protocol for "physician extenders" developed in the early 1970s. As can be expected, the statistical tool builders regarded the criticism and tools of Feinstein cum suis with scepticism: their counter-critique is discussed in the last subsection.

The Haze of Bayes and the Aerial Palaces of Decision Analysis: A Criticism of Statistical Reasoning[16]

> Decision analysis has led us into Vietnam: where will it lead health care?
> (graffito in the Harvard School of Public Health, 1979)[17]

Feinstein attacks both the practical functioning of the statistical tools and their inscribed ideal-typed rationalities and views of medical practice. In his view, medical practice is nothing like a mathematical activity, and the ideal of objective statistical inference is an utterly mistaken goal to strive for. His criticism, developed in many years of prolific writing, can be somewhat crudely summarized in two main points.[18]

First, *why replace judgmental decisions with a logic of statistical reasoning when the latter requires judgmental decisions as well?* The statisticians' critique on clinical judgment, Feinstein argues, is that it "often may be applied in an inconsistent, unstandardized, or even capricious manner. The great appeal of ... mathematical models is that they offer a standardized mechanism for this process" (1987a, 113-4). The problem with this endeavour, Feinstein states, is that it *also* requires many ad hoc, non-mathematical

judgments. First of all, mathematical tools always imply a wide range of assumptions. Medical practice, however, more often than not violates these assumptions:

> To apply the Bayesian concept in general clinical diagnosis requires assumptions about nature that are incompatible with realistic clinical activities. The types of data and the required "independence" of different variables in the Bayesian diagnostic calculations can be obtained only if clinicians ignore the epidemiologic realities of human ailments: many people have ... [undiscovered] disease, and many people have multiple co-existing diseases. (1967, 372)[19]

When are such violations severe; when allowable? Questions like these continually require pragmatic, subjective judgments. Clinical decision analysis faces such questions even more frequently: as the statistical tool builders already charged, setting up the tree, estimating subjective probabilities, and eliciting utilities constantly involves judgmental decisions (*cf.* Feinstein 1987b).

Stressing the fact that he is a clinician himself, Feinstein argues that statisticians hardly understand how inevitable and frequent these violations will be. The logic of statistical reasoning will simply drown in the messy reality of actual clinical practice: it completely overlooks the pragmatic, fluid nature of clinical reality (1977). Statistical tools, for example, often require the input of sensitivity and specificity values of the relevant diagnostic tests (measures of the capacity of the test to distinguish between the absence or presence of e.g. a disease). "The mathematical quantophrenia," however, leads us to overlook that the sensitivity and specificity of a test are often highly dependent on the reason *why* a physician requests the test (whether for discovery, exclusion, or confirmation), the severity of the disorder, and so forth. In different situations, we are dealing with different groups of patients, which thoroughly affects the magnitude of these measures. When all these matters are taken into account, Feinstein asserts, "we can promptly begin to discern the scientific inadequacy of all the mathematical folderol" (1985, 609-18).[20]

Similarly, diagnostic and therapeutic actions are often interwoven in medical practice, making a Bayesian "diagnosis-machine" irrelevant (Feinstein 1973a). Here, decision analysis is not the solution it claims to be, since building decision trees demands the utterly unrealistic attempt to foresee all possible outcomes, all intermediate actions, and decide upon all the required probabilities and utilities (1977).

The pointlessness of this task is even more glaring when one considers the building blocks used. Medical nosology is a "potpourri," Feinstein continually insists, with ambivalent terminology and vague criteria which mean something else for every other physician. Any attempt to build a clean, precise, mechanized inference tool upon this quagmire will be thoroughly thwarted from the outset (1973a).

Feinstein argues, thus, that "the use of the mathematical models merely transfers the judgmental decisions from one intellectual location to another" (1987a, 114). The "messiness" of medical practice is much more widespread and unyielding than the statistical tool builders appear to think. As a consequence, the highly undesirable situation arises that these decisions are transferred to the hands of people who do not know anything

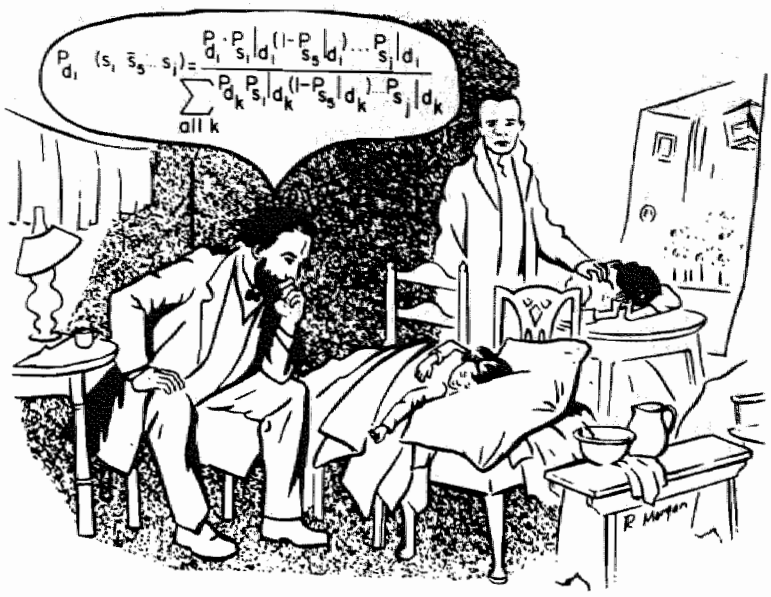


Figure 6. Cartoon criticizing a Bayesian approach to medical practice. From van Bemmel and Willems (1989, no original reference).

of the intricacies of medical work: statisticians. The tools we end up with, he concludes, are nothing but

> a splendid array of castles in the air - having all the abstract academic virtues of aerial palaces and none of the gritty dirt associated with a strong foundation or firm roots in reality. (1967, 370; 1977)

Feinstein's second main point, closely related to the first, is that *statistical goals often conflict with clinical goals*. So called "soft data" (such as visual impressions or psychosocial information), Feinstein argues, are of crucial importance in medical practice. Such data invariably tend to drop out the statistical tools since they are "too vague," thereby threatening to further dehumanize medical practice (1967, 370-379). Moreover, and crucially, Feinstein argues that the "decontentualized" nature of the statistical tools is unacceptable and utterly unscientific in a clinical context:

> The mathematical goals are aimed at eliminating details, using standardized models, and producing maximum reductions of variance in the available data. ... [If] the clinician wants to preserve details, observe direct evidence of relationships, ... and arrive at conclusions that are clinically both cogent and consistent, the conventional mathematical goals will not always be satisfactory. (1987a, 113)

Physicians do not simply want to hear that a symptom is correlated with a disease. They want to *understand*. The essence of scientific action, after all,

is *explaining* this correlation through for example a pathophysiological mechanism. Statistical methods merely "label," Feinstein argues (1973a). Their conclusions "may be accurate but impertinent, ... statistically satisfactory but scientifically defective, diagnostically inadequate, and therapeutically hazardous" (1973b). More often than not, clinicians are not very interested in the "number issued by the calculations of probability"; their actions are guided by "many features much more subtle than a simple calculation of diagnostic probabilities" or a "clinically alien arrangement of statistical scores" (1967, 372-3; 1987b).[21]

Feinstein stresses that he is by no means an "antistatistical nihilist." It is just the "unscientific manner" in which statistics are being applied to medical practice which bothers him (1973b; 1994). Developing good tools, Feinstein argues, requires

> adapting statistics and computers to the practical realities of clinical medicine rather than forcing clinical phenomena into Procrustean modifications to fit the theoretical concepts of statistics and computers. (1973a)

In the next subsection, I offer an example of the type of tool Feinstein supports. Subsequently, the specific view of medical practice and rationality contained in this undertaking is explicated.

The Clinical Grounding of the Protocol

In 1974, Greenfield's team reports on their experience with the use of a protocol for "physician-extenders."[22] Patients with upper-respiratory tract complaints entering the walk-in clinic of the Beth Israel Hospital in Boston are seen by a health assistant: usually a briefly trained high school graduate. The health assistant first scans a list of "chief complaints" indicating suitability for the protocol (Figure 7). If one of these complaints is present, a data collection form (Figure 8) subsequently guides the health assistant through a series of questions regarding symptoms and relevant history (does the patient "ache all over," is (s)he taking antibiotics, and so forth). Furthermore, the protocol requires some signs to be assessed through physical examination: whether for example sinuses or neck nodes are tender. Finally, the health assistant follows the steps outlined in the "decision

Acute cough (< one-week duration)
'Cold'
'Flu' or influenza
Hay fever
Hoarseness
Postnasal drip
Sinus trouble
Sneezing
Sore throat
Streptococcal troat infection
Stuffy or runny nose
Tonsil trouble or tonsillitis
Request for throat culture

Figure 7. List of chief complaints, indicating suitability for protocol. From Greenfield et al. (1974).

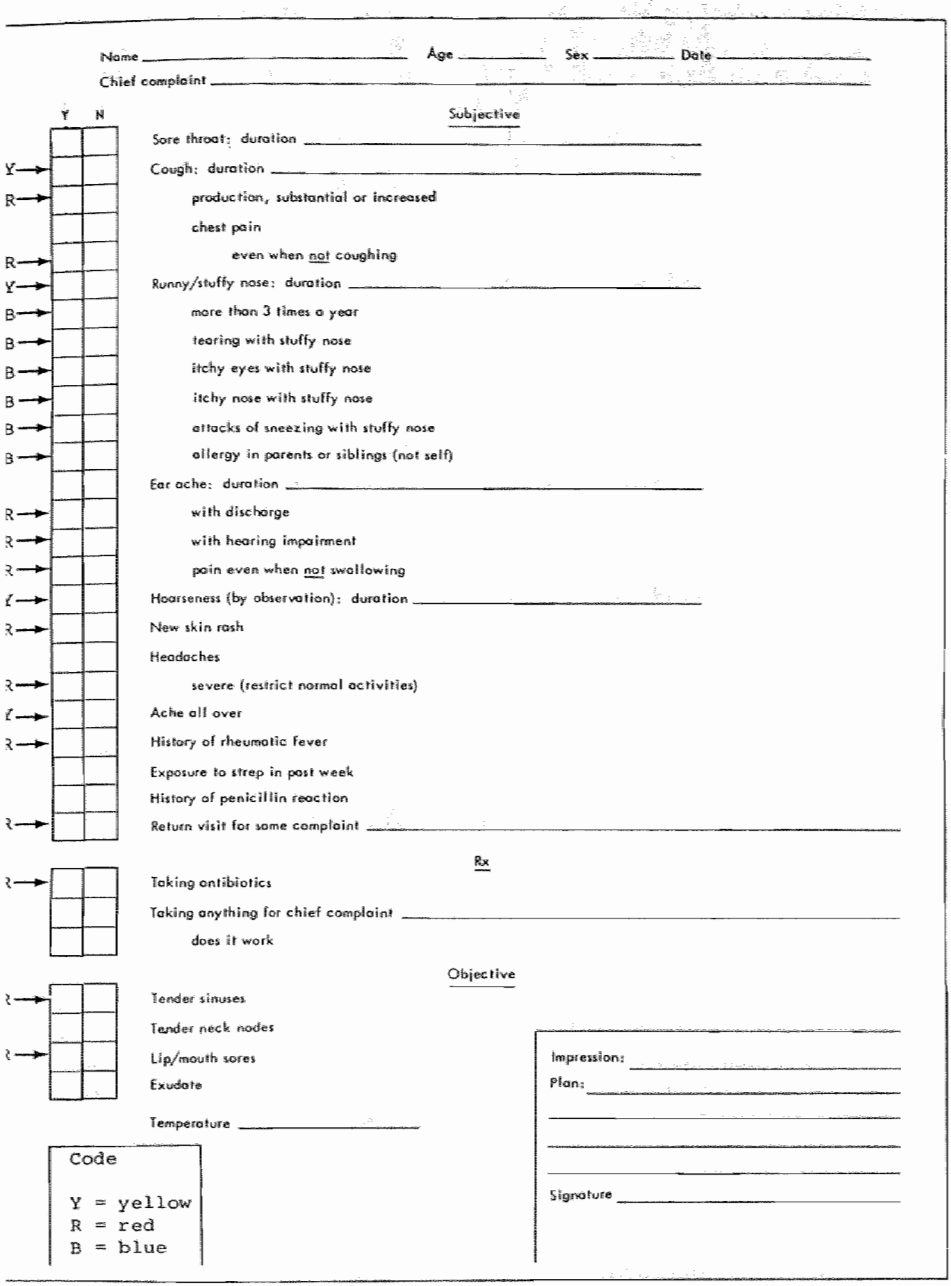


Figure 8. Data collection form for protocol (Greenfield et al. 1974). Information is denoted as present ("Y") or absent. "R," "Y," "B" indicate the colors (red, yellow, blue) of these boxes used in the decision-making algorithm.

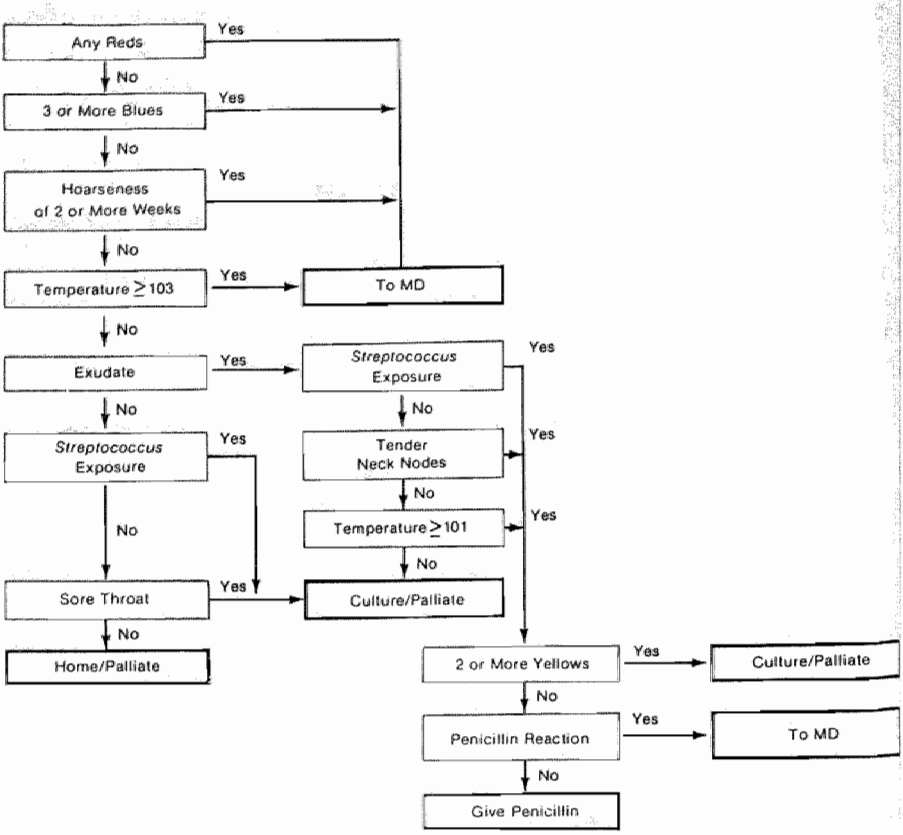


Figure 9. Decision making algorithm of Greenfield protocol (ibid.).

making algorithm" (Figure 9). The algorithm checks the absence or presence of (sets of) symptoms and selects one out of four possible "medical actions": send home without specific treatment, make throat culture, give penicillin or refer to the physician.

The major "medical reasons for accurate work-up of upper respiratory tract infection," the authors argue, "is the detection of streptococcal pharyngitis": a bacterial infection of the throat for which antibiotic treatment is effective. Moreover, "if the patient does have a streptococcal infection, he might be at higher risk for rheumatic fever."[23] The protocol attempts to ensure that patients who might have something "serious," or something requiring a different diagnostic workup, are immediately sent to the physician: one symptom or sign marked "red," or three or more symptoms and signs marked "blue" (Figure 8), are indications that something like pneumonia, herpes infection, otitis media or allergy may be present.

The algorithm makes "a differential diagnosis based on the assumption that clinical judgment flows from an analysis of many overlapping features that may be present to varying degrees." In this case, the challenge

is that the symptoms and signs of a viral throat infection (indicated by the presence of more than two yellow symptoms) may appear similar to those of a streptococcal infection. "In choosing what attributes are to be included for directing the protocol logic," the authors continue:

> decisions are derived from a combination of practical and purely medical considerations, [which were subjected to peer and consultant review]. These may not be universally acceptable to all physicians; however, if they are explicit, each can be rejected or modified to fit the view of the individual physician. (Greenfield et al. 1974)

So, the authors argue that "under ideal circumstances" one would probably never treat immediately with penicillin: the diagnosis would first have to be confirmed through a positive culture. Only then would the patient be treated. However, they state,

> lost cultures, lost patients, and a respectable rate of false-negative single cultures (about 10%) make waiting hazardous. For these reasons we treat these patients immediately, in accordance with the practice of many, and in what we believe is a sound convergence of theoretical and practical considerations.

A culture is taken only if the algorithm ends up with a clinical picture with "overlapping features": when a distinction between viral or streptococcal infection cannot be made.

Greenfield's team tested the protocol by having it administered to 226 patients, who were subsequently seen by the physician for evaluation. All in all, they say, 42% of the patients would have been sent home by the protocol - of which none, according to the physicians, had pathological conditions other than uncomplicated (non-streptococcal) upper-respiratory tract infection. This would have meant a considerable savings of physicians' time. Furthermore, they argue, the protocol is safe since

> the decisions about data-base collection and disposition are not made by the health assistant. They are made by the protocol, derived from local experience and peer consensus.

What is the image of medical practice associated with these types of protocols?[24] What is a *rational* practice, and how is this to be achieved? Good algorithms, Feinstein argues,

> describe good clinical reasoning: rules that are specific enough to manage the standard situations, broad enough to encompass the common exceptions, and flexible enough to allow separate decisions for the rare. (1974)

The protocol structures medical practice in such a way that its essential *clinical* nature is retained, while introducing the "reproducibility and standardization" which will make a true science out of medical practice. "The object is to preserve the vitality of clinical reasoning while enhancing its scientific effectiveness" (ibid.). So, the protocols contain pragmatic but optimal, sequential paths through data gathering, interpretation, requesting additional paraclinical tests, prognostic estimations, therapeutic actions, and behavioral strategies. Pathophysiological theories are and have to be connected to practical considerations, and diagnostic decisions often come

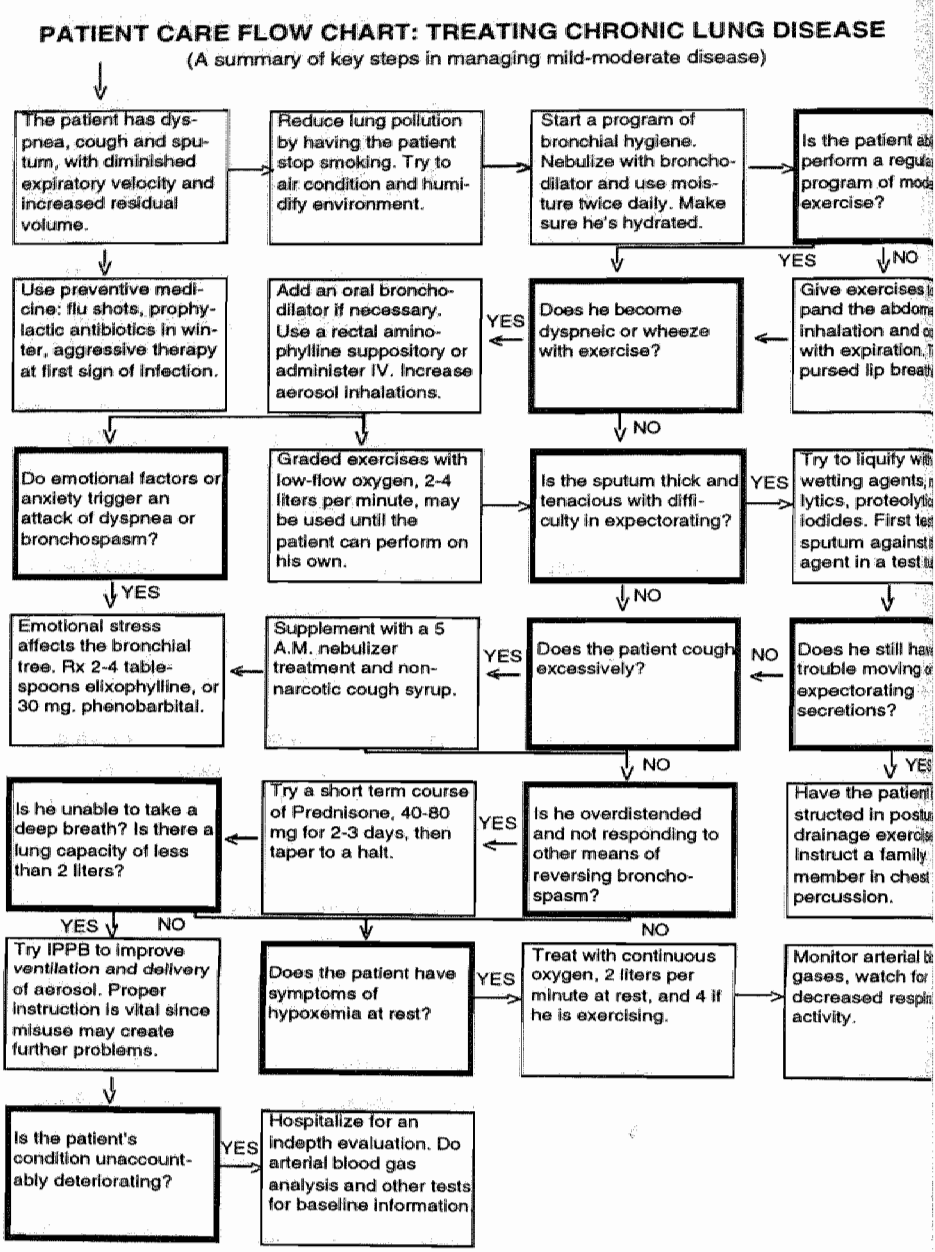


Figure 10. Feinstein (1974) gives this example of an "algorithm" for therapeutic management of chronic lung disease to illustrate the way different decisions are interwoven in a singular protocol.[25] Note also how the protocol's course of action stretches out over several days, and involves several different members of the medical staff.

only after the therapeutic act. Protocols try to strengthen this clinical rationality, without trying to *replace* it by some alleged superior rationality as the statistical tools did. A protocol simulates the physician's actions, but includes "the rationale for each decision which it makes" (Komaroff et al. 1973). As Feinstein argues:

> The addition of suitable justification, containing citations of data or principles to substantiate each decision, is the activity that converts a flow chart [a term for the representation of a protocol as in Figure 9] from an arbitrary set of rules into a scientific document. (1974)

"Rationality," thus, coincides here with attention to "soft" data, with the need to understand *why* (against the statisticians ideal of statistical inference), and with a pragmatic stand forced upon one by the demands of the clinic.[26] In addition, the ideal-typed protocol *intervenes* in a rather different way than statistical tools. The latter, in principle, focus only on the moment of *decision*. At that moment, when physicians are at their weakest, the tool momentarily takes over. On the contrary, the protocol standardizes a whole sequence of actions. A protocol can contain several decisions at once, since "an efficient clinician regularly intermingles many other decisions with the diagnostic reasoning" - such as prognostic estimations, choices of additional tests, selection of therapeutic agents, and planning for the personal interaction with the patient (Feinstein 1974; *cf.* Margolis 1983). Although many current protocols do have a single, often diagnostic decision as their end point (as Greenfield's protocol), the realities of clinical work will require an increase in the production of protocols as exemplified in Figure 10 (Feinstein, ibid.). Rather than intervening at a single moment, the moment of decision, the protocol submits the user to its control for an extended period, guiding a whole series of actions and intermediary decisions.[27] Moreover, at the moment of decision the statistical tools retract into their own, statistical realm, inaccessible to the physician. The protocol, on the other hand, never leaves its users. They can follow the protocol's path, drawn out on the pieces of paper before them.

The Statistical Counter-Critique; The Development of Consensus Reports

The protocol as described here was thoroughly criticized by those who held more closely to the statistical approach. First, statistical tool builders scorned that the protocol is far too inflexible and simple a tool to be of much help in medical practice. "The algorithmic approach has limitations when one is dealing with uncertainty," Kassirer and Kopelman argue (1990). The "branching tree structure demands arbitrary cutoffs for clinical variables," thus presenting clinical situations "in a black-and-white fashion, when we all know how gray they often are." Protocol builders underestimate the complexity of medical practice - a complexity which can never be caught in simple, rigid recipes, where "the progression ... is always the same" (Sultan et al. 1988). Moreover, since these tools are so simple and rough, they end up being nothing but a form of additional regulation - creating rigidity, stifling new development and creating a "self-perpetuat-

ing bureaucracy" (Kassirer and Pauker 1978, see Chapter 1). Statis tools, on the other hand, are highly flexible and can in principle handl infinitely diverse array of cases. They do not attempt to rigidly fix a w range of actions, but only intervene when the important decision has t made.

The second major criticism concerns the foundations of the proto Their design, it is argued, is often far too eclectic. Science, after all, is cise, rigorous and universal: creating ambiguous, pragmatic and l tools is nothing to be proud of. All the local efforts of setting up proto have only led to a plethora of subjective, conflicting rules to govern mec practice (Wennberg 1991).

These types of protocols, furthermore, only too often rest "on a gerous tautology": they conflate what physicians *should do* and what *are doing* (Eddy 1990d). They merely restate what physicians have b doing all the time. This tautology can only be broken by using the uni sal, rational tools of decision analysis, Eddy argues. Kassirer and Ko man agree:

> a principal difference [between algorithms and decision analysis] is that the l is based on solid theoretical grounds, namely on probability and utility theo one accepts the precepts of these theories and if one accepts the numerical va for probabilities and utilities used in an individual analysis, one should, if rati accept the outcome of such an analysis. The theoretical underpinning for algori is less defined and rarely made explicit. (ibid.)

Finally, in translating knowledge into the shorthand notation flow-chart-like structure of a protocol, Kassirer and Kopelman arg "something is invariably lost" (1990). The ambiguity which results fr this loss in precision facilitates misinterpretation.

We "should delve much more deeply into the theory and basis o gorithm design and content before we grow more trees of this kind" th counter-critics conclude (ibid.). In their view, subordinating a statist logic to a clinical one is putting the cart before the horse.[28]

One influential development which originated against the ba ground of this critique[29] was the phenomenon of *consensus reports*. Star in the USA in 1976, and spreading rapidly to many different countries, development of consensus reports is a means through which physici are being kept informed of the most up-to-date methods and technique treat and diagnose classes of patients. A group of experts on a given main meet and form a "consensus" on what the current scientific opin is of a selected problem. This "processing" of scientific information, cording to the proponents of this method, is necessary to fill in the gap our scientific knowledge and to "translate" scientific information i practical, usable guidelines. Protocols have been designed for the decis to do a cesarean birth, the treatment of breast cancer, and so forth.[30]

These consensus-reports can be seen as an attempt to create a m "scientific" alternative to the protocols described above. The explicit tempt is to design a *universal* (or at least national) protocol, which is no eclectic mixture of physicians' subjective hang-ups. Rather, the rep should be maximally based upon evidence attained through the epitom statistical rationality for clinical research: the clinical trial (*cf.* Jacoby 19

Lomas et al. 1988). In addition, decision analytic techniques can be used to find the most desirable management option (*cf.* Jacoby and Pauker 1986).[31]

Nevertheless, this adaptation is not very impressive from the statistical logic's point of view. The idea of the consensus report, with its attention to universality, its clinical trial foundations, and its openness to decision analytic methods, is seen as a step ahead. The idea of "consensus," however, of filling in the gaps and "adapting" scientific knowledge for practical purposes, still sends shivers down the spine of the more forthright upholders of the statistical ideal. All this tinkering just opens the floodgates of subjectivity and unscientific idiosyncrasy. "Who can say that a guideline developed by an expert panel is correct?" Eddy argues. "Indeed, what does a consensus of a group whose perceptions [of e.g. the effect of a therapy] might vary from 0% to 100% even mean?" (Eddy 1990b, interview Pauker, October 29, 1993). The need to assemble a workable recipe, it is feared, will lead to reports which are more based on "compromise" than on actual consensus.[32] Moreover, the structure of the protocol has not changed. Simple, rigid recipes are simply unsuitable for assessing "clinical problems that are controversial" (Kassirer and Kopelman 1990).

Finally, consensus reports tend to be much more general, much less detailed, than the protocols described above (*cf.* Everdingen 1988, 161). This is no coincidence. The tight, direct and encompassing control of the protocols for physician extenders was utterly unacceptable to physicians. Consensus reports, on the contrary, have a very different status. In order to get the physician to cooperate, such protocols are usually rather indeterminate guidelines with a non-compulsory character. From the perspective of the statistical counter-critics, then, this is a dead-end street. Loosening the grip of a bureaucratic straight-jacket will not solve the protocol's essential shortcomings, since this can only result in more space for subjectivity and bias.

Expert Systems: The Best of Both Worlds?

> The Mind came in on the back of the Machine.
> (George Miller)[33]

Rooted in work done and alliances made during World War II, the field of Artificial Intelligence was born in the 1950s with the work of Newell, Simon and others.[34] As briefly described in Chapter 1, these authors argued that the human mind functions as an Information Processing System. Rather than modeling the human mind to a calculating "mathematical device," as the early statistical researchers did (see above), these authors drew upon the metaphor of the programmed digital computer (*cf.* Baars 1986). Simon expresses this position succinctly:

> Like a modern digital computer's, Man's equipment for thinking is basically serial in organization. That is to say, one step in thought follows another, and solving a problem requires the execution of a large number of steps in sequence. ... There is much reason to think that the basic repertoire of processes in the two systems is quite similar. Man and computer can both recognize symbols (patterns), store symbols, copy symbols, compare symbols for identity, and output symbols. These pro-

> cesses seem to be the fundamental components of thinking as they are of com tion. (1979; *cf.* Pylyshyn 1980)

According to these authors, the way to study human intelligence was t to imitate human problem-solving abilities by a computer program.

Building upon this information processing framework, the e> system became a favourite research object of the Artificial Intellig community in the 1970s. The turn to expert systems implied a shift i search for a way to build programs which could behave intellige Earlier researchers, such as Newell and Simon, had tried to look for a general, problem-independent mechanisms underlying all human p lem solving behavior (1972). Expert system builders argued that thes tempts of the Artificial Intelligence pioneers had not been very succes This work "was dominated by a naive belief that a few laws of reaso coupled with powerful computers would produce expert and superhu performance" (Hayes-Roth et al. 1983, 7). This was too grandiose an dertaking, expert system builders profess. Rather, "high performance soning systems" can only come into being when sufficient attentic given to the *knowledge* required for expert functioning. Simultaneously, shift to a "knowledge-based approach" also implied a shift to nar specialized domains of expertise. As Duda and Shortliffe argue:

> Many human experts are distinguished by their possession of extensive knowl about a narrow class of problems. It is this very limitation that makes it feasit provide a computer program with enough of the knowledge needed to per those tasks effectively. (1983)

Expert system pioneers ventured into medicine since they felt th was an optimal domain to try out their theories. The expertise in this main, the medical knowledge of specialists, seemed optimally structu It is circumscribed, and relatively explicated: it deals with specific fa related in relatively clear ways (*cf.* Davis et al. 1977).[35] Different syst were developed in the early seventies (both medical and non-medical) of them in the United States.[36] In the next subsection, I introduce MYC one of the most famous of these "first generation" systems.

An Early Promise: MYCIN

MYCIN, the dissertation project of Edward Shortliffe, was developec Stanford University, California, as a tool to aid physicians in the select of antibiotics for patients with severe infectious diseases (Shortliffe et 1973; Buchanan and Shortliffe 1984). The required knowledge was ga ered through interviews with physician experts, literature reviews, a scanning patient files of actual cases (Yu et al. 1979). MYCIN's knowlec was contained in some 500 "procedural" (if.. then..) rules, which cove the areas of meningitis (infection of the cerebrospinal membranes) a bacteremia (bacteria in the blood). Each rule embodies

a single, modular chunk of knowledge and states explicitly in the premise all necessary context ... It forms, by itself, a comprehensible statement of some piece of domain knowledge. (Davis et al. 1977)

Rephrased in English, such a rule looks as follows:

IF:
1) the infection is meningitis, and
2) organisms were not seen in the stain of the culture, and
3) the type of infection may be bacterial, and
4) the patient has been seriously burned,

THEN:
there is suggestive evidence that *Pseudomonas Aeruginosa* is one of the organisms that might be causing the infection.

(From: Duda and Shortliffe 1983).

The program begins by asking a series of questions to the physician who has invoked MYCIN's help (see Figure 11). Through attempting to apply rules, the program tries to determine the type of infection and the possible organisms involved. In the case of the rule stated above, the program tries to see if the premises hold: whether the patient indeed has meningitis. The program can either directly ask the physician if this is the case or, if the physician does not know, it can try to infer this through applying other rules. When, finally, the premise is fulfilled (as is the case, in Figure 11, after question 14), the program checks if the second premise is fulfilled as well, and so on. As Duda and Shortliffe explain, MYCIN "attempts to achieve any goal by applying all the directly relevant rules. The need to establish the premises of those rules sets up new subgoals that are treated in the same way" (1983). Once MYCIN has completed its diagnostic phase, it determines an optimal combination of antibiotics to be administered to the patient (see Figure 12).

In 1979, MYCIN's medical accuracy was subjected to what Yu et al. proudly presented as "the first clinical evaluation by independent experts" ever performed of a medical expert system. MYCIN and eight specialists had to choose antibiotic treatment for ten "diagnostically challenging" paper cases. These prescriptions, subsequently, were evaluated by "prominent infectious disease specialists," who did not know the identity of the prescribers. Compared to the physicians, MYCIN performed well. The system did not fail to cover a treatable organism, whereas some of the physicians did. In addition, its prescriptions were deemed "acceptable" by the expert panel more often than those of the physicians (65% versus an average of 55.5%).

Notwithstanding these encouraging results, MYCIN was never used in practice. Apparently, Buchanan and Shortliffe acknowledge, showing that a system performs well is not enough. Physicians found it irritating that they could not offer information: they had to wait for the program to ask them. Moreover, MYCIN asked a relatively large number of questions (usually more than 50), which took time - a scarce resource for most physicians. These and other problems led the MYCIN team to abandon

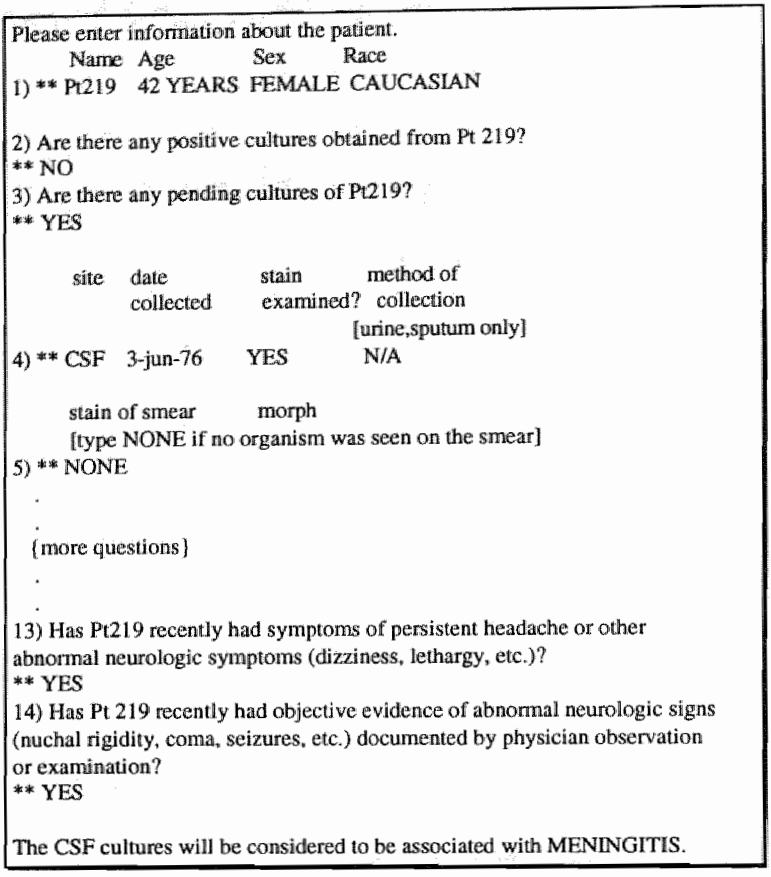


Figure 11. Excerpt from a session with MYCIN. From Duda and Shortliffe (198

this project and shift their attention to other tools (Buchanan and Shor 1984, 691-5).

The Expert System: Encoding Clinical Reasoning

The expert system is positioned at an interesting point vis-à-vis the t described above. On the one hand, in agreement with Feinstein cum s the expert system builders criticize the feasibility and desirability of st tical tools. The ideal of statistical inference, they argue, is unattainable:

> the failure of the use of the pure probabilistic decision making schemes lies in voracious demand for data ... For even a relatively small problem - e.g. 10 hyp ses and 5 binary tests - the analysis requires 63,300 conditional probabil (Szolovits and Pauker 1978)

Assuming, for example, that symptoms and signs of a disease are independent (i.e. that having symptom A does not alter your probability of having symptom B) cuts this number down considerably. But although this assumption is almost always taken for granted, it is "usually false." The probability that a person has a given symptom only too frequently *does* change with the outcomes of other tests. Many symptoms are interrelated through pathophysiological and anatomical mechanisms, Szolovits and Pauker argue; having a rash increases your chance of having fever, since "rash" and "fever" often go together in infectious conditions.[37]

Statistical tools create "artificial simplifications of the problem," Szolovits states:

> Attempts to extend these techniques to large medical domains in which multiple disorders may co-occur, temporal progressions of findings may offer important diagnostic clues, or partial effects of therapy can be used to guide further diagnostic reasoning, have not been successful. The typical language of probability and utility theory is not rich enough to discuss such issues, and its extension within the original spirit leads to untenably large decision problems. (1982, 7)[38]

The expert system builders, thus, join Feinstein in questioning the suitability of abstract statistical formulae for the complex, messy nature of medical practice. "Diagnosis needs to be only as precise as is required by the next decision to be taken by the doctor," Szolovits and Pauker argue (1978). Tool builders should take the pragmatic particularities of medical practice seriously: "the simple passage of time, 'creative indecision', often provides the best diagnostic clues because [it] adds a whole new dimension to the other available information" (ibid.).

Expert system builders join Feinstein's other major point of critique as well. The statistical nature of the tools, they agree, is often directly antagonistic to what clinicians try to achieve. Statistical tools have been unsuccessful because their reasoning is devoid of clinical meaning. Expert system builders share with the protocol makers we encountered the fundamental notion that *content* is primary: all usage of decontentualized statistical theories should be subordinated to clinical experience or pathophysiological knowledge.[39] This would improve both the performance of decision tools and their acceptability to physicians. With regard to statistical tools, Davis et al. argue, it is often

> not clear how each of the symptoms (or some combination of them) contributed to the conclusion. ... The problem, of course, is that statistical methods are not good models of the actual reasoning process ... [They are] "shallow," one-step techniques, which capture little of the ongoing process actually used by expert problem solvers in the domain. (1977)

Notwithstanding this alliance with the protocol, the expert system builders have fundamental problems with the set-up of that tool as well. Protocols, they sneer, "have been largely rejected by physicians as too simplistic for routine use" (Shortliffe 1987). The rigid, simplistic structure of the protocol causes it to break down as soon as a non-routine event occurs, so that "the difficult decisions are left to experts" (Shortliffe et al. 1979). When many factors play a role, when uncertainty comes into the picture, "then the rigidity of the flowchart makes it an inappropriate deci-

My therapy recommendations will be designed to treat for organisms that either are very likely or, although less likely, would have a significant effect on therapy selection if they were present. It is important to cover for the following probable infection(s) and associated organism(s):

INFECTION-1 is MENINGITIS
+ <ITEM-1> MYCOBACTERIUM-TB [from clinical evidence only]
+ ... COCCIDIOIDES [from clinical evidence only]
+ ... CRYPTOCOCCUS [from clinical evidence only]

43) Does Pt219 have a clinically significant allergic reaction to any antimicrobial agent?
** NO
[Considering ETHAMBUTAL-AND-INH-AND-RIFAMPIN for use against INFECTION-1...]
44) Is Pt219 pregnant or breast feeding?
** NO
45) Do you have reason to suspect that Pt219 may have impaired renal function?
** NO
46) Pt219's weight in kilograms (or <number> POUNDS):
** 51.3

[REC-1] My preferred therapy recommendation is as follows:
Give the following in combination:
1) ETHAMBUTAL
Dose: 1.28 g (13.0 100mg-tablets) q24h PO for 60 days
[calculated on basis of 25 mg/kg]
then 770 mg (7.5 100mg-tablets) q 24h PO
[calculated on basis of 15 mg/kg]
[the dose should be modified in renal failure. I can adjust the dose only if a creatinine clearance or a serum creatinine is obtained].
Comments: Periodic vision screening tests are recommended for optic neuritis.

2) INH
Dose: 513 mg (5.0 100mg-tablets) q24h PO
[calculated on basis of 10 mg/kg]

3) RIFAMPIN
Dose: 600 mg PO q24h
Comments: Administer dose on an empty stomach

Although I suspect that fungus may be a possible pathogen, in the absence of definitive evidence, I will not recommend antifungal therapy at this time. Please obtain samples for fungal, TB, and viral cultures, cytology, VDRL (blood and CSF), coccidioides complement-fixation (blood and CSF), cryptococcal antigen (blood and CSF), viral titers (blood and CSF). An infectious disease consult may be advisable.

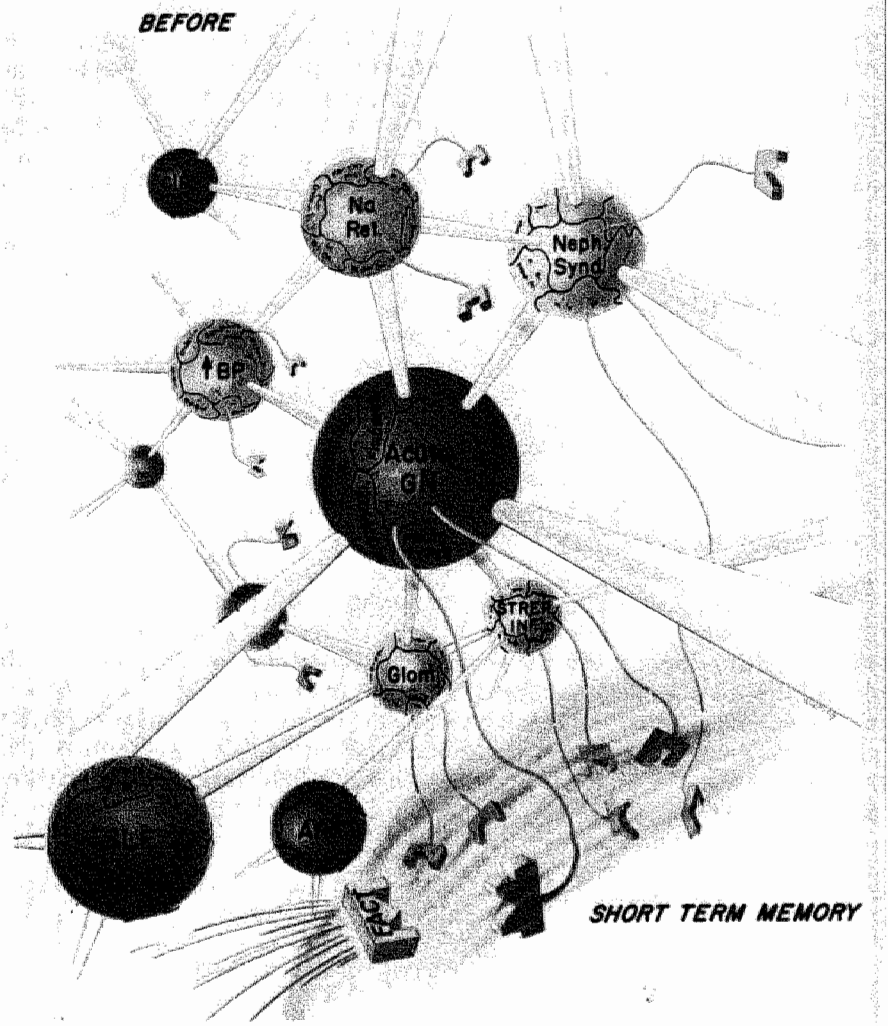
Figure 12. MYCIN's therapy advice. Note the additional questions and e tory remarks. From Duda and Shortliffe (1983).

sion making instrument" (Szolovits and Pauker 1978). Like the statistical tool builders, expert system designers feel that protocols put physicians in unyielding strait jackets. Protocols deny physicians the flexibility they require when problems become difficult - when, in other words, there is a need for a decision support tool in the first place.[40]

The expert system builders' position is thus juxtaposed to the other tools in an intriguing way. On the one hand, expert system designers share the view of Feinstein and others that medical practice is a non-quantitatively structured domain, typified by its pragmatic, judgmental, *clinical* character. Ameliorating the practice of medicine, then, requires building upon the concepts, theories and know-how physicians themselves employ - and not implementing decontentualized, "shallow statistical theories."

On the other hand, expert system builders join the proponents of statistical tools when it comes to the explicit focus on *decisions*. They position their tool in the flow of medical practice like statistical tool makers do. Protocols try to rationalize medical action by attempting to attain uniformity in diagnostic and therapeutic procedures. For expert system builders, on the contrary, rationality is in a fundamental way not the same as standardizing sequences of action. Here, rationality is flexible, intelligent *reasoning*. Expert system builders hope to optimize medical decisions without imposing restraints on more than just the decision itself. Given the clinical nature of medical practice, they want to outperform the physician's cognitive, decision making capabilities - which, after all, are limited (*cf.* Pauker et al. 1976; Szolovits 1982, xiii-xv).[41]

The development of the expert system and its Artificial Intelligence-predecessors, then, was inextricably tied to the emergence of the second main variant within the cognitivist discourses encountered in Chapter 1. With the coming of digital computers, the physician is described as an *information processor;* with the further development of expert systems, the physician is said to organize his/her knowledge in "scripts" or "frames" (*cf.* Figure 11). Again, medical practice is redescribed in the tool's image; again, the perfect supplement to the physician is the device (wo)man is modeled to in the first place.



Text of original caption (Pauker et al. 1976): *Hypothesis generation. BEFORE: in the nascent condition (when there are no hypotheses in short-term memory), tentacles (daemons) from some frames in long-term memory extend into the short-term memory where each constantly searches for a matching fact. AFTER: the matching of fact and daemon causes the movement of the full frame (in this case, acute glomerulonephritis) into short-term memory. As a secondary effect, frames immediately adjacent to the activated frame move closer to short-term memory and are able to place additional daemons therein. Note that, to avoid complexity, the daemons on many of the frames are not shown.*

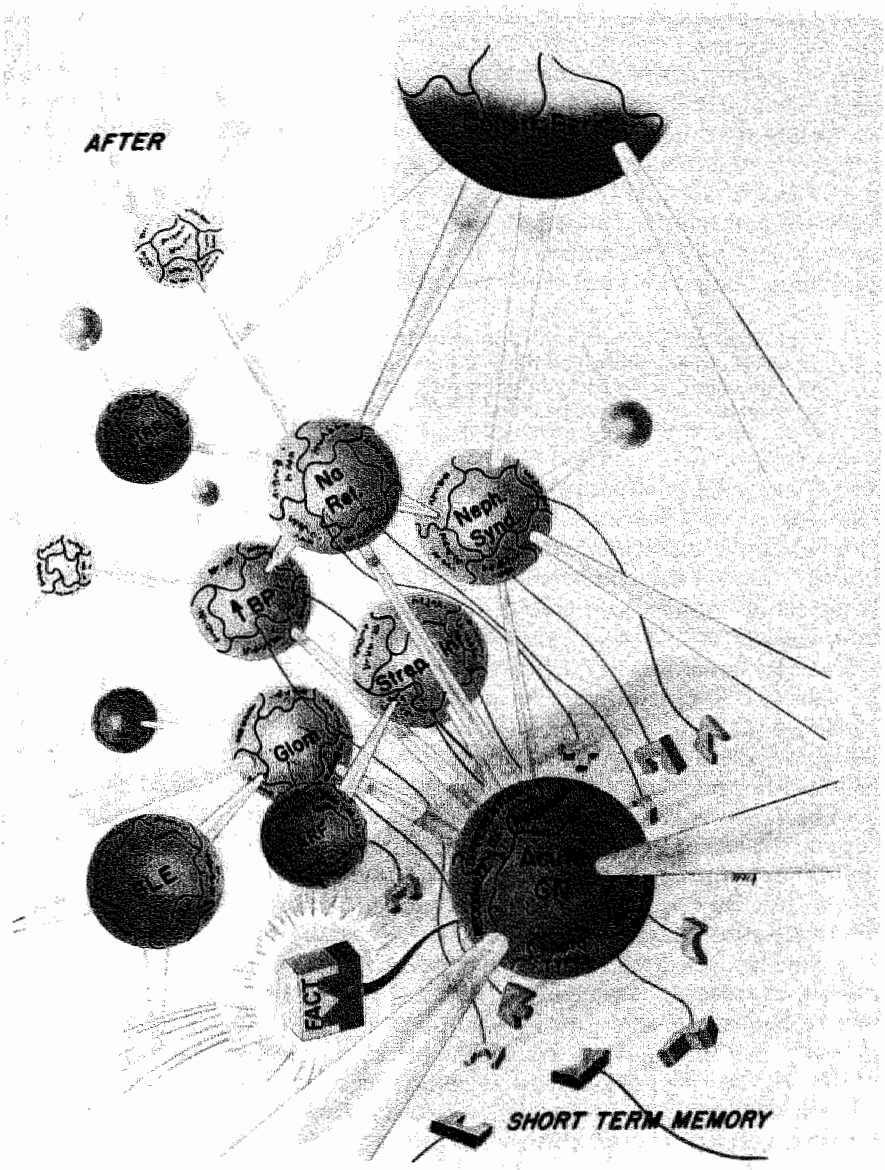


Figure 13. The continuous intertwinement of the construction of computer-based tools and new (descriptive *and* normative) models of medical practice closely follows the divergent developments of these tools. So, the information-processing notion that physicians organize their knowledge in cognitive "frames" or "scripts," which are "packages of closely related facts," has come up only with expert systems being programmed with such devices (*cf.* Pauker et al. 1976). This figure (ibid.) describes hypothesis generation of physicians: the intertwinement of models of reasoning, of the brain and of computers is prominent. The "frames" are portrayed as a kind of nervous cells, and terms like "short term memory" are used to indicate parts of the program.

Since both the expert system and the protocol as depicted above embrace a clinical logic, the statistical criticism of both expert systems and protocols ran along the same lines. As protocols, expert systems were attacked for being too superficial, eclectic, and overly pragmatic. Local experts' subjective opinions are just black boxed into a computer, critics argued. No attempt is made to *improve* upon the scientific nature of these ideas, or to make a system which is universally applicable from the outset. Modeling physicians' behavior in a computer also implies modeling his/her biases:

> There may be no assurance that information will be selected (e.g., tests ordered) in a cost-effective manner, and there may be only little assurance that gross errors will not be made... In modeling intuitive judgment to avoid the problems of formal decision analysis we might be trading a headache for a case of poor vision. (Politser 1981)

Many evaluations of the "first generation expert systems" acknowledged these criticisms. So, Clancey and Shortliffe state that unavoidably, "knowledge bases are incomplete, approximate, and biased models of the world." They "always reflect the values of their designers" since they "inevitably contain judgmental knowledge relating to social costs and benefits" (1984, 5).[42] Likewise, ad hoc tinkering and errors are inescapable since in systems like MYCIN "one must include all the necessary context for a rule's application in its antecedent clauses" (Reggia and Tuhrim 1985, 20). As a consequence, MYCIN lumps together "causal mechanisms, the taxonomic structure of the domain, and the problem solving strategies" (Buchanan and Shortliffe 1984, 396). From a statistical point of view, of course, this tinkering mingles many matters which should have been kept separately. As Shortliffe acknowledges, MYCIN's rules "were imprecise meldings of probabilistic and utility notions." Aggressive bacteria, for example, were given more weight so as not to miss them (Buchanan and Shortliffe 1984, 217; Shortliffe 1991).[43]

Several attempts have been made to design expert systems which meet this criticism. Some have tried to replace the "implicit value judgments" in their programs with explicit utility-values, and the "vague guesses" with numerical probabilities.[44] This development is reminiscent of the development of consensus reports: there as well, a clinical logic was "upgraded" according to statistical standards. Yet, as was the case there, statistical critics are probably not impressed. After all, the tool's backbone still consists of clinical, judgmental (and thus "vague," "imprecise") reasoning. If not, all the proposed advantages of the expert system, like its focus on content and its sensitivity to the pragmatics of clinical practice, would be abandoned. Adorning this basis with some well-meant gestures towards a "more fundamental" statistical rationality will not, in the view of the latter, lead very far.

Others have tried to delve deep into the "clinical logic" to find the scientific rock-bottom beneath it. The hope is to find a pure ground; a weighty, equal alternative to the clean world of quantitative inference. It is argued that what is needed is a "model of disease or clinical reasoning": "humans seem to exploit several different representations of the same phenomena"

(Duda and Shortliffe 1983; Szolovits et al. 1988). Likewise, then, programs should take into account epidemiological, pathophysiological, anatomical and clinical data and knowledge; and, as humans, they should be able to "reason from cause to effect." This can only be done through creating highly complex programs, in which for example "multiple pure hierarchies" (one for "anatomical site of involvement," one for etiology, and so forth) stand side by side and continually interact (Szolovits et al. 1988). The programming challenge, however, appears to be staggering; as of now, no system can be said to have accomplished this task.[45]

Rational Medicines

Decision support techniques are not the unequivocal "rationalizing" tools they might appear to be. They do not jointly underwrite a homogeneous view of Rational Medicine. Rather, the "rationality" of medical practice appears to have many faces. The decision support tools are associated with fundamentally different views of what medical practice looks like, how it should be transformed and what a rational medical practice is.

The differences between these tools do show some general trends, which can be laid out upon two main dimensions.[46] First, all decision support tools position themselves in one way or another to the ideal of *statistical inference*. According to this view, optimal decisions can be taken through the usage of a mathematical formula, free from both context and specific content. The ideal-typed statistical tools form the embodiment of this notion. These tools incarnate a thoroughly statistical view of medical practice, and want to improve medical practice by replacing the physician's judgmental and erratic decisions by solid, scientific, statistical reasoning. In contrast to the statistical ideal, the expert system and protocol builders discussed argue for tools which are grounded in *clinical* reality; which embed a clinical logic as opposed to a statistical one. For them, medical practice's raison d'être lies in its pragmatic, substantive nature - and this should be strengthened, not replaced.

The ideal of statistical inference, however, has become an epitome of Scientific Medicine. Through for example the institution of the clinical trial as the "Queen of Rational Therapeutics," the development of clinical science has become strongly intertwined with this ideal.[47] So, we see many tools which originated at the clinical side of this axis moving towards the statistical side. The development of consensus reports in the case of the protocol, and the development of expert systems dealing explicitly with probabilities and utilities: all these are attempts to bring these tools closer to the ideal of statistical inference. On the other hand, we also see movement away from this aim, as the case of decision analysis showed. Here it was the sheer impossibility of close adherence to this ideal which made clinical decision analysts willing to re-introduce "vagueness" and "subjectivity" - through the notions of subjective probabilities and utilities and the content-rich structure of the tree.

A second dimension cuts across the types of tools in a different way. Both statistical tools and expert systems, I have argued, are constituent features of the cognitive discourses described in Chapter 1. Both tools seek

to do the work that allegedly goes on in the physician's mind when (s)he makes a decision - and to do it better. In principle, then, both tools intervene only at the very moment of decision. The physician gathers all the data needed, and the tools process these data in a different space - in *their* brains, as it were. The tool generates output, after which the physician takes over again.

The protocol advocated by Feinstein and others, on the contrary, yields a different story. This tool does not primarily focus on a physician's *decision* but on the standardization of a sequence of actions. Feinstein and Weed never focused so much on individual decision making; as elaborated in the previous chapter, their notion of "scientific medical practice" resides not in the mind but in the structure of medical action. In a protocol, one decision can be involved, or more, or even none - that is not crucial. What is crucial is that the protocol does not intervene at any one point in time, but stretches out across a whole period. Also, since the protocol as described here does not model itself on the physician's brain (or vice versa), the "decision" is not made "elsewhere," inside some black box (whether literally - a computer, or more figuratively - some statistical formula). There is no realm where this tool momentarily retracts itself. The protocol grasps physicians by their shoulders and leads them here right, there left. Physicians may or may not understand why, but the path is laid out in front of them, on the piece of paper containing the recipe.

This dimension is also a source of mutual quarrels. Protocol builders are accused of promoting stifling bureaucracy, while decision-focused tool builders are criticized for overlooking the wider context in which these decisions take place. And, as on the previous axis, there is movement between the positions here as well. This movement, however, is more subtle, more elusive, and less easily witnessed when the focus is on texts describing tools. For this, we have to shift our perspective, and study the construction and implementation of actual decision support tools: the topic of the next two chapters.

A second main point of this chapter was that the diverse rationalities and images of medical practice have come into being *concurrently with* the development of specific tools. The statistical tools or expert systems were not called upon to fix some pre-given, long since recognized flaws in physicians' performances. Rather, these tools provided the metaphors for the working and failing of the physician's mind in the first place. Nor was the protocol "invented" as an answer to medical practice's problems. Feinstein's protocol was part and parcel of the standardization practices which had given medical practice its scientific footing; in bringing the protocol to the practice of medicine, notions of "medical practice" were transformed concurrently. The view of medical practice as a scientific process of distinctive, clear-cut steps is the inseparable counterpart of the notion of the protocol as an organizer of stepwise actions, as the fulfilment of medical practice's scientific character. There were no problems simply waiting for a solution: the development of the different tools was interwoven with the emergence of new rationalities and new views of medical practice. With the construction of the solutions, the specific shape of the problems was co-produced.

Notes

1. The empirical material used for this chapter consists mainly of journal articles dealing with the tools studied during the period 1945-1990. In contrast with Chapter 1, the editorials do not have a central position; nor do the *NEJM* and the *JAMA*. I have gathered articles using the "snow-ball-method." In addition, some of the key figures in the stories recounted here were interviewed.

2. Discourses cannot be disentangled from the (material) practices in which they have taken shape; vice versa, rationalities and views of medical practice can be inscribed in concrete tools. *Cf.* Edwards' notion of "discourse" (1994, MS50); *cf.* Akrich (1992).

3. See for instance Anon. (1956b) and MacMahon (1955). The reliance upon statistical methods to decide upon the validity of a hypothesis bypassed the clinical judgment of physicians. Löwy argues that clinicians were willing to surrender some of their judgmental power partly due to the increasing public visibility of medicine after the Second World War. More public funding of both medical care and medical research, she argues, meant more political vulnerability for the medical profession. Acquiring objectivity and universality of judgment through statistical methods was a means to strengthen the profession's position vis-à-vis potential critique of idiosyncratic wastage of public funds (1993). In his analysis of the history of the clinical trial, Marks qualifies the notion that statistics became so dominant a science for therapeutic research *because* it was seen as the route to medical truths. Rather, Marks argues, "to contemporaries, the improvements in experimental method offered by statisticians represented an elegant technical fix for a host of previously insoluble organizational and social problems" - as the problem to discipline individual physicians, the variations between methods and goals of involved groups, and so forth (1988; forthcoming).

4. See e.g. Warner et al. (1964), Gustafson et al. (1972), Van Way III et al. (1982). Contrary to de Dombal's tool, most of these other tools have never functioned in actual medical practice. Some used different statistical techniques as discriminant analysis, matching procedures, factor analysis, etcetera. Bayes' Theorem, however, was used most often (see note 5). These specific differences are not vital to the arguments made in this chapter; those notions prevailed which came to be linked up with successfully developed tools. For overviews, see e.g. Lusted (1968), Ledley (1965) and Jacquez (1964; 1972).

5. Bayes Theorem, it is argued, smoothly fits medical practice since it answers just the question a physician is looking for: what is the chance that this person, presenting with this and this symptom, has disease D (Ledley and Lusted 1959)? When we look at the probabilities figuring in the formula, these authors note that P(S|D) is just the relation between symptoms and disease given by medical knowledge: this relation can be found in medical textbooks. It "depends primarily on the physiological-pathological aspects of the disease complex itself." P(D), subsequently, is a factor incorporating local differences between populations. "This factor explains why a physician might tell a patient over the telephone [that a headache and fever probably indicate the flu, since it is around the community]. And the physician is more than likely right; he is using the [P(D)] factor in making the diagnosis" (ibid.).

6. See for this history e.g. Edwards (forthcoming, Chapter 6), Heims (1980, Chapter 12) and Mirowski (1992). Mirowski convincingly shows how the link with the military shaped the form of current Game Theory.

7. Early pioneers were Ledley and Lusted (1959); see also Lusted (1968; 1971) and Schwartz et al. (1973). A classic textbook in this field is Weinstein et al. (1980). The technique can be applied by the individual physician (computer-aids are available), or delivered by a "clinical decision consultation service" (Plante et al. 1986; Lau et al. 1983; Kassirer et al. 1987).

8. Some decision analysts and diagnostic tool makers would argue that they do not intend to say anything descriptive about the nature of the medical decision process. Most, however, do - either implicitly or often also explicitly. *Cf.* Bell et al. (1988) and Gigerenzer and Murray (1987, 150-62) for fundamental discussions of the not-so-clear line between descriptive and prescriptive models.

9. In the case of de Dombal's system, making the diagnosis *is* for a large part deciding upon therapy: when appendicitis is the most probable diagnosis, the patient should be operated upon; in most other diagnoses, less urgency is required. For an attempt to create a decision analytic extension of de Dombal's tool, see Clarke (1989).

10. Kassirer et al. (1987), *cf.* Politser (1981). See for a study of the so-called "framing effect" McNeil (1982). Some authors doubt whether "subjective probabilities" and "utilities" are completely independent quantities (Fischoff 1988). "Sensitivity analysis" is often evoked as the solution out of these problems. This is a method by which the input probabilities and utilities can be varied over broad ranges, so that one can see for which values the conclusions of the decision analysis hold. This added complexity, however, is only easily interpretable when the outcomes are clear-cut, i.e., when the conclusions of the analysis stay the same for a wide range of utilities and probabilities. More often than not, according to some, this is exactly the problem (*cf.* Doubilet and McNeil 1985).

11. In an off-hand demonstration somebody once gave me, the best policy came out to be "not having the disease at all"- a sympathetic but alas hard to follow advice.

12. In the eyes of some, this saturation with content explains the relative validity of many decision analyses. Shortliffe et al. argue that the usage of probabilities in a decision tree is less problematic than in a Bayesian, diagnostic tool: the "powerful knowledge structure" of the former compensates for the imprecise subjective probabilities. In a tool like de Dombal's, however, "the 'knowledge' ... lies in the conditional probabilities alone" (1979).

13. Of course, many "in between" tools exist. Subjective probabilities, for example, have been used for diagnostic tools as well: see e.g. Gustafson et al. (1972).

14. See e.g. Eddy (1990c), Field and Lohr (1990), Komaroff (1982), Komaroff et al. (1974). Emphasizing that protocols draw upon conditional (if... then...) rules as well, Lundsgaarde (1987) speaks of "noncomputerized expert systems."

15. It may be possible to trace links between the processes of standardization, Taylor's scientific management, and the emergence of protocols in the work of Feinstein and Weed. Taylorism strongly influenced the attempts to standardize hospitals at the beginning of the 20th century (*cf.* Stevens 1989), and Taylor's attention to flow-charts to represent workprocesses reappeared in the 1960s and 1970s in the work of computer scientists to represent the flow of information in organizations (Friedman and Comford 1989, Chapter 10). Feinstein comments that it is from computer science that he picked up the term "algorithm".

16. The first part of this title is derived from the title of Feinstein (1977).

17. Quoted in Vandenbroucke (1988).

18. Feinstein is by far the most outspoken and eloquent of these critics. Similar criticisms can be found in e.g. Cummins (1990), Margolis (1983), Ingelfinger (1975) and Rizzo (1993).

19. For example: the probabilities of being obese and having hypertension are not independent, since being obese increases your chances of having a high blood pressure. The independence assumption is only required when a simplified version of Bayes' Theorem is used (which is most often the case) (Cornfield 1972).

20. Horrocks et al., introducing the Leeds' acute abdominal pain system, themselves state that "unfortunately," their computer "is merely indicating that *if* the patient has one of the six or seven listed diseases in a 'database', then the probabilities are as stated" (1972). In other words, the physician (or other health worker) first has to make a judgment whether the patient belongs to the population of "potential appendicitis patients." Only *within* this population do the stored probabilities have any meaning. Moreover, the computer cannot diagnose a patient having a disease which is *not* one of the seven diseases listed - it will, then, erroneously (and in the same vein as in regular cases) mention one of its seven options as the most likely diagnosis.

21. Feinstein is somewhat more sympathetic to the basic idea of decision analysis, since it incorporates some of the pragmatism and clinical sensitivity he fights for.

22. *Cf.* Winickoff, McCue and Perlman (1977, 22-39) for a later, more extended version of this protocol.

23. Rheumatic fever is a potentially serious complication of streptococcal ear, nose and/or throat infections, sometimes resulting in progressive formation of scar tissue on (and concurrent deformation of) the cardiac valves.

24. Other contemporary examples of protocols for "physician-extenders" are e.g. Sox Jr. et al. (1973), Komaroff et al. (1974), Grimm (1975) and Strasser et al. (1979).

25. Feinstein derived this algorithm from D. F. Egan et al. 1970. What more can you do for your chronic lung patients? *Patient Care* 18-63.

26. Protocol builders often see no conflict between the design of protocols and the uniqueness of individual patients: through their branching logic, and through the leeway its rules allow, protocols can assure or allow for individuation of the steps to be taken.

27. See also Field and Lohr (1990, 36) and Eagle (1991).

28. See Mol and Berg (1994) for an analysis of "clinical," "statistical" and "pathophysiological" logics bearing family resemblance to the logics discussed here. See also Dodier's sociological analysis of the "clinical frame" (forthcoming).

29. The criticism had a powerful appeal. The influence of the statistical ideal of objective inference was and is expanding. In the booming field of quality assurance, for example, it was also increasingly felt that what proper medical action entails is not something that can be left in the hands of local physicians to decide. Objective, statistically processed "outcome figures" are needed to decide upon these issues (*cf.* Williamson 1973). See Gigerenzer et al. (1989) for acute observations of how statistics has come to "rule the world."

30. See for an overview and evaluation of these American consensus conferences Kanouse et al. (1989). More recently, the American Congress created the Agency for Health Care Policy and Research (AHCPR), with the development of "practice guidelines" as one of its functions (*cf.* Field and Lohr 1990 and Clinton 1992). On the European experience, see Vang (1988). In the Netherlands, "standards" for general practitioners are created by the NHG (Dutch General Practitioners Association), and guidelines directed more at specialist medicine are created by the CBO (Central Advisory Organ for Medical Audit). *Cf.* Grol (1989); Zwaard et al. (1989) and Everdingen (1988).

31. Eddy proposes such a combination of decision analytic techniques and protocols into what he calls "practice policies." Clinical problems should be analysed "in advance" using decision analytic techniques, and the preferred steps to take should be written down in a practice policy (1990a).

32. For similar criticism from within the medical profession, see May (1985), Rennie (1981) and Oliver (1985). The "consensus" approach is dismissed as "non-scientific" by Feinstein as well (interview November 11, 1993).

33. Quoted in Arkes and Hammond (1986, 3).

34. See Haugeland (1985, 176). For an analysis of the links between the US political discourse during World War II and the Cold War, and the origins of Artificial Intelligence (and the coming of the computer in general), see Edwards (forthcoming).

35. It should not be too explicated and clearcut: that would make the task of designing the system a trivial one (Davis et al. 1977).

36. INTERNIST, for example, was an attempt to create a system which could make diagnoses within the broad field of general internal medicine (*cf.* Schaffner 1985), and PIP (Present Illness Program) was to be a system which searched for the diagnosis in patients with edema through taking the history of the present illness (Pauker et al. 1976). See for a detailed overview Lipscombe (1991).

37. Szolovits and Pauker explain: "To the extent that anatomical and physiological mechanisms tie together many of the observations which we can make of the patient's condition and to the extent that our probabilistic models are incapable of capturing those ties, simplifications in the computational model will lead to errors of diagnosis" (1978). These authors also agree with Feinstein c.s. that in medical practice, the "basic premises of the applicability of Bayes' rule ... are often violated." Since many patients have more than one disease at a time, the requirement that the hypotheses are "mutually exclusive" is often not satisfied, and errors in diagnosis will abound (ibid.).

38. *Cf.* the critiques of Simon (1986) and Elstein et al. (1978, 290-1).

39. It was, after all, their explicit focus on the substantive core of a domain which separated the expert system builders from the earlier attempts to create artificial intelligence.

40. For some expert system builders, the potential flexibility of expert systems also surpasses the possibilities of statistical tools (*cf.* Pauker et al. 1976, Szolovits et al. 1988).

41. *Cf.* Simon's emphasis on rationality as "adaptive behavior"; as behavior which is demanded by the environment (Newell and Simon 1972, Chapter 3; Simon 1981). In his

turn, Feinstein distrusts the expert system approach, criticizing it as an "academic" at tempt to "bring outsiders' models in which do not work" (interview November 11, 1993).

42. For an early, similar critique, see Kulikowski (1977). For a harsh critique on th "unscientific" nature of the whole expert system approach, see White (1988).

43. Similarly, MYCIN used so-called "certainty factors" to "permit a conclusion to b drawn with varying degrees of belief." In the rule described in the text, for example, th conclusion talks about "suggestive evidence," which stood for a certainty factor of e.g 0.7. In combination with other rules, this allowed for "the accumulation of evidence" (Buchanan and Shortliffe 1984, 209). In their book, Buchanan and Shortliffe argue wh they found a more "formal" probabilistic theory like Bayes' Theorem both too limited an unfeasible, while also critiquing their own "groping effort to cope with the limitations o probability theory" (ibid., 209-62).

In addition, decision analysts criticize expert system builders for not paying explici attention to the patient's perspective. Clinical decision analysis, according to Pauker, i the only decision technique which offers a "hook" for the patient's preferences. "I don' know any other technique which allows shared decision making, that provides a forma process for combining the utilities, the values, the preferences of one individual with a formal structure and probabilities provided by another individual" (interview Pauke October 29, 1993).

44. See e.g. Langlotz (1989), Shortliffe (1991) and Schwartz et al. (1987). See e.g. Gorry e al. (1973) on an attempt to create a computer-based decision tool with a decision tre built in. In such a system, however, as with the attempt to base protocols on decisior analytic outcomes, the "underlying decision models generally have been prespecified Thus the program's usefulness is limited to those cases that correspond closely to th decision tree provided" (Shortliffe 1987).

45. For an attempt, see e.g. Patil et al. (1982). On the general trend to elucidate the intri cacies of the "clinical logic" further, see also de Vries and de Vries-Robbé (1985) and Reggia and Tuhrim (1985). See Chapter 4 for additional reactions to some of the per ceived problems of the first generation expert systems.

46. More dimensions could be drawn out: the different tools contain different images o "the patient"; they contain different images of the position of the physician vis-à-vis other health personnel, and so forth.

47. See also note 29. The term "Queen of Rationality" was coined by professor Vanden broucke (Leiden, the Netherlands; interview August 4, 1992). See on the history of the clinical trial Marks (forthcoming).

3

Getting a Tool to Work: Disciplining a Practice to a Formalism

Advocates of decision support tools often sketch an image of medical work in which the decision support tool smoothly fits. The tools do what physicians do - but better. The expert system is supposed to capture the knowledge of the "best minds available." Likewise, the statistical tool builders argue that their formulas contain the best route to decision making: the route physicians themselves would take under ideal conditions. The protocol, finally, is similarly depicted as a means of describing "good clinical reasoning." It is a vehicle through which the current standards of Good Medical Practice can be distributed: it merely explicates that what was already implicit in the practice it derives from (Grimm et al. 1975; Kanouse et al. 1989). Modestly repairing the course of action taking place anyway, the tools smoothly upgrade medical practice to its scientific status.

On the other hand, critics such as Dreyfus argue that decision tools *cannot* work. Such tools attempt the impossible. In the unruly reality of everyday medical practice, critics claim, achieving meaningful action through preset rules and formulas is just not feasible.

This chapter challenges both positions. I will demonstrate that getting a decision support tool to function in particular medical practices involves a thorough and specific transformation of these practices. From tool builders' literature and discussions in medical journals, the scene of action now shifts to the work of designers creating an actual tool and getting it to operate in a specific medical workplace (or series of workplaces). Contrary to the ideal-typed views of system builders, tools do not simply slip into their predestined space within a practice. Rather, getting a decision support tool to work, *constructing* a niche in a local medical practice, involves continuous *negotiations* with all the various elements which constitute this practice - nurses, physicians, patients, but also laboratory tests, blood cells and auscultatory sounds.[1]

This focus throws new light on the debate between critics and advocates about the reach of decision support tools - about which domains are formalizable, and which are not. I challenge the critics' emphasis of the impossibility of these techniques, and discuss some consequences of the negotiation processes described.

To include some of the diversity in decision support techniques, I center primarily on three tools. ACORN, a computer-based "chest pain advisor", is introduced in the following section. The second tool discussed,

de Dombal's acute abdominal pain system developed in Leeds (a statistical, computer-based tool), has already been introduced in the previous chapters. (Since this tool has not one name, I refer to it interchangeably as "de Dombal's tool" or "the acute abdominal pain system from Leeds" or some such combination.) The last example I draw upon is a research protocol.[2] This protocol treats patients with locally advanced breast cancer (that is, with involved local lymph nodes but no apparent distant metastases) with an experimental therapy called "peripheral stem cell transplantation."[3] These cases are compared with similar patients treated with "conventional" chemotherapy (see Figure 1 for an overview of this protocol).

I focus on a *research* protocol since these types of protocols are highly detailed and widely used in medical practices, which turns them into excellent objects for study. Also, research protocols are supposed to fulfil all the functions of a well-designed standardizing protocol (see Chapter 2): they articulate actions over time and place, give detailed recommendations

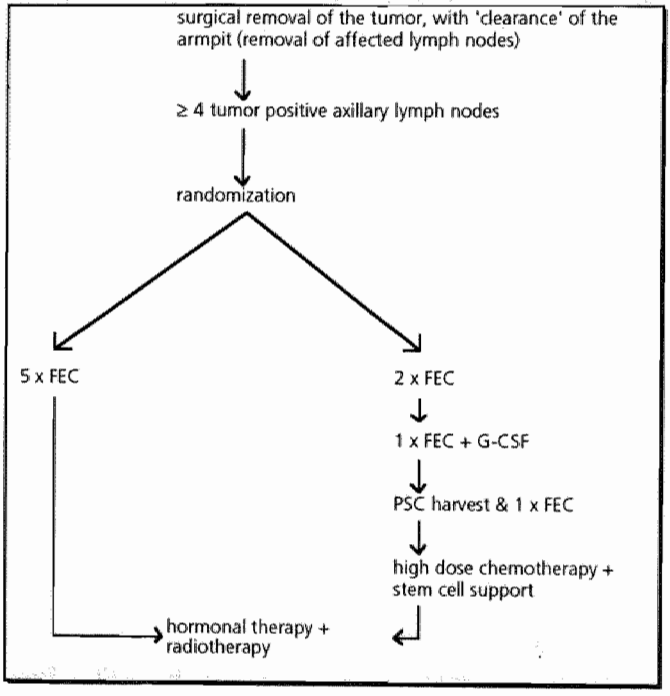


Figure 1. Overview of breast cancer protocol. FEC is the abbreviation for the combination of three chemotherapeutic agents used. PSC stands for "peripheral stem cell." The left arm is the "conventional" treatment; the right arm is the experimental therapy. In the latter, blood stem cells are filtered ("leukapheresis") from the patient's bloodstream and deepfrozen. Subsequently, the patient is treated with massive ("ablative") doses of chemotherapy. This would ordinarily kill the bone marrow cells, responsible for the production of the bloodcells. In this treatment, however, the unfrozen blood stem cells are given back to the patient after the highly toxic chemotherapy, which is hoped to affect the tumorcells equally deadly.

given specific situations, and transport (expectedly) optimal therapeutic regimes to an array of medical practices. Their widespread usage, moreover, seems to contradict the critics' repudiation of the feasibility of formal techniques - which makes this type of protocol an even more interesting case.

Disciplining a Practice to a Formalism

The Story of ACORN

Acute chest pain patients form a well known problem for UK emergency department personnel. Decisions about whether or not such a patient needs urgent admittance to a coronary care unit (CCU) must be made quickly and on limited information. CCU beds are scarce and expensive, so unnecessary admissions should be kept to a minimum. On the other hand, delaying a needed admission can be a matter of life and death. Acute myocardial infarct, dysrythmias and unstable angina all have a high early mortality rate. For these patients, the benefit of being treated in a CCU is greatest in the first few hours after an attack.

Delays and mistakes, however, are frequent phenomena in emergency departments. Patients are admitted unnecessarily, are erroneously discharged, or have to spend long periods of time waiting before a decision is taken. There appeared, Wyatt (1989) notes in retrospect, "to be a role for a decision-aid that could rapidly process relevant information about a chest pain patient, and help casualty staff to solve the urgent problem: should a patient be Admitted to the CCU OR Not?" Emerson, a chest physician working in an emergency department in London, had previous experience with decision analytic methods. He started work on a Bayesian tool (similar to the then already well-known acute abdominal pain system in Leeds) to help nurses make these decisions.[4]

The data items for this system had to be either examinations nurses could perform or questions the nurse could ask the patient.[5] Emerson set out to ask some of his experienced colleagues (all physicians) which items they deemed relevant. The three lists of items gathered in this way, however, were often in conflict with each other. There was widespread opinion about how significant different symptoms and signs were. "If an attempt had been made to build a system from the collected personal constructs of this group of experts," Wyatt and Emerson (1990) state, "the resulting repertory grid would have defied all known methods of analysis." Through a first rough selection as to their expected importance and ease of collection, this list was cut down to 54 items: 45 questions on the history and the characteristics of the pain and nine "simple" nursing observations about pulse rate and rhythm, temperature, blood pressure and the general appearance of the patient. In February 1984, a questionnaire containing these items was drawn up to collect data from presenting patients complaining of chest pain.

Using this list, a database of some 400 patients was created. Analysis of this database soon made it clear that many of the physicians' insights were not reliable. Compared with statistically determined powers of discrimination, physicians often overestimated the relevance of isolated signs and symptoms. They attached much weight to the type and the duration of pain, while the statistically measured predictive value of these items for a cardiac condition requiring admission to the CCU appeared to be low. Also, many data items suggested by physicians could not be used, since the nurses were not able to elicit them consistently. Many questions seen as relevant by the doctors failed this requirement: a question like "is the pain sharp in nature" was found to have a repeatability of only 66%.[6] Finally, the large number of items ran counter to the original goal of the system: having to ask 45 questions in addition to nine investigations was not very helpful in reducing delays!

To deal with these problems, the designers set out to curtail the item list. By eliminating those items which either had statistically low predictive values and/or could not reliably be collected, they managed to reduce the number of items to 22.

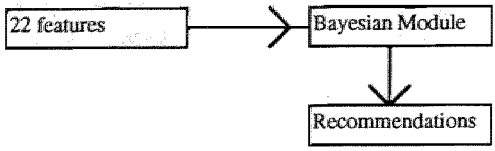


Figure 2. The structure of an early version of ACORN.

With these 22 items, a Bayesian formula (see Chapter 2) was devised. Figure 2 schematizes the setup of the system at this point.

Now, however, Emerson and his co-workers ran into another critical difficulty: the tool did not work. Using patient data from the database, it appeared that this "simple Bayesian approach failed to produce a viable decision aid, because over 30% of patients fell into a middle probability band between the 'Send home' and 'Admit to CCU' thresholds" (Hart and Wyatt 1989).[7] Moreover, ACORN would sometimes suggest clearly impermissible actions, such as sending a patient home who looked very ill and in shock. This could happen, for instance, when the patient would further have only a few positive items, thus yielding an overall Bayesian score below the threshold of "Admit to CCU."

The designers concluded that "a system based on probabilities alone was not sufficiently accurate and that additional symbolic rules were needed to complement the Bayesian analysis" (Emerson et al. 1988).[8] This hybrid system started out just like the first version: by processing a list of (now 12) indicators with Bayes Theorem. It decided whether the probability of the current case being at high cardiac risk was either low, middle, or high. Subsequently, a small expert system, containing some 200 rules, processed up to 12 further clinical features. This analysis would result in one of the following advices:

i) admit to the CCU immediately as a case of acute ischemic heart disease. No further investigation required.
ii) classification as a non-cardiac case not requiring an ECG (with a recommendation about whether or not to order a chest X-ray). [The ECG (electrocardiography) yields diagnostic information about the rhythm of the heart, and about whether e.g. the heart seems to be ischemic].
iii) do an ECG (of those patients of whom the computer had insufficient information to decide upon either i) or ii)).

When the last advice was given, the nurse performed an ECG, had the doctor interpret it, and went back to ACORN to enter this information as well. The structure of ACORN at this point is outlined in Figure 3.

This system was installed in the department and tested to see whether it would be of help to the nurses. To the disappointment of the system designers, three new,

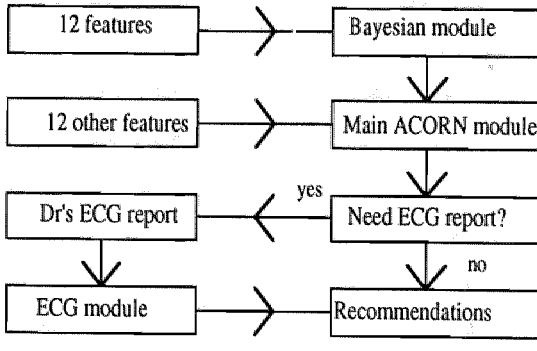


Figure 3. The structure of the "hybrid" ACORN (after Wyatt 1989).

serious problems emerged. First, ACORN performed poorly. It admitted and discharged patients wrongly far too often to be an acceptable decision aid. Also, the nurses often did not use the system immediately. When they finally entered the data, they had often already made the important decisions. Third, the nurses complained that the system was "too complex" and told them to do things they were not allowed to do: admitting a patient to the coronary care unit, they argued, is formally a doctor's job.

Confronted with these problems, the designers made some fundamental changes to the system. First of all, they decided that a main reason for all the problems mentioned was the fact that the program still tried to accomplish too complex a task. It was trying to do too much with a relatively limited set of data. For example, ACORN tried to figure out whether pneumonia of pulmonary infarct was likely to be the cause of the chest pain - in which case it would ask for an X-ray. As Wyatt and Emerson came to realize, however, they "could not get anywhere near accurate diagnosis of non-cardiac pain with the items used" (interview Emerson). They decided to limit themselves solely to the question of whether the patient could have a cardiac condition requiring acute admission to the CCU. A more general approach to the problem of acute chest pain, they felt, was not feasible.

Closely linked to this modification was the decision to stop trying to determine whether a patient needed an ECG or not. "It was more and more becoming practice to do an ECG on virtually anybody with chest pain," Emerson noticed (interview ibid.). Moreover, incorporating the ECG results from the very beginning would speed up the process which, as described above, now sometimes involved two separate sessions with the system.

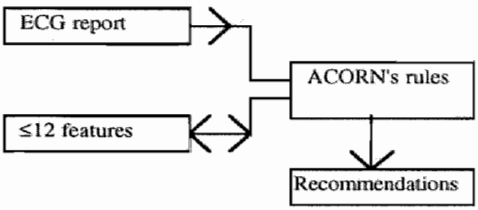


Figure 4. The structure of ACORN after the abandonment of its statistical part.

To make this change feasible, yet another main problem had to be tackled. ACORN's builders found that the delay and inaccuracy of ACORN's advice was primarily due to the fact that the ECG depended upon physicians. Since their reading was often both inaccurate and late, the designers decided to eliminate this dependency. Instead, an automated ECG interpreting machine was installed, which provided "more accurate ECG reports within seconds" (Wyatt and Emerson 1990). The nurse, now, had to take an ECG, answer ACORN's questions on the ECG, and subsequently answer up to 12 questions on the history, symptoms and signs. By now, Wyatt had left the team, and Emerson had completely abandoned the statistical part of their system; it now solely consisted of symbolic rules (the schematic outline of the new ACORN is shown in Figure 4).

Other changes were made as well. Since part of the delay was caused by the computer being too slow and located in a separate room, ACORN's ease of use and mobility was improved by installing the program on a fast laptop PC fixed on a trolley with wheels (see Figure 5). Finally, it was decided that the thirteen advice options originally provided by ACORN "did not fit in with the nurses' 'triage' system." Suggestions to perform additional tests and conclusions about possible diagnoses appeared to be too elaborate. They were reduced to three: "admit urgently," "see the doctor soon" and "wait in the queue to see the doctor." The designers furthermore arranged that a new policy was formally agreed upon: when ACORN advised the nurses to do so, they could admit a patient to the CCU without the intervention of doctors.

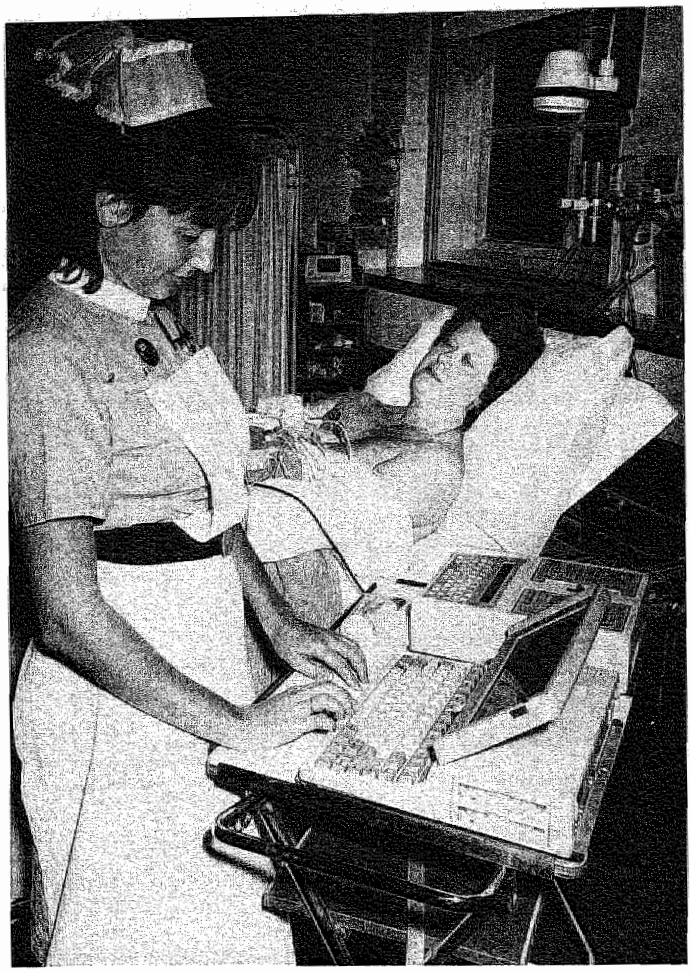


Figure 5. ACORN in use.

***

The introduction of a technology into an existing practice is a process of continual negotiations. Decision support techniques such as ACORN or the breast cancer protocol contain specific notions of what nurses, for example, are and are not allowed to do, and which patients' signs are relevant and which are not. In other words, decision support tools have an inscribed *script* delineating who or what the relevant elements in the involved practice are and what their respective roles consist of (Akrich 1992; Akrich and Latour 1992).[9] When a technology such as ACORN or a protocol is introduced, the originally inscribed script may be challenged by any of the heterogeneous elements it affects. In the negotiations which follow, these elements, including the technology itself, can all be transformed.

So, the *nurses*, at first, did not subscribe to ACORN's script: in the first field test, they often did not use the tool. The tool had to be modified so that it would be more mobile and gave less elaborate advice. Moreover, the designers arranged an official re-delegation of responsibilities between physicians, ACORN and nurses: now, when ACORN told them so, nurses no longer had to call upon physicians to admit a patient to the CCU.

The elements constituting this practice also include the *patient's chests*. These could also challenge ACORN's script. If ACORN would too frequently send healthy hearts to the CCU and return ischemic hearts home, it would be a failure. The chests, however, did not speak for themselves: a host of medical data were the elements which spoke for them. These at first made the creation of ACORN seemingly impossible. They were too numerous, they contradicted each other and often were unfit for use since the nurses could not gather them consistently (or in time, in case of an ECG). Here again, the designers negotiated with these elements in order to get some version of ACORN to work.[10] To make ACORN possible, the list of items first had to be cut down. In order not to lose the physicians in the process, ACORN kept the list of medical data cited by them as its central core. Subsequently, using statistical techniques, those data items which spoke most clearly and uniformly were selected - and at one point, all items dealing with non-cardiac diagnosis were thrown out. Finally, in order to keep the nurses' role feasible, those data items that spoke too indeterminately were thrown out, and the ECG-interpretation was delegated to a machine.

Negotiations were also part and parcel of the construction of the breast cancer protocol. In the following fragment, the setup of the protocol is discussed in an inter-institutional research meeting of oncologists. A first agreement on the desirability of this project has already been reached:

> Paula[11], who chairs the meeting, asks the group "how many lymph nodes" they want to settle upon. This number stands for the amount of axillary nodes in which, after the first surgical resection of the tumor, cancer cells were found. (On the whole, it is assumed that the more lymph nodes are involved, the poorer the prognosis will be. Settling on e.g. "ten," would imply, in this case, that only patients with ten or more involved lymph nodes would be eligible. The number chosen is dependent on several different considerations. One wants to reserve an experimental, intense treatment as this (with potentially lethal complications) for patients with a poor prognosis. On the other hand, the poorer the overall prognosis and condition of the patients, the less effect the therapy will have.)
>
> "If we set the limit to ten," one oncologist starts the discussion, "we might get too few patients. After all, the category of 'more than ten positive nodes' contains only some ten percent of our patients. We might have too few cases for our study." "How many patients are we talking about?" somebody else intervenes. Some numbers are mentioned, but nobody is really sure. "But we can ask the national registration office," Paula says. "Although I'm actually not even sure that our pathologists do not stop counting nodes at some point. They might not even reach ten; they might get bored beforehand, and just make a rough guess. I'm not sure." Some doctors laugh, and a fourth oncologist asks: "How many patients would have more than seven nodes?" Paula shakes her head: "the registration does not register that. 'Five' or 'ten' they will know. But that will be too much; we'll be flooded in patients. We cannot do more than so many of these treatments per year." "Not if we set the eligibility-age lower," the same oncologist counters. "That clears away much."
>
> Another oncologist joins the discussion: "if you would set the number of nodes lower than, say, ten, our center would get in trouble. We've already got some local trials running for these categories; this protocol would interfere." John, the main au-

thor of the protocol, interrupts: "That's no problem. You can enter all patients you want. We will deal with the differences in entry between centers. And about the expected number: we have to realize that some twenty percent will drop out, and that we've got a control group [receiving the conventional treatment], and that we need quite a number of patients to show the small difference in survival we expect. I would argue for going for, say, four nodes. Which is also a cutoff point often used in the scientific literature - where we want to be, right? If we settle on seven, for example, we cannot compare ourselves to other studies. Finally, if we have more eligible patients, we can finish our study sooner. This is good news both for our financiers and for our publication plans - and it also supports the choice for a low number of nodes."

This episode deals solely with the number of lymph nodes to be used as cutoff point. Already in this short episode, however, it becomes clear how the script of this protocol interweaves a wide, heterogeneous spectre of elements. The number of nodes, a pathological-anatomical criterion, was related to the statistical power of the trial, the workload of the centers, the habits of the registration office, the position of other, local protocols, the financial situation of the whole group, their alignment with the scientific literature, the fate of individual patients, and the (not explicitly discussed) question of which patients can be asked to "suffer" this therapy. All these issues had to be aligned with each other for the protocol to become feasible; all these issues were affected by the seemingly simple choice of four versus ten nodes. The centers would get their extra share of work, and they would have to renegotiate, within their own departments, how this protocol would interact with other, already existing protocols. Similarly, everybody had to agree to administer this treatment to the larger group of (relatively) less ill patients. If all went well, they would receive financial support, with which some centers could buy the equipment necessary to perform the peripheral stem cell transplantation. Also, they might succeed sooner, and more convincingly, in showing that this treatment would be beneficial.

In the construction and implementation of decision tools, then, the practices within which the techniques become embedded are thoroughly transformed. In introducing their tools in the emergency department and oncology wards, the system builders we encountered here did not restrict themselves to working solely on their tool. They redesigned the wards' practices as well. They tinkered with the wards' elements in order to accommodate them to the needs of the tool.[12]

These changes are not arbitrary: decision support techniques are *formal* tools. They operate using a collection of explicit, symbolic rules or statistical formulas, which turn input data into output. Given, amongst others, the number of lymph nodes (input), the breast cancer protocol, according to its (branching) logic, delivers the judgment on eligibility (output); somewhat further along its path, yet other rules determine the dosage of the chemotherapy (see Figure 6). As to ACORN, feeding in the requested patient features generates a piece of advice; in the Leeds' acute abdominal pain system, the input of some thirty data items results in a diagnosis.

Formal tools carry some specific requirements. First, they require a well-defined set of clear-cut, elementary bits of information as *input*. Whether the rules are statistical or symbolic, input items must at all times match a possible, discrete starting point contained in the rules. All the eligibility criteria of the breast cancer protocol, for example, are made exact

	day 21	day 28
WBC $\geq 3.0 \times 10^9/l$ and plat. $\geq 100 \times 10^9/l$	100%	100%
WBC $\geq 2.0 < 3.0 \times 10^9/l$	delay one week	75%
WBC $< 2.0 \times 10^9/l$	delay one week	off study
plat. $< 100 \times 10^9/l$	delay one week	off study

Figure 6. Table from protocol indicating dose reduction or delay in the FEC regime. FEC is administered in one day, once every 21 days. As its main side effect, FEC leads to a suppression of the blood cell production by the bone marrow. Given the number of white bloodcells (WBC) and platelets (plat.), the table determines what action to take. When, for example, at "day 21" (when a next dose should be given) the number of WBC is below 3.0×10^9 per liter (indicating poor recovery), the protocol instructs to "delay 1 week." At day 28, the white blood cells are measured again; given the results, the patient is either given the full dose of FEC, or a 25% reduced dosage, or sent "off study."

through laboratory tests and specific checklists: the cancer should be "histologically confirmed stage II A, II B or III A adenocarcinoma of the breast," the performance status should be "ECOG-ZUBROD 0 or 1" (appendixes of the protocol tell you what these codes mean), the "creatinine clearance" should be "≥ 60 ml/min." (a measure for renal function), and so on. There is no room for "the kidneys look all right," or "this patient is in-between stage IIA and B": the protocol has no rules to deal with such vague or undetermined statements. Similarly, the current version of ACORN asks the question "what is the site of the pain?" It then draws fourteen little puppets on the screen, each with a different shaded area over their chests. The nurse should select *one*. Actually, (s)he cannot do otherwise: the program will not proceed until (s)he has done so. ACORN also imposes a definite set of input items, and a definite sequence in which they are to be obtained. The ECG comes first, for example: without an ECG ACORN does nothing.

Second, the *output* of formal systems is pre-defined within the rules. Generally, a system contains a circumscribed set of pre-fixed statements, from which it selects one or several.[13] ACORN's final advice always consisted of one of at most thirteen possibilities. And often the protocol merely says either "yes" or "no"; sometimes, specific branches might make intermediate options possible (as when, in the case of a reduced white blood cell count, the FEC dosage is reduced by 25% (see Figure 6)). A seemingly innocent consequence of this feature is that at all times, all possible output must be feasible in the practice in which the tool functions. If, for example, one of the output statements requires the performance of a laboratory test not available in that particular setting, the decision tool is useless. Similarly, staff members need to have the time and skills to follow the tool through. Users have to be persuaded to act upon the advice - and to do so in the way intended by the formal tool (*cf.* Collins 1990). The hybrid version of ACORN, for example, sometimes offered the advice to obtain an ECG.

When the users of the system always already do that, however, the whole purpose of this advice becomes meaningless.

A formal system thus requires disciplined users to enter well-defined input, and expects the practice to be able to live up to preset output statements. In addition to these demands, a formal tool also requires a certain stability in the relations between the data items input and the content of the output statements. A data item like "the location of the pain" is only of use to ACORN or de Dombal's tool if it can be linked, through one or several (mathematical or symbolic) rules, to one of the output statements the system contains. If the location of the pain has no relation whatsoever with the cause of abdominal pain it is obviously senseless to include it in the Leeds' data item list.

So, a data item is useless if the existence of a relation is dependent on a host of contingent factors. For example, the breast cancer protocol sets the time for the leukapheresis (the "harvest" of blood stem cells) at the moment that "CD34 positive [blood] cells are detectable in the peripheral blood." (CD34 stands for a group of immunological markers which, according to the protocol's authors, constitute a standardized way of checking the presence of the blood stem cells: when a certain level of CD34 positive cells show up in the blood, the chances of gathering sufficient blood stem cells are high.) The amount of CD34 positive cells, however, can only be linked to leukapheresis in the breast cancer protocol if all institutions measure these cells with the same techniques, and if the different immunological markers used indeed function identically - something which was questioned by some of the oncologists. If not, all centers might be harvesting stem cells at different times - which might have all kinds of unanticipated effects on the success of the harvest and the outcomes of the treatment. In such a situation, an item like "CD34 positive cells" cannot be used by a formal system since it does not behave predictably enough. Similarly, to build a system like ACORN, the question whether the patient can enter the CCU cannot depend on too many circumstances. Whether the CCU is full or is expected to be full, whether the patient wants to be admitted, or whether the patient lives near another hospital where (s)he would preferably be sent to cannot be too important. If many such contingent considerations played a role, there would again be no way to link the input data items with the possible advices given as output.[14]

All in all, thus, the tools require many of the diverse elements constituting the medical practices to behave in a uniform, stable, and predictable way. A broad range of elements is affected: ensuring the proper execution of the CD34 test is just as much a problem of getting the laboratory staff to act in similar ways as a problem of ensuring that different immunological markers behave identically. Similarly, administering the FEC cycles requires both the patients and the hospital organization to adhere to this "rigorous three-weeks schedule" - and to be ready for a one-week delay when the blood cell count so indicates. It is not only the "nature of the clinical problem" which has to be well-defined, as computer-based system builders often argue when justifying the choice of a specific "domain" of interest.[15] It is not just a matter of finding a group of diseases with clear-cut symptoms and explicit criteria for deciding on diagnosis and therapy. *It is*

a whole, hybrid practice which must be made sufficiently docile - including nurses, physicians, data items and organizational routines.

Moreover, whether a practice is disciplined enough for the tool to work is not a pre-fixed, given fact. In the active "redescription" of the oncology and emergency practices, many heterogeneous elements needed to be *transformed* to make their behaviour definite, uniform and predictable enough for their protocol to work. A sufficiently disciplined practice is an *actively achieved accomplishment:* pathologists have to be instructed to count precisely, CD34 has to be measured in a similar way in all centers, and the medical personnel involved has to be taught to meticulously measure and document "side effects." In the case of ACORN, the ECG interpreting machine and the system's answer-options "digitized" input, and the department's organization was restructured to make ACORN's output feasible.[16]

Similarly, to ensure suitable input, the group from Leeds had to train physicians so that they would all mean the same thing when stating, for example, that the pain was "severe." Without adequate training, the observer variation was far too large for any decision tool to work with. When three untrained physician observers witnessed one clinician interviewing a patient, they "were unable to agree in 20.4% of circumstances as to whether or not a particular question was asked. [Moreover, they] were unable to agree in 16.4% of instances as to whether the patient's answer was positive or negative" (Gill et al. 1973). As de Dombal's team states, it took new staff "some six to eight weeks during which the personnel familiarised itself with the terminology" and during which "the system performed much poorer" (de Dombal and Gremy 1976, 155).[17]

In the implementation of a decision support technique, then, the involved practice is *disciplined to this formalism.* The networks of elements constituting the medical practices involved have to be made sufficiently "tight" for the tools to function (*cf.* Callon 1991). Instead of delimiting what decision support tools can or cannot do, thus, I argue that we have to focus on how domains are *made* "formalizable." Getting a tool to work is neither impossible nor a simple actualization of an already perfect match. The limits here are not preset. We are, rather, dealing with a *moving frontier*: a place of struggle to fulfil the prerequisites of formal tools.

Building Simple, Robust Worlds

> Writing a good protocol is not so difficult.
> What is hard is getting and keeping it in place.
> (Interview Dr. Howard, oncologist)

In disciplining practices to decision support techniques, several recurring patterns can be traced. In this section I point at three related ways through which elements are attempted to be made uniform, stable and predictable: reinforcing bureaucratic hierarchies, materializing the tool's demands, and shifting decision power to the most uniform and predictable elements. Together, these patterns emphasize the accomplished and heterogeneous work of getting a tool to function in an actual workplace.[18]

Reinforcing Bureaucratic Hierarchies

A first recurring pattern in the disciplining of practices to a formalism is the instalment or reinforcement of specific bureaucratic hierarchies (*cf.* Horstman forthcoming). In the case of the breast cancer protocol, eligible patients have to be registered at a national "trial office." Here the randomization takes place, and the name, age, diagnosis, and number of positive lymph nodes are registered. Also, it is checked whether "informed consent is obtained." The patient's therapy is thus determined at the national level; the physicians involved do not even get to do the randomization. Moreover, registration ensures that, from that moment on, the patient is "entered": the physician will now have to explicitly justify every non-prescribed action. For every "protocol-violation" the study-coordinator will have to be contacted.

Hierarchical relations can also be installed or reinforced *within* institutions (*cf.* Kling 1991). The following fragment from one of the centers participating in the breast cancer study illustrates how supervising relations among physicians are often enforced to ensure compliance:

> In a discussion between oncologists and residents, one of the residents asks whether they should not be given the freedom to modify chemotherapeutic dosages when side effects occur. This resident works at the oncological outpatient-clinic, where she is responsible for the administration of these medications. One of the oncologists disagrees: "I feel that we cannot give that responsibility to you. If we write a prescription, you have to be able to follow that blindly. Often some protocol is involved, which you might not be aware of; in these cases different rules apply for whether or not you can continue treatment in the light of side effects. So, I would say that you just call us when you're not sure. Don't go changing things on your own."

Finally, it is only *through* this rigorously hierarchical set up that many decision support tools are feasible at all. Both in the case of Greenfield's upper respiratory tract infection protocol as in the case of ACORN, supervising physicians have to be constantly available as a "back up." They have to deal with all patients the health assistant or nurses cannot deal with by themselves, and intervene whenever contingencies occur which are not foreseen by the tool. So, while here health assistants and nurses get to do tasks they were not allowed before, this is only possible through a thoroughly hierarchical anchoring of these newly gained responsibilities.

Materializing the Tool's Demands

Embedding the tool's exigencies in material arrangements as instruments and other artefacts is a subsequent recurring, often effective means of ensuring compliance. Having a tool's demands materialized prestructures the medical personnel's work environment so that the decision technique becomes an unavoidable (and often unnoticed) part of daily practice (*cf.* Fujimura 1988; Suchman 1993b). By requiring the usage of specific forms, for instance, medical personnel can be directly guided in the taking of a history, or the sequencing of a therapy. De Dombal's team introduced a specific form "in which the patient's case history could be to some extent

'formalized'" (Horrocks et al. 1972, see Chapter 2, Figure 1; see also Greenfield's protocol). Through this form, they forced investigating physicians to enter their findings in the required "structured and well-defined" way (de Dombal 1990). This "structured way" also implied another requirement secured through the form: all data items had to be entered at the same time. Physicians could not fool around with this demand: they had to hand over the form to a research assistant, who encoded the form and sent it to the main computer. Finally, de Dombal argued, to ensure ongoing commitment it should be made impossible *not* to use the form:

> The trick is to get the form used as a permanent part of the case record. That is absolutely crucial. If the young doctor [the emergency ward resident] has to fill out that form and then write it out again in the case record, they won't do it (interview July 18, 1993).

Forms are only one way of materializing a tool's demands. In the next fragment, a different material arrangement ensuring the strict performance of a protocol is already in place:

> When nurses change the "Hickmann catheter" [a catheter inserted in the subclavian artery, just below the collar bone, through which chemotherapy and bloodcells can be easily administered], they always use what they call the "Hickmann set." Usage of this set is prescribed by the protocol on Hickmann catheter-replacement. The "set" is a sterile, pre-assembled package containing the material required for changing the catheter according to the protocol, such as bandages, small trays for the disinfectant, tweezers, and so forth.

Similarly, the tools' output possibilities can be materialized. In research protocols, for example, medications can be pre-packaged and centrally distributed. In this way, participating institutions obtain the (expensive) drugs free of charge, in a form adjusted to the protocol at hand.

Materializing a tool's demand is often intertwined with the reinforcement of hierarchies: after all, hierarchies can be (and often are) materialized as well. Forms requiring signatures from superiors, pre-packaged chemotherapy which can only be modified through consultation with the study coordinator: all these materializations embed and sustain hierarchical relations. This intertwinement is forcefully illustrated by the protocol for "physician-assistants" designed by Sox Jr. and his team, where the assistants' actions were automatically monitored. To check compliance, the checklists filled in by the physician-assistant were screened by a computer for omissions and errors in following the protocol's branching logic (1973; Tompkins et al. 1973).

Reshuffling Spokesmanship: Shifting Decision Power Amongst the Elements

A third recurring pattern in the disciplining of a practice to a formalism concerns the decision power encoded in the tools: the input items the formalism weighs heavily in choosing a branch in the protocol, selecting an advice, or reaching a diagnostic statement. Here we see that tool builders have a preference for "trustworthy" elements: that is, those elements which, from the perspective of the tool builder, exhibit the most predictable

and unequivocal behavior. "Trustworthy" does *not* imply "better" or "more true," but points at the tool builders' tendency to delegate decision power to those signs, symptoms and tests which best, or most easily, fit the formal tool's prerequisites. As a result, spokesmanship is often redelegated from staff and patients to laboratory tests or machines. Rather than letting physicians decide when to harvest peripheral stem cells in the breast cancer protocol, CD34 positive cells settle the matter. The protocol builders deemed the criteria used by individual physicians to be overly idiosyncratic, and saw the number of CD34 positive cells as a test which would yield identical results in all centers. Rather than allowing "clinical judgment" to have its way, also, the number of lymph nodes and other quantitative, laboratory-derived values determine eligibility for this protocol. Similarly, rather than have physicians interpret ECG's for ACORN, the designers opted for an automated ECG interpreter, thus ensuring fast, unequivocal results. Whenever certain "non-human" elements behave predictably and uniformly enough, decision power is often taken out of the hands of the physicians and nurses (frequently deemed hard to discipline) and handed over to them. More often than not, in search for the most unequivocal and constant elements available, a laboratory test provides the more uniform, "digitized," and predictable behavior.

Likewise, patients' abilities to control the specific course of events is often limited. Patients can always refuse cooperation with a decision support system or step out of a protocol - but that is the only active role they can take. Besides that, they have little room for influence.[19] In the breast cancer protocol, the patient signs an informed consent form asking her to participate in the trial. She is told that "she has the right to withdraw cooperation at any moment." But in the tight network required for the protocol to work, there is no room for additional desires such as "a somewhat less intense second course of chemotherapy." If the laboratory tests do not register side effects, the second course will be as the first; and if they do, the dose modifications are already precisely prescribed. In ACORN and de Dombal's tool, similarly, patients can only answer preset questions about the location and duration of the pain. The preset output statements cannot deal with idiosyncratic patient desires or needs: they are geared towards a generic patient.

In addition to a shift away from patients' and medical personnel's voices and towards cells and chemical reactions, there is also a selection of the most trustworthy element *among* the physical signs and laboratory tests. Laboratory tests and physical signs are not more trustworthy per se: as many studies have shown, their robustness is itself the result of much work.[20] The CD34-test is debated among the oncologists exactly because its trustworthiness is debatable. Since the centers might use different techniques to measure these cells, an oncologist remarked, "center [A] might be looking at something quite different than [B]." Similarly, data items should speak in a clear voice: a large part of ACORN's story was concerned with the selection of those data items that would speak for the heart's condition most unequivocally and constantly. Those data elements were selected which both spoke clearly enough for the nurses to gather them, and were elected by statistical techniques as "truly representative." Auscultatory sounds, the classic voice of the heart from the physician's perspective, were

not included since nurses could not elicit them.[21] In an interview on the development of the Leeds' acute abdominal pain tool, de Dombal recounted a similar story:

> First we created a long list with items mentioned in the literature. Then we got rid of those items the majority of our clinical colleagues wouldn't do, or where they could not agree on the method of elicitation. The reproducibility of the item is important: we have thrown out typifications of the pain as "boring," "burning," "gnawing," "stabbing." They haven't gone because people don't use them - they've gone because people can't say what they are. We do observer variation studies on these things, and we ask people to define them. And if people can't agree what the definitions are, then we kick the items out. The observer variation on burning, stabbing and boring pain is about 50% - they're useless. Another example which fell off was backpain with straight leg raising: an often mentioned sign. But nobody agrees on what they are talking about. What should the result of the test be. A figure? The angle the leg makes with the table? But when should the angle be measured? When the pain starts or when it becomes bad? What angle to measure exactly? How to lift the leg? Or should the patient lift the leg himself? We could not get a group of rheumatologists, orthopedic surgeons, and general practitioners to agree about what they should call 'straight leg raising', so we abandoned that.[22]

Through a plethora of means, thus, practices are disciplined. Whether through reinforcing bureaucratic hierarchies, through materializing the tool's demands, or through delegating decision power to stable and uniform elements, potential obstinacy or unpredictability is averted.

Having said all this, it is important to avoid a specific misunderstanding. I am not claiming that the disciplining of a practice is a purely local, consciously planned endeavor of the tool builders and/or some local supporters. The process of getting a decision tool to work, of building a niche in a local network cannot be understood only from the perspective of those who attempt to implement the tool. Of crucial importance is the way these local networks fit into other, intermeshed practices. Oncology practices, partially through their historical tie with research protocols, are already heavily prestructured in ways congenial to new protocols. Medical personnel are used to these intervening instruments, data items or criteria are often taken over from earlier protocols, and central registration agencies are already in existence (as was alluded to in the excerpt). Moreover, research protocols often thrive on the existence of large institutions that aid in the proper execution of protocols, financially reward compliance, distribute centralized equipment, and so forth.[23] Similarly, practices gradually become more amenable (more "pre-disciplined" in the required fashion) for computer-based decision tools through developments such as the expansion of administrative bureaucracy, standardization efforts, the increasing usage of uniformly "packaged" technologies, increasing computerization of laboratories, and so forth.[24] These phenomena, including the crucial, active role of other participants in the involved practices, are discussed later in this book. In the following section, I look again at the transformations occurring in the "rationalized" practices, to discuss some consequences of the recurrent disciplining patterns I have pointed at.

Changing Practices: The Specific Exigencies of a Formal Tool

In the disciplining of a practice to a formalism, we have seen, there is a tendency for individual nurses, physicians and patients to loose direct influence on the course of events. Hierarchies are reinforced and material settings are rearranged so as to ensure compliance with the tool's demands. Physicians have to hand over the judgment whether a patient should enter a protocol to an interplay of laboratory values and lymph node counts. Similarly, bringing the patient's own voice back to either "yes" or "no" brings about the unequivocality needed for formal tools to work. In the new network, these actors are repositioned: inevitably, the requirement for stable and predictable elements predisposes the taming, or even the silencing of these potential sources of contingency (*cf.* Star 1989a).

These changes point to a fundamental challenge to the views of the tools' advocates. Decision support techniques are not the inert mediums they are proclaimed to be. They are not simply "carriers" of optimal knowledge, mirroring the reflection of a Good practice to wherever it is needed. Rather, the mirror transforms the reflection by merging it with the exigencies deriving from its own formal structure. In these transformations, *medical criteria are specifically altered as well.*

The age limit in the breast cancer protocol, for example, was set to 55 - a straightforward decision criterion for in- or exclusion. In a non-protocollized situation, such a strict cut-off point would not have been regarded as meaningful. As one oncologist remarked, it might have been better to differentiate between pre- and post menopausal women. "Biologically spoken, it's a different disease," he remarked. Determining the "menopausal status" (through measurements of the levels of sex hormones), however, is "tricky," as another oncologist stated. It is time-consuming, not always reliable, and it has none of the simplicity and clearness of simply setting an age limit. It would thus endanger the smoothness and tightness built up which is so crucial for the optimal functioning of the protocol.

Similarly, counting the number of lymph nodes is an easily performed test, which will yield similar results in different institutions - that is, if the pathologists do their counting properly. As a criterion for suitability for treatment, however, it is a rather rough and blunt means to distinguish between patients. In medical practices most such tests are rough and blunt: there are no simple ways to precisely determine how far cancer cells have spread. However, when not working according to a protocol, physicians will often try to construct some image of the extent of the spread through physical examination (can small tumors be felt elsewhere?), clues in the history (pain may indicate spread), different imaging techniques (are the bones affected?), and so forth. In this way, some more specific idea may be generated of whether and, if so, to what extent cancer cells have actually disseminated to distant organs. This plethora of methods, however, is much less clear-cut, much harder to predefine and much more difficult to standardize than counting the number of nodes.

Finally, deciding on admittance to the CCU through the analysis of an ECG and up to twelve clinical and historical features puts much weight on a highly delimited number of data items. Auscultatory data are omitted, as are questions about the situation at the patients' home, as is the X ray as a

potential source of additional information. Again, these possible additional data can and often will be used by a clinician to determine whether acute admission is required - but their richness is too indeterminate and unstructured to be tamed. To function, ACORN ignores such sources of information and leans heavily on those it does take into account. So, physicians often reckon with the social history of a patient in their decision to admit a patient or not. If a patient is panicky, and if his situation at home is unstable, they may decide on admitting a patient even though his situation might be not all that alarming. They might even quote a "social indication" as the reason for admittance (*cf.* Chapter 5). ACORN, however, has no room for input information such as this. In its turn, ACORN does take the prior history very seriously: when a patient has had a myocardial infarction before or is known to suffer from angina pectoris, ACORN is strongly directed to admitting the patient. As Emerson remarked, "if they have had it before, it takes a lot to get the computer on a different track" (interview July 21, 1993). Whereas physicians might say that a certain patient is merely anxious because of his previous heart attack and worries too much, ACORN would admit.

Again, these examples do not show that these tools are wrong or stupid. I am not arguing, here, that they miss a crucial point. Rather, these examples illustrate how formal tools cannot be conceived as inert carriers of some "good medical practice." Delegating tasks to a formal tool *transforms* the nature of those tasks. The introduction of a decision support tool generates a propensity to refocus medical criteria on those elements which behave in predictable and easily traceable ways. Formal tools contain a predisposition to build *simple, robust worlds*, without too many interdependencies or weak spots where contingencies can leak back in.[25] In doing so, in selecting the measurements and indications which best fit its prerequisites, the breast cancer protocol *redefines* what eligibility for bone marrow transplantation treatment denotes - what, thus, "potentially curable disseminated breast cancer" *is*.

Settling on a minimum of four lymph nodes as inclusion criterion selects a different group of patients than inclusion criteria based on extensive testing for dissemination would: the lymph nodes have become the salient sign determining the patient's fate. A patient with three positive lymph nodes but a large, fixed lump in the breast, for example, might be deemed a good candidate for this bone marrow treatment by some physicians - but the protocol would rule her out.

Likewise, opting for the CD34 test above the physician's judgment transforms what it means to be suitable for stem cell harvest: the two methods will yield different results in many patients. The age limit, also, carves out significantly different cut-off lines than a measurement of menopausal status: different criteria are invoked, which will create different groups of in- and excluded patients, and rewrite the significance of particular signs and symptoms.

Finally, focusing on the ECG and some additional clinical and historical features transforms the indications for admittance to the CCU. Social considerations, for one, can no longer play a role, while a previous history of angina or myocardial infarction gains importance. This again rewrites the significance of these signs and symptoms, and again carves out a dif-

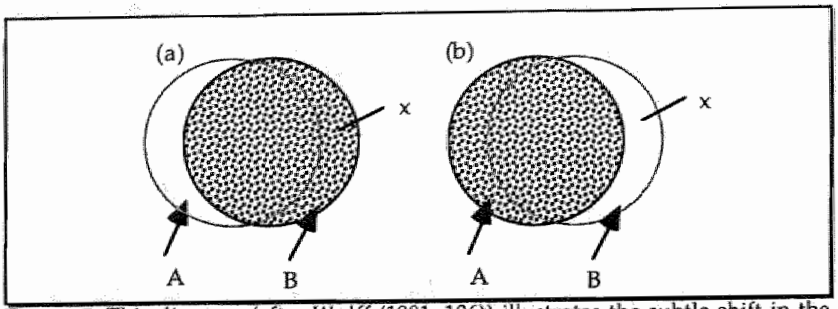


Figure 7. This diagram (after Wulff (1981, 106)) illustrates the subtle shift in the medical criteria for "treatable breast cancer" or "chest pain requiring admission to CCU." (a) is before and (b) after the implementation of the decision support tool. The shaded areas indicate the group of patients to which the diagnostic label is applied; the shift from A to B indicates the subtle, often unnoticed change in the use of this diagnostic label. Patient x would be treated (or admitted) in (a), but not in (b) - an effect which is wholly due to the specific characteristics of the tools in question.[26]

ferent dividing line between patients who are admitted and patients who are not (see Figure 7 for a graphical representation of this argument).

How should these redrawings be valued? In her study of decision making in neonatal intensive care units, Anspach comments critically on the development of decision support tools which utilize statistical formulas to predict an infant's prognosis:

> [These developments] may take place at a certain price and may result in the loss of certain potentially valuable sources of information. For only the measurable features of infants' conditions are [usable for the formulas] ... Precisely because many of the interactive cues nurses note are rarely entered into patients' charts systematically, they cannot be validated as prognostic signs. (1993, 79-80)

"Subjective" signs, she argues, will then loose out against an ever increasing emphasis on technological, "hard data."

The tendency to build simple and robust worlds, nevertheless, is not *necessarily* equivalent to relying solely on technological data items.[27] In the case of the breast cancer protocol, the age limit of 55 was preferred as a cut-off point over the measurement of hormones: here, a "technological" source of information was *not* used. Similarly, de Dombal's tool only uses data items achieved through a questionnaire and a simple physical examination: his team has refrained from using any of the additional technologies (as laboratory tests or X rays) available. As a final example, the designers of a research protocol I studied were not able to define precisely how to use "sufficient pulmonary function" (as defined by a set of technological, pulmonary tests) as an entry criterion without loosing either too many or too few potential patients. As a way out, they decided to leave this issue up to "the physicians' discretion." Sometimes, then, technologies are more unruly than humans; sometimes, "technological" data are less disciplined (or more costly to control) than "subjective" information.

More fundamentally, opposing "soft" to "hard" data, "subjective" to "technological" information, draws a dichotomy we should rather attempt

to overturn.[28] These are not two, wholly different types of data: rather, the difference is a gradual one. In fact, arguing that formal tools cannot deal with "soft" data is nothing but a tautology, since *what we call "hard" data are simply those data whose production already has been disciplined.* This disciplining, we have seen, is a highly heterogeneous affair; "hard" data are only robust since so much work has been put into stabilizing them. There is no intrinsic "softness" or "hardness" to data: what *is* measurable (and by what or whom) is the *outcome* of the negotiation processes involved in the construction and implementation of new diagnostic procedures, for example - or new decision support techniques.

The disciplining of a practice to a formalism, moreover, is not a process to be condemned in and of itself. The often drawn generalizing conclusion that formal tools are "dehumanizing" or "objectifying" misses the point. Through disciplining the nurses to ACORN, their responsibilities may be enhanced; through the implementation of this tool, physicians may be relieved of some duties they did not really cherish. Through their compliance to the breast cancer protocol, residents get to handle new, promising drugs they would otherwise not even have heard of, and some patients might willingly give up some right of voice in exchange for a potential straw of hope. The tools open up some worlds and close others; disciplining nurses, for example, might relieve them of a *prior* form of discipline (in the case of ACORN, for example, having to find a doctor for each and every admittance decision). On the other hand, situations may come up where Anspach's criticism should be taken at heart: when the disciplining of nurses, for example, is only intended to make someone else's work easier (see Chapter 6). Especially when the scope and depth of the transformations taking place are taken into account, global judgments as to which worlds are preferable are hard to make.[29]

Finally, the disciplining of a practice to a formalism is never complete. The technologies do not simply transform medical practice to their liking: in their construction and use, they are themselves transformed as well. The latter transformations are the topic of the next chapter.

Notes

1. Auscultatory sounds are sounds heard through a stethoscope. Following the work of Latour (1987), Callon (1991) and others, I conceptualize medical practices as networks of heterogeneous elements (physicians, patient files, X ray machines); elements which are interconnected in manifold ways, and which, taken as a whole, *are* the workplaces I focus on. The heterogeneous nature of these processes is a theme which recurs several times throughout the remainder of this book. I speak of "negotiations" with nurses and laboratory tests, and I mention physicians, ECG machines and auscultatory sounds as being the "spokespersons" for the heart. This terminology, probably odd to readers unfamiliar with recent studies of science and technology, is an attempt to keep a more symmetrical perspective. One of the central tenets of these studies is that the development of a technology cannot be properly understood from perspectives which treat Nature and Society as separate realms, and which confine explanatory power to either one of them. Just what the transformation of a practice (or of the tool) exactly comes down to can never be understood from standpoints which center merely on, for example, the nature of a medical problem or, conversely, the actions of medical personnel. Rather, the physical and social aspects of "computerized" or "protocollized" medical practices are the *outcome* of the

historically contingent events which have led to their current configuration. Only by treating them symmetrically can we hope to gain an understanding of the "success" or "failure" of decision support techniques; only by focusing on the way these heterogeneous networks take shape and break down can we come to terms with the fundamental issues at stake in the production and use of technical systems in (medical) practices.

2. I have not focused on clinical decision analysis since this technique is very rarely used in the form as described in Chapter 2 (see also Chapter 6). The stories I tell about the tools are based on articles their builders have written about them, observation of the tools, observation of the processes of tool-construction (in the case of the research protocol) and interviews and correspondences with the tool builders.

3. This therapy is a recent alternative for bone marrow transplantation, where bone marrow cells are harvested from the patient's pelvic bone. This new technique, it is claimed, is more convenient for the patient and results in a more rapid restoration of normal blood cell levels.

4. As both Wyatt and Emerson remarked, Emerson had been thinking of building a system like ACORN for some time already. The first few paragraphs of "ACORN's story" as depicted here, constitute an "origin myth" which could easily be told otherwise (it could be argued, as Wyatt did, that the program started out to reduce the number of tests done in the emergency department). The problem, however, is that *each* alternative could be told otherwise - such is the nature of origin myths. Since all the issues mentioned did, at some stage, play an important role, and since the arguments of the paper are not dependent on the specific origin myth chosen, I chose to keep the first paragraphs of the story more or less as Wyatt (1989) recounts it.

5. "Could," here, implies not just the ability to gather the item, but also the legal and professional issues involved. Nurses, for instance, could not be given the responsibility to judge an ECG (electrocardiography) report - something which, in practice, experienced nurses often do much better than junior physicians (see e.g. Hughes 1988). On the skill of ECG reading, see Hartland (1993b).

6. "Repeatability" was measured by the mean percentage agreement between a first and a second elicitation of the data item; in the case of a "yes/no" question, of course, 50% repeatability is expected by chance.

7. A test set of 133 patients had been kept apart from the set with which the Bayesian formula had been devised.

8. The rules were predominantly directed at "further reasoning about the ECG and about specific combinations of data," for example about those data which might indicate unstable angina. See Wyatt (1991a, 121-) for a more detailed description of ACORN's rulebase.

9. A technoscientific script is not simply equal to the prescriptions in the protocol. The interests at stake, the redistribution of costs, the research careers involved, the technologies which are skipped, the laboratory tests which are deemed more crucial than others: these issues are not explicated in the text, yet are of prime importance for an understanding of the development and use of the tool. *Cf.* Timmermans and Berg (forthcoming).

10. See Latour (1987) for the notion of "spokesperson;" see also note 1.

11. Names in this case-fragment are fictional.

12. Hence the term "heterogeneous engineering," coined by Law (1987).

13. Not all formal systems contain a prefixed list of statements as their outcome. Decision analysis does not: there, the outcome is a set of numbers, generated through the application of the decision analytic method on the input data (see e.g. Weinstein et al. 1980). The outcome is just as much predefined within the rules, however: it is *always* a similar set of numbers.

14. On similar issues in MYCIN, see e.g. Lipscombe (1991). And see Davis (1977) who, in discussing MYCIN's rulebase, states that a domain should have "a limited sort of interaction between conceptual primitives." If more than six clauses start appearing in the premises of a rule, the "premise becomes conceptually unwieldy." In other words, when too many factors can impinge upon a relation between a data item and, say, a given outcome, these relations become practically infeasible to hold on to. Similarly, he argues, it is needed that "the presence or absence of each of those factors can be established without adverse effect on the others"- otherwise the results will "depend on the order in which the evidence is collected."

15. Szolovits and Long (1982), see also Schwartz et al. (1987), Engle (1992), de Dombal (1990).

16. The term "digitization" is from Haugeland (1985, 52-8). Collins (1990, 22-9) takes up this notion as well. For a discussion of Collins' views, see Chapter 6, note 22.

17. See also Horrocks et al. (1976); Wilson et al. (1977). As with ACORN, it was an interwoven, heterogeneous assembly which needed disciplining. De Dombal's tool, for example, could not work with data items which took time to gather. Like ACORN, this tool dealt with emergency situations: data items could only be used if they could be collected and entered quickly. Unlike Emerson's team, however, de Dombal and his co-workers had physicians doing the investigation. In this way, more physical examination items could be used. Also, it appeared that with these data items, enough stable relations could be drawn to cover the great majority of acute abdominal pain patients coming to the emergency ward. Here again, it was only through training the physicians, introducing specifically developed forms and hiring research assistants that they were able to do so. Here again, the tool could only work due to the disciplining of a very diverse and entangled range of elements.

18. Parallels to the analysis given here can be found in Marks' work on the coming of the clinical trial (1988; forthcoming). The analysis draws strongly on the insights gained in the studies of the stabilization of scientific facts and technologies: see e.g. Latour (1988), Shapin and Schaffer (1985), Bijker and Law (1992). See also Zuboff (1988, Chapter 1) for a historical background to the depth and scope of disciplining practices (Zuboff focuses primarily on the disciplining of workers); see Woolgar (1991) for a study of how users are configured. Coming from an entirely different (cybernetic) background, Beniger's analysis of heterogeneous technologies of control yield interestingly similar conclusions as those drawn here (1986).

19. In the case of decision analytic techniques, of course, patients' preferences are central - but here as well the patient's role is highly prestructured (see Chapter 2 and 4).

20. See e.g. Fujimura (1992) and Latour and Woolgar (1986). See also Horstman on the construction of reliability in the case of the urine test in insurance medicine (forthcoming). A beautiful attempt to come to terms with the way matter is *not* regularly tamed, is Haraway's discussion of nature as "coyote/trickster" (1991, 197-201).

21. See note 5.

22. Interview July 18, 1993. Text is edited somewhat to increase readability.

23. Pharmaceutical industries, local, national and international research associations, and so forth. In fact, these links to interinstitutional and international protocols partially explain why (for example) oncology wards look so alike all over the Western world - much more alike, often, than different wards within one and the same hospital (Fujimura 1987; 1992). See further Chapter 6; *cf.* Timmermans and Berg (forthcoming).

24. *Cf.* Bowker's notion of infrastructure (1994b), and see also O'Connell's study on the creation of universality through the circulation of particulars (1993).

25. This simplicity, it should be clear, is a *consequence* of the work performed to achieve this robustness - what O'Connell, after Latour, calls "metrological practices" (1993).

26. Note that there is no definite way to ground either judgment. There is no way to determine what is the "more *truly* indicative" set of data items, for example. There is no solid, independent ground from which to make some ultimate judgment on a new selection of indicants. Hence, redelegating spokesmanship will result in subtle but thorough transformations in disease definitions and criteria. Lynch (1985) shows this point forcefully in his research of laboratory science. See also: Latour (1988), Pasveer (1992) and Hirschauer (1991).

27. Feinstein's life work could be summarized in this sentence (see e.g. his *Clinimetrics* 1987a).

28. See note 1 above.

29. On the impossibility of sweeping judgments about a related class of tools see Kling (1991).

4

Of Nodes, Nurses and Negotiations: The Localization of a Tool

"In light of these concerns this guideline seems crazy"[1]

Some years ago, Eddy worked on the development of a guideline on the use of contrast agents (used in radiology to visualize e.g. the stomach) for a large group practice health maintenance organization. Decision analytic techniques were central to his efforts. The issue was whether the benefits of using an expensive type of contrast agent (which caused fewer adverse reactions) would outweigh the money saved using the cheaper type. In the course of creating the guideline, several problems were encountered. It appeared difficult to figure out what the exact adverse reaction rates were. Eddy and his colleagues had already decided to divide the group of patients into two: those who had a high risk of adverse reactions[2] and those with a low risk. Now, "everyone knew that the risks were higher with ionic agents [the cheaper kind, MB], but there was uncertainty about how much higher." Analyzing the evidence in the literature appeared problematic: as is always the case, "virtually every study has something wrong with it." Study designs were imperfect, or did not match the group practice's local conditions. Moreover, hardly any study divided patients into low and high risk groups, requiring much decisions about how to interpret the data.

In addition, estimating the money saved was difficult. "To estimate the cost of a moderate reaction such as nausea and vomiting, we would have to conduct time-motion studies to determine how long an episode of vomiting ties up a radiological suite and the cost of janitorial service."

Comparing costs and benefits, subsequently, proved forbidding. Taking the decision analytic premises seriously, Eddy set out to question members of the health maintenance organizations about *their* preferences. Questions were prepared like "would you be willing to pay more for an X ray in order to reduce the chance of side effects." Slightly differently phrased questions, however, yielded completely inconsistent answers.[3] Members would say "no" to the question mentioned, but they would simultaneously be willing to pay slightly increased premiums to ensure getting the expensive agent whenever they would need it in the future. In addition, members would simply misunderstand the questions asked - however much effort was spent in making them clear-cut.

After much time and deliberations, Eddy and his colleagues ended up with a guideline recommending the usage of the cheap agent for low risk patients. Reflecting on the process, Eddy stated that this application of decision analytic techniques "should have been about as simple as any application ... ever could be": it was methodologically simple and politically desirable. Still, they "agonized over this guideline for more than a year." Being explicit about all benefits and costs, making the invisible visible, appeared incredibly complex - and often simply impossible. In addition, the radiologists remained highly reluctant to *not* use the expensive agents - no matter what Eddy's calculations said. They were the ones who saw the patients face-to-face; they were the ones who felt they were "withholding optimal care" from their patients. Actually, Eddy's calculations demonstrated that even for high risk patients, using the expensive agent as a standard would not yield

benefits justifying its costs. However, Eddy stated, "we had pushed rationalism about as far as we could on this problem. Words like 'high risk' have a psychological impact far beyond their actual numbers ... and we were approaching burn-out."

***

In the previous chapter, I outlined how in the processes of construction and implementation a practice is transformed after the image of the tool. In and through the continuous negotiations, however, the decision support technique is transformed as well: in getting a tool to work, it is inescapably *localized*. Inevitably, tools that function in a practice have had to give up much of the original, ideal-typed ideals about the power, range and/or transferability of the tool. Many tools appear to work only in one specific medical practice instead of in a universal range of practices; other tools end up confining themselves to a small part of the spectrum of medical problems they may have wanted to address. Localization, I argue, occurs in one or more of three (often intertwined) ways: in *space*, *scope*, and *rationale*. The phrase "the localization of a tool" points to the double meaning of the term: the tool's functional universality diminishes, and the tool becomes more and more particular. Avoiding localization implies increasing the disciplining of the involved practices to the formalism - but in the unruly worlds of medical practices, designers never succeed in exerting enough control. Localization, thus, is a widespread and inescapable phenomenon, continually thwarting attempts to spread the usage of these tools over broader terrains.

I first demonstrate the processes and different types of localization. Subsequently, I discuss how localization affects the advocates' claim that decision support techniques bring a unified, rational medical practice within reach. In the last section, the processes of disciplining and localization are reconsidered, and some central assumptions of the tools' advocates and critics are addressed.

Localization: "One starts with great expectations"[4]

Every protocol is a political compromise.
(Interview Dr. Bear, oncologist)

Localization in space is a process which already starts at the first conception of the technique and continues well after its introduction into practice. The script of the final ACORN can be read as reflecting the continual need to "give in" to local needs in order to get ACORN incorporated in the emergency department. ACORN uses rules and data items elicited from local experts (and then again only those items elicitable by local nurses, with locally available techniques, and speaking for local hearts[5]), its advice is adjusted to local customs, and so forth. Likewise, many protocols are developed by local (groups of) physicians, resulting in protocols reflecting local ideas about "what is best," local organizational configurations, local skills and local patient-characteristics. Both Eddy's guideline and the upper respiratory tract infection protocol of Greenfield et al. (1974) are such tools. As these authors themselves assert, their protocol "is derived from

local experience and [local, MB] peer consensus." When it would be used in different contexts, it would need "modification both to suit the needs of the individual physician and to accommodate changes in causal factors, local conditions, and therapeutic and laboratory advances" (ibid., *cf.* Komaroff et al. 1973).

A system can become so thoroughly imbued with specifically local idiosyncrasies that it can only work at that one site. This, in fact, is what has happened to most decision tools. Only very few expert systems or computer-based statistical tools have been able to deal with "the unresolved question of transferability of a successful system from its initial development site to other geographically distant locations" (Reggia and Tuhrim 1985, 35). The large majority of computer-based decision tools are in operation (if at all) in just one location. Similarly, in overviews of the current state of the art of protocol development, authors complain that the explosion of local initiatives has resulted in a chaotic multitude of often incompatible local protocols:

> Groups currently developing guidelines vary widely in the goals, methods, formats, and degrees of precision of their guidelines. Current guidelines from different groups contain important clinical differences on such common topics as screening for breast cancer. (Pearson 1992; *cf.* Audet et al. 1990)

However, some tools have explicitly attempted to overcome this confinement to individual sites. De Dombal's acute abdominal pain tool is one of the very few computer-based tools which has (partially) succeeded in doing so. Also, interinstitutional protocols have been construed (research protocols being a prominent example) and national institutions have developed "standards" or "consensus reports" (*cf.* Chapter 2) which should be used by every physician dealing with the specific clinical problem addressed. Focusing on such tools sheds light on what these efforts amount to.

A first phenomenon de Dombal and his co-workers continually ran into is that "databases don't travel" (interview de Dombal). Their system calculates the probability of a given patient having disease D by drawing upon a database listing past patients' symptoms and diagnoses (see p. 44-46). Every time the system was implemented in a different hospital, de Dombal stated, you needed to check whether the computer still performed as well as it did in Leeds - and usually it did not. When the system was tested in a Swedish hospital, for example, the computer's accuracy dropped from 80% to 50%.

The reasons for this phenomenon are manifold. For one, it often happened that a different type of patient was encountered. Sometimes the system was located in an emergency department, while elsewhere it was installed on the surgery ward. One difference between these locations is that in the former department patients can just walk in, while in the latter all patients have been referred by a physician. This preselection thoroughly affects the prior probabilities: the referring physician will have "stopped" many non-serious cases of acute abdominal pain from entering the ward. The chances of such a patient having acute appendicitis on a surgery ward, therefore, are much higher than in an emergency department. Also, de

Dombal and his co-workers encountered huge terminological differences among regions. Different groups of physicians appeared to use completely different definitions of "acute abdominal pain." Some included urological patients, while others did not.[6] Similarly, while introducing the system elsewhere, different forms had to be designed, symptoms had to be carefully defined and physicians had to be trained anew to ensure that in saying "the pain is colicky" they would mean the same thing as the Leeds doctors had been taught to do. Finally, the presentation of some diseases seemed to vary between places. Pancreatitis, for example, behaved differently in Copenhagen than in Leeds, which caused the Leeds program to falter.[7]

The non-transferability of databases, thus, is due to a complex, interrelated amalgam of differences in organization, terminology, disease-presentation, and so forth. Differences in opinion as to what constitutes "good medical practice" can also play a role: Shortliffe et al. (1979) argue that a system built by Bleich (1972) had never been in widespread use due to its "'Bleich-in-the-Box' feature." In their introduction to a volume on computer-assisted medical decision making, Reggia and Tuhrim sum up this thorny issue:

> This is more than the usual problem of physical portability of a program: it also involves resolving differences in medical definitions, varying standards of practice, and differences in patient populations. (1985, 35)

Non-transferable databases, however, are not the only obstacle to the implementation of a tool beyond a singular site. Other problems originate from differences in the way the tool is (to be) used. Consider the following remarks made by a physician about the difference between the usage of de Dombal's acute abdominal pain system in a surgical ward (Airedale) and in an emergency department (Leeds, where this physician worked). He points at the myriad detailed organizational features which all inhibit smooth transfer from one location to the next:

> The constraints of using the system are, compared to Airedale, not as strict here. [The form] is not part of the patient file, for example. They have to separately pick up the form here; they can very well do without it. These are different worlds, the ward at Airedale, and the emergency department here. There is no one standard way of implementing the system. At Airedale, physicians have ten to twenty minutes to talk with each patient. Plenty of time. Here, five minutes, and you've got to make a decision. It's a mess here. You can't find the form, what do you do? You've got to deal with the patients; it's so much more hectic in a place like this. (N., July 16, 1993)

According to this physician, this tool's script fits with the way Airedale's organizational routines are structured. Since work in the Leeds' emergency department is more disordered and performed under greater time-pressures, physicians are less willing to for example fill in the required thirty-some data items on the form. In this emergency department, the tool's script fits markedly less.

At the time I interviewed Emerson, no attempt had been made to export ACORN to another hospital, so whether this could be done remained

an open question. When asked about this possibility, Emerson sighed and remarked that the situation would be very different in other hospitals:

> If you would have someone who's enthusiastic for it, then maybe ... But it would be a lot of trouble. ... Just dropping it there wouldn't work. (interview July 21, 1993)

Wyatt (1991a, 116-7) lists some reasons why exportation might be problematic. ACORN, he argues, was for a large part directed at reducing delays, which might not be a problem in other hospitals. These might be better staffed than ACORN's emergency department in London, or they might have spaces in the emergency department where more intensive supervision and treatment of possible myocardial infarction patients is already possible. Furthermore, in other hospitals more experienced medical staff might be available to interpret ECGs, which might make ACORN's reliance on the automated ECG interpretation undesirable or unnecessary.

Similar issues affect the construction of protocols intended for usage in multiple sites. A protocol requiring a certain array of tests at a certain time will not work in a setting where one or more of these tests are not (routinely) available, or where the organizational arrangements are such that the protocol's time-schedule is unworkable. A protocol which requires a lumbar puncture (where, through a puncture in the lower back, spinal fluid is withdrawn) on "day 4" of the treatment scheme will generate problems in smaller hospitals where these punctures are not done routinely in weekends. Similarly, a demand to do "radiotherapy" at a certain, fixed time, might run into problems in hospitals which do not have their own radiotherapy facilities but need to refer these patients to other clinics.[8]

In the cases of interinstitutional protocols and tools like de Dombal's, then, the problem of transferability does not disappear: it is merely replaced. In these instances, decision support techniques limit their reach to those hospitals or departments that are alike enough to make common usage possible. The "FAC-scan" needed to perform the CD34-test in the breast cancer protocol, for instance, is an expensive instrument which is (at least in the Netherlands) only available in university hospitals and some of the larger general hospitals. Even less widely distributed is the technology to perform a peripheral stem cell transplantation. Finally, the diagnostic and therapeutic concepts contained in an interinstitutional protocol might be dismissed as "overly academic" by physicians working in rural areas, or, conversely, as "too low-grained" by physicians working in university hospitals.[9] The latter typically sneer at standards or consensus statements produced by national institutions as being for "ordinary hospitals." For them, *not* adhering to such protocols is what gives them extra status as first rank, university-based clinicians. Localization in space, here, is not so much confinement to a single site, but confinement to a limited series of local practices whose configurations are similar enough to allow for a single script.[10]

Constructing a feasible, workable decision support tool, to conclude, always implies that *specific contexts are built into the technique*. Inevitably, idiosyncratic, unique features of the specific sites involved become embedded in the tool's script; inescapably, in the construction and implementation of decision support tools, the tools' radius of action is reduced.

As the example of de Dombal's tool shows, however, combating this type of localization is possible - but it requires much additional work. Swedish physicians had to be trained to adapt the same terminology and investigational techniques as their Leeds colleagues, organizational setups had to be screened or altered, and emergency department physicians had to be enticed to use the system as their surgery ward colleagues did. Also, much larger, international databases have been gathered to ensure their validity across regions - albeit with varying success.[11] To overcome localization in space, additional disciplining of the involved medical practices to the formalism is needed. Stable relations between data items and outcomes now have to stretch out over regions. Physicians in different locales have to be trained to use the system and the involved terminology uniformly; organizational differences have to be flattened out.

Similar attempts are made to overcome the idiosyncracity of local protocols: Eddy tried to base his tool on international data (in which he succeeded only partially), and several institutions have developed "guidelines" on writing "guidelines."[12] It is through these processes that localization in space can be countered: only in and through this work of aligning organizations, terminologies, and procedures does a tool acquire its potential to "work everywhere."

On the other hand, "everywhere" is always "somewhere": the built-in context always locates a tool in (a) specific place(s). Research protocols embed the bureaucratic structure of modern hospitals: chemotherapy courses are always repeated in seven, not nine day intervals; infusions are given over 24 hours, never over 27. Also, when the required disciplining fails, localization in space is one way to cope with this failure. There were many instances where de Dombal's team did *not* succeed in adjusting the practices to the requirements of the widespread usage of their tool. In messy and complex practices such as hospital wards, control is never complete. Nurses, physicians, patients and hospital administrators all have their own tasks and agendas, to which the decision tool might be accommodated. The concerns of medical personnel for what *they* think should be done for the patient might overrule their alliance to the tool, or they might see the tool as imposing too many additional duties on them, and so forth (see the next chapter). Staff members' conformity to the tools' demands might be too hard to obtain - just like the controllability of data items and the patients' afflictions is sometimes beyond reach. In those cases where this unruliness could not be overcome, de Dombal and his coworkers went back to their system and tinkered with its setup so that these problems would not result in the total breakdown of the project. Whenever they did not manage to flatten out the organizational differences between different locations, for example, they just created different tools for each different site. Likewise, when the breast cancer protocol builders were not able to prevent some institutions from giving their own, local protocols higher priority, they just reduced the number of sites where the protocol would be used.[13]

Tools can also become *localized in scope*. This occurs when the field of application of a system is limited to an increasingly small subset of tasks. This may mean, for example, that the range of disorders the system can handle is reduced, or that the scope of actions it can undertake is limited.

In ACORN's practice, the designers did not succeed in incorporating non-cardiac causes of chest pain. Also, the tool had to be made "less intelligent": the thirteen detailed advice options were cut down to three. Here again, ACORN is quite typical of the overall category of computer-based decision tools. Within limited fields, tool builders often say, their tools perform well. "Skilled behavior" is seen in "many relatively well-constrained domains." However, they then continue, most programs exhibit the "plateau and cliff effect": "the program is outstanding on the core set of anticipated applications, but degrades rather ungracefully for problems just outside its domain of coverage" (Szolovits 1982, 15).[14] Increasing the scope leads to problems: unforeseen interactions between rules appear, the system looses precision, or fails to function altogether. De Dombal's team repeatedly tried to enlarge the scope of their system beyond the small domain of acute abdominal pain: they attempted to include, for instance, gynaecological disorders, dyspepsia or Crohn's disease. None of these attempts were very successful: the system would always perform much better with the added categories left out.[15]

Like localization in space, localization in scope is a result of failing control. It is a matter of unruly, heterogeneous elements which cannot be brought in line with the system. It is never "just" a *"technical"* matter of the "broad and complex" nature of clinical problems, as decision support tool builders often state (Szolovits et al. 1988; *cf.* Forsythe, 1993b). It is not that ACORN failed in the realm of non-cardiac causes of chest pain because of the "inherent difficulty" of that problem. The situation is both more complex and more mundane. If ACORN's designers had succeeded in arranging nurses to elicit additional data-items, as e.g. an X ray, blood tests or auscultatory sounds, it might have been possible to create a set of data items which *could* be used to distinguish between several pulmonary afflictions. This was not feasible, however: the physicians would not let nurses do these tasks, and the nurses themselves did not want them much either. Also, it would have taken additional time to gather these items, and it would have generated new problems such as how to turn X ray interpretations (which are notoriously vague) into clearcut, unequivocal pieces of input.[16] All in all, the failure to increase the scope of a tool is thus as much a failure of *social* control as it is a problem of the intractability of the relations between data items and disease categories. Including pulmonary diseases in ACORN would have required a stronger grip on nurses, physicians, the interpretation of X rays, and the links between test data and their implications, than was feasible in this place, at this time. In this sense, ACORN's localization in scope was not so much due to a mysterious "inherent difficulty" in the nature of chest pain, as to a *practical* failure to adequately constrain the closely intertwined, heterogeneous elements of this emergency department.[17]

Similarly, the breast cancer protocol authors first included all patients with surgically removed breast cancer - whether they had undergone a "radical mastectomy" (removal of the whole breast) or a "breast conserving procedure" (where the tumor is excised but the breast is not removed). The radiologists of one of the participating centers, however, did not agree with this. These physicians felt that patients who had had the breast conserving procedure required immediate radiotherapy to kill tumor cells possibly

remaining in the breast. In the proposed protocol, radiotherapy would come at the end, being delayed some 16-18 weeks (see Chapter 3, Figure 1). Although not all radiotherapists felt that this radiation of the breast was so urgent, the authors of the protocol felt obliged to give in in order to retain the cooperation of these radiologists. Consequently, in the next version, the protocol now excluded patients treated with a breast conserving procedure.[18]

The third and last type of localization I want to discuss is *localization in rationale*. This occurs when designers must agree to a setup in which the potential, ideal-typed precision, accuracy or "scientific quality" of the system is not realized. As in the previous types of localization, localization in rationale inevitably occurs as a result of the ongoing negotiations during the development and implementation of a tool. Faltering attempts to adequately discipline parts of the practice may be managed in this way as well. This comes out clearly in the account of Eddy's guideline: in order to keep the guideline feasible, he had to swallow so many compromises that the guideline "seemed crazy." As one physician put it, protocols are always "a political compromise." They will always have to be tinkered with to be not too different from what physicians "are already doing," and from what is organizationally "doable." Wiersma (1992) describes how the "working group" responsible for the Dutch general physicians' consensus report on serum cholesterol tests and treatment continually struggled with these issues. The working group, for instance, found that for women, high serum cholesterol levels are much less strongly correlated with coronary heart disease than is the case for men. But, Wiersma says, having "different guidelines for screening women was felt to be both socially unacceptable and impossible for the practitioners to advocate to their patients" (ibid.). Finally, Komaroff et al. (1973) argue that developing a protocol also means deciding which patients the protocol "might be expected to miss." In the case of the upper respiratory tract infection protocol, they state, it would have been possible, in principle, to check every patient for a-symptomatic otitis media (inflammation of the middle ear). Doing so, however, would have required looking in all patients' ears - an activity usually considered to be too difficult for physician-extenders to perform. Wanting to find these patients would thus have required dropping the whole idea of the physician-extender, since the physicians would have to see each patient anyway. As a compromise, Greenfield's protocol simply asks the patient about ear ache. By missing the a-symptomatic otitis media cases, by reducing some of the protocol's potential precision and accuracy, they saved their project.

Similar phenomena occurred in the case of ACORN. When Emerson started out building his tool he was "a devoted Bayesian." He was interested in rationalizing medicine; in the "mathematical approach to decision making" (interview ibid.). Already when selecting the items, however, Emerson and his co-workers had to give up some of this ideal. Potentially powerful data items had to be left out since the nurses either could not or would not elicit them. This "corruption" of the Bayesian ideal grew worse when they had to incorporate all kinds of symbolic, ad hoc rules to prevent impermissible actions like sending patients home who looked gravely ill. Finally, Emerson and his co-workers completely abandoned the statistical

ideal and embraced a *clinical* logic (*cf.* Chapter 2): ACORN's final version is a rule-based expert system in the classical sense. Judged from this clinical ideal, however, the tool's rationale was limited as well: too many data items with potential clinical value had to be left out to keep the tool workable.

The ideal of statistical inference was also compromised from the outset in the case of de Dombal's tool. Its limited scope of six or seven diseases "unfortunately" led to the requirement that the physician should make the first selection: is this patient suitable for the tool at all? Does this patient belong to the category of "patients possibly suffering from acute appendicitis" (Horrocks et al. 1972)? This necessity for the subjective judgment of the physician is, of course, a pollution of the purity of the ideal-typed statistical logic. Also, over the years, de Dombal's tool has evolved towards a rather different philosophy. In recent articles, it is continually stressed that the tool is *not* a decision maker. It is "just a test," or even only an educational tool; it yields nothing more than an indication as to what might be the case for a given patient (de Dombal 1991; de Dombal et al. 1991). The decision itself is returned completely to the hands of the physician. From the perspective of the original Bayesian ideal, this is a complete capitulation: an invitation to let the painfully fought arbitrariness back in.

A similar movement can be said to characterize the whole domain of rule-based expert systems. Many authors are abandoning the traditional Artificial Intelligence approach in favor of building tools which function more like electronic textbooks. The time of the "Greek Oracle model," these authors argue, is over. No longer should we work on computers employing "seemingly superhuman reasoning capabilities to solve the physician's diagnostic problem" - while the latter would be merely a "passive observer."[19] Here as well, the tool has stepped back from the claim that it outperforms the physician at the moment of decision. Here as well, the original ideal of mimicking and even ameliorating expert reasoning has been thoroughly watered down.[20]

The localization of a tool is inextricably tied up with the disciplining of a practice to that formalism. In the process of getting a decision support tool to work, both phenomena inevitably occur. The medical practice rewrites the tool's script and vice versa; the final result is a configuration in which both have transformed each other to their own liking. The different types of localization (in space, scope and rationale) can be fought through trying to strengthen control, through tightening the network of nurses, patients' bodies, physicians, ECG-machines, and so forth. As said before, the limits on the "formalizability" of a medical domain are not a preset given: they constitute a moving frontier, a place of struggle for control. Moving the frontier involves an exponentially increasing amount of work: Eddy and his colleagues *could* have fought the corruption seeping into their guideline by *doing* the time-and-motion studies to estimate costs, or by performing trials to determine the risks for adverse reactions themselves. On the other hand, difficulties to adequately constrain the heterogeneous constituents of medical practices can be coped with by localizing the tool. In a practice disciplined to a formalism, what is "legal" and what is "illegal" behavior is clearly distinguished. Only some movements are allowed by the script;

as Bowers (1992) phrases it, only "some orders or compositions can be realised or recognized in a particular formalism." When resistances occur, when "illegal" conduct (whether of people or e.g. laboratory tests) is too frequent and unamenable to correction, the tool can be tinkered with to save it from failure. By incorporating "craziness" into the tool (the final guideline was only partly in agreement with the decision analytic calculations), Eddy kept himself from burn-out. By reducing the scope or the rationale of the tool, or by accommodating local idiosyncrasies in the tool's design, a project can survive.

The different types of localization are not meant as an exhaustive set of mutually exclusive categories. Nevertheless, they allow us to view the development of both individual and classes of decision support techniques as attempts to avoid some type of localization. These attempts frequently result in another type of localization seeping in through the back door. More often than not, for example, the widespread distribution of a protocol is "bought" by reducing its specificity. As Bowker and Star have argued for the case of disease classifications, "standardization-procedures must be tailored to the degree of granularity that can be realistically achieved" (1994).[21] A certain, necessary "uncertainty and ambiguity" (ibid.) is needed - and more so where more different locales or different specialities have to be included within one, "universal" protocol. And, indeed, authors involved in the making of nation-wide consensus statements or "standards" speak about the need for "sufficient space for individual variation" which has to be built into such tools (Casparie and Everdingen 1989, *cf.* Kanouse et al. 1989 and Chapter 2, p. 62-63). Only by reducing the script's tightness can such protocols ever hope to achieve widespread use. Similarly, the oncologists working on the breast cancer protocol decided in an early stage to give recommendations for the radiotherapeutic treatment - but to include the statement that "modifications of the radiotherapy according to local views are permitted." They succumbed to their conviction that they would never be able to get the radiologists in the different centers to agree on one set of prescriptions.

These examples illustrate how a more widespread usage is achieved through inscribing vagueness into the protocol; how localization in space is fought through increasing localization in rationale. Vice versa, ACORN's designers, in statistically pruning the list of items generated by the physicians, can be said to have traded localization of its rationale with localization in space. In order to increase the predictive power of the data items used, the designers utilized a statistical technique which took into account only a very specific population: patients with chest pain coming to this specific emergency department. So, while they increased the accuracy of their tool, they simultaneously tied it more tightly to the local context of the specific emergency department they were dealing with.

Changing Practices: Decreasing Diversity?

In Chapter 3, I argued that disciplining a medical practice to a formalism tends to transform medical criteria in distinctive ways. The predisposition to build simple, robust worlds results in their association with a limited set

of simple, clear-cut variables. In addition, medical criteria also change in yet another way. To bring this into view, I take another look at the oncologists' discussion of the breast cancer protocol presented in the previous chapter (p. 85).

At the end of their deliberations, the oncologists settled on four positive nodes. To reduce the large number of eligible patients this criterion would yield, they cut down the upper age limit from 60 to 55. Also, the originating center had set a maximum time of 14 days between the surgical removal of the tumor and the start of chemotherapy; a trade-off between the time required for the patient's recovery of surgery and the need to prevent potentially remaining cancer cells from proliferating. Other centers, however, objected to this, arguing that this time span would be too short for obtaining patients from regional hospitals. As one oncologist argued:

> If we want their patients as well, and I think we do need them in order to reach our quota, then fourteen days is too short. I mean, you'd have to wait for the regional oncology meetings; that's how it goes. And we have these meetings twice a month.

Finally, they settled for six weeks.

How do these observations relate back to the hopes and goals of the tools' protagonists? The contingent nature of the protracted, ongoing process of negotiations, I argue, is incorporated in the core of the tool itself. Trying to get a protocol to work is a process of making ad hoc compromises, going back to the tool, and tinkering to get the medical practice's elements in line. The number of lymph nodes is juggled so as to articulate the heterogeneous issues involved: the tool makers take whatever opportunities they perceive in order to adequately constrain the links between the diverse, constituent elements of the medical practices. The protocol's "final" state is the highly *contingent* outcome of all this tinkering. Radiologists were given their way, entry criteria were modified, and so forth. Continually, contingencies erode the adjustments accomplished; continually, idiosyncracity seeps in from all sides. This unending, ad hoc compromising leads to a tool *nobody had planned beforehand.*

And because we are talking about decision support techniques, we are talking about embedded decision criteria nobody had planned beforehand. We are talking about the number of lymph nodes or the age limit which determine entry in a treatment-scheme, or about the amount of drugs administered, or about the type of radiological contrast-agent used, as in the account of Eddy. We are dealing, thus, with transformations in the *criteria* to enrol patients in an intense treatment schedule, and to exclude others, and with changes in the treatment given - with, therefore, what "untreatable" breast cancer *means*, and what proper treatment *is*. There is no direct, straight line between the first blueprints of the tool and the tool the practice ends up with. There is, similarly, no simple "ironing out" of idiosyncratic variations. Instead of changing a substandard practice Y to an optimal practice X we end up with practice Z: a hybrid which contains traces of X, Y, and all the struggles that have been fought. Instead of the transparent, optimal, unified Clinical Rationality hoped for, we end up with opaque, impure, *additional* rationalities. Instead of imposing order

where there was disorder, an order is achieved which *incorporates* the very messiness it started out to curtail.[22]

Similar conclusions can be drawn regarding the computer-based tools. Returning to the early Bayesian version of ACORN, we saw how the creation of the data item list began as a list generated by physicians, which was subsequently cut down by checking the individual items' reliability in the hands of nurses and by statistically determining their discrimination power. Through negotiating processes between nurses, physicians, the physical capabilities of ACORN, the possibilities of the statistical techniques and the patients' hearts, a small list of items, with their mutual relations, was obtained. The knowledge incorporated in the practice including this new tool, thus, was a compromise between the needs of the heterogeneous elements involved. Seen from a "purely" statistical point of view, the end result would have been a practice which comprised highly corrupted decision strategies. The ongoing negotiations during construction and implementation stripped all purity away: the statistical techniques were not invoked until the list of data items, created by questioning the clinicians, appeared too large. Moreover, the list of data items was cut down further by a very pragmatic requirement: the repeatability of the data item in the hands of nurses.

From a "purely" clinical point of view, the situation is equally incomprehensible. The construction of ACORN's later, expert system version started out with a reduced version of the data item list described above, and worked from there. So, it began with building blocks which already formed an impure mixture of statistical and clinical logics. Moreover, the endless ad hoc cycles of refining the rules, debugging unwanted interrelations between them, and altering the advice options created a system which further distanced the system from a simple "simulation" of physician's reasoning.[23] As the end result of the complex negotiations in the implementation of computer-based decision tools, we encounter practices which embed decision strategies unknown and unfathomable for any physician. Questions are omitted which are deemed essential by physicians (such as about the nature of the pain), and other data items take on relevancies unintelligible to doctors.

In the work of getting a computer-based decision tool to function, knowledges and decision strategies thus change beyond recognition. The ensuing logic of ACORN's practice is of a new, irreducible kind: clinical and statistical logics are mingled with each other and infused with a host of ad hoc alterations resulting from resistances encountered in the medical practices involved. This generation of idiosyncratic, impure logics, aside from being an intriguing phenomenon in itself, again runs counter to the hopes and goals of those who want to "cleanse" medical practice of its problem of medical practice variations. Rather than "cleansing" medical practice of diversity, ACORN introduces a *new, additional* variety - an approach to patients with chest pain equivalent to that of no physician or practice anywhere else. Stated somewhat more generally, computer-based decision techniques do not do away with the heterogeneity of logics in medicine. Medical practice variations will not be abated through these tools. As in the case of protocols, they do not bring the sound, unitary Logic of Science hoped for. They incorporate the very idiosyncrasy the tools' ad-

vocates wanted to expel from medical practice. They introduce new logics, new rationalities, and thus continue, or maybe even intensify, the multiplicity they started out to erase.

Of Nodes, Nurses and Negotiations: Some Conclusions

The creation of a niche in a local practice for a decision tool, I argued, can be understood as a dual process of disciplining a set of heterogeneous elements constituting the local medical practice and, simultaneously, the localization of that tool. Three recurring ways were distinguished through which the required disciplining is often attempted: reinforcing bureaucratic hierarchies, materializing the tool's demands, and shifting decision power to the most uniform and predictable elements. Neither of these "strategies" is fool-proof, however: hierarchies and forms can be circumvented or resisted, and seemingly "stable" elements can be impossible to gather. Building simple, robust worlds is a hard task indeed. The tool and the practice mutually transform one another to each other's image: the practice is redesigned to mirror the tools' formal structure, and the specific contexts from which the tool emerges are inscribed in its design. Drawing upon a notion coined by Bowker (1994b), I call this the *convergence* of tool and practice.

The processes of convergence lead us into a different world, which was not there before. The tool does not simply come to mirror some optimal or pre-existing practice, nor does the practice come to wholly reflect the tool's original script. On the contrary: the tool's formal setup restructures the knowledges it carries in distinctive ways. Similarly, the knowledges become "corrupted" with contingent cut-off lines and idiosyncratic considerations through the ad hoc tinkering which is part and parcel of getting a tool to work. Through the mutual redescriptions, the medical knowledge embedded in the tool is inextricably interwoven with its "context" (*cf.* Akrich 1992). The decision technique inhabits this new world as an active entity - but the tasks it performs are not simply "tasks taken over" from medical staff. In this taking over, the tasks are fundamentally transformed. This issue, again, generally eludes both protagonists and critics, who remain stuck in arguing which human tasks a formal tool can take over and which it cannot. In thus reifying the debate, they render the question meaningless; in equating the tasks, they ignore the crucial transformations that are taking place before their eyes.

Two final points remain. First, I want to point at a particular paradox confronting the statistical tools and the expert system. These tools, I argued in Chapter 2, contain an ideal of rationalizing medical practice through single, discrete interventions in the work of medical personnel. At the moment of decision, some type of mechanical inference (whether statistical or symbolic) should take over from the physician and replace or support the latter's imperfect cognitive capacities. Whatever happens "outside" of this cognitive realm, like the physical gathering of data or administering of a selected therapy, is of secondary importance and need not be tampered with. In fact, it is an important claim that these issues *should not* be tampered with. Builders of decision-focused tools criticize the protocol for

striving for standardization for its own sake. Rather than constricting the physician in bureaucratic webs, these designers want to intervene only and superbly at that single moment where they are most needed: the moment of decision. The expert system builders especially phrased this desire explicitly: ideally, these authors hoped, the rules within the tools would become so fine-grained and so all-encompassing that the tool would smoothly slide into the practice, its input and output features articulating effortlessly and powerfully with the ongoing activity.

This hope, however, has remained just that: an aspiration which has lost most of its luster. The desire to create formal tools which hide their formality is like the desire to create light without heat.[24] The introduction of decision-focused tools inescapably entails that the heterogeneous elements of the involved medical practice be kept in check - including increasing restraints on the actions of nurses, physicians and patients. Moreover, this disciplining is not merely confined to the direct dealings with the tool: laboratories which supply data can be involved, the whole encounter with the patient can be transformed, and so forth.

The fact that the decision-focused tool's logic is more often than not fundamentally unintelligible to the physician dealing with the tool only strengthens this phenomenon. Confronted with such a tool, the physician really has no choice but to *either* accept the tight lead offered by the tool, or to abandon the tool altogether. An expert system or statistical tool often simply does not allow for short-cuts and unforeseen situations. The physician is presented with a prefixed series of questions to which (s)he has to enter predefined input. And if, for example, the physician can get away with not or partly answering a question, (s)he has no way of estimating the consequences (see also the next chapter). The tool's rationale is hidden within a literally opaque, black box. Moreover, the whole process of interaction with the tool is unidirectionally geared towards the "single-moment intervention"; due to the "plateau and cliff effect," any deviation in the route towards this moment might push the system inadvertently off course.

Paradoxically, thus, we are witnessing tools which, in order to generate a one-moment intervention, require the disciplining of a much broader realm of activities. One could say that their failure to live up to their ideal, projects outward into the surrounding practices. To compensate for the inner limitations of these tools, a tight hold on their environment becomes necessary to maintain their functioning. And as the localization of the tool's rationale renders its interior more and more obscure, this hold inevitably becomes firmer. The decision-focused tool becomes somewhat like a Trojan Horse: while pledging to turn medical practice into a science, while hoping to free medical practices from the "encroaching control" from outside the medical profession, the tight network it necessitates will coalesce with this control rather than detain it.

A final corollary of the scope and depth of the mutual transformations I want to point at is that decision support tools are, and are perceived to be by the personnel affected, thoroughly *political* tools. The scripts of decision tools inevitably embed prescriptions as to who is in charge, and who gets to use which technologies (and thus gets the status that goes with that).[25] In prescribing (expensive) drugs and tests, the tools predefine where what

costs are made, who gets to make these decisions, and who is accountable to whom. Similarly, the tool redefines which laboratory tests are important and which medications are used - and, therefore, whose labs are passed, and whose offices are filled. Physicians working with research protocols are often "wielders of hope": in their hands lie new drugs that might help where other medications have failed. Furthermore, the tool determines which patient is eligible and which is not; how much risk and suffering is tolerable; who is "high risk" and who can wait in queue for the doctor; who gets "a last chance" and who does not.

Contrary to the rhetorics of the tools' advocates, the tools do not "confine" themselves to carrying Rational Knowledge. Rather, the scripts of the tool redistribute responsibilities and redirect patient flows - political choices, therefore, are inescapably intertwined with the "knowledges" carried. Bringing such choices into sight is of importance especially since the way these tools are presented hides these implications from view. The emphasis on their rationalizing potential, the "scientific authority" their discourses endow them with, shunts attention away from these very real consequences. The seemingly straightforward legitimation of being part and parcel of Rational Medicine can destroy the memory of what was needed to make the tools' functioning possible (Bowker 1994a).

These politics are just as much an unplanned, contingent result of the negotiation processes and an unexpected consequence of the formal characteristics of the tool as the decision criteria embedded in the tool. The precise responsibilities of nurses as set by the final version of ACORN and the leeway radiotherapists would get in the breast cancer protocol could not have been predicted beforehand. The politics in the lines of the protocol, or within the wires of the computer-based decision tool, cannot be directly mapped onto the "intentions" of the tool builders.

Focusing on the politics of individual tools is much more useful than debating global questions as whether formal decision tools are "deskilling" or "empowering," "dehumanizing" or "restorers of hope," and so forth. The sheer complexity and scope of the negotiation processes pointed at results in local networks which are thoroughly changed - and in often unforeseen ways. The way to "judge" individual decision support tools is to look at how individual elements' juxtapositions vis-à-vis one another change. Whether the reshufflements leading to the silencing of some voices outweigh the increases in responsibility elsewhere; whether the way the "patient" is described in the script of one tool is preferable to another. To whom happens what; what capacities are attributed to what element? Getting to the politics implies scrutinizing the way the new local network's relations redefine its constituents. Whether the nurses' newly gained tasks outweigh both the tight lead ACORN imposes on them and the transformed relations vis-à-vis the emergency ward physicians. Whether the protocol's criteria include too many women in order to ensure a sufficiently large flow of patients. When the tools' straightforward universality is a chimera, we cannot expect our global judgments of them to carry very far either.

Notes

1. This story is derived from Eddy (1992).

2. This included both reactions due to the high osmolality of the iodinated contrast agents and allergic reactions. The nonionic, more expensive agents supposedly caused fewer adverse effects in both categories.

3. This is the so-called "framing effect"; see Chapter 2, note 10.

4. Interview Emerson July 21, 1993. The work of Bowker and Star (1994) on the International Classification of Diseases has been especially influential in the writing of this section; see also Fujimura (1987; 1992), Star (1989; 1991a), Star and Griesemer (1992), Kling and Scacchi (1982), Webster (1991). Studies of information technologies in medical domains in which the same phenomena as described here can be seen to occur are Dent (1990), Rudinow-Saetnan (1991), Bloomfield (1991), Lipscombe (1991) and Hartland (1993a). The latter two cover a whole range of different medical expert systems.

5. This may sound somewhat cryptic, but the statistical method by which the design builders selected the "powerful" from the "weak" data was based on a database containing only a sample of the *local* population of "patients with chest pain coming to this emergency ward."

6. When asked about the definition of acute abdominal pain, de Dombal replied in 1976: "Of course everybody else has their own definition, and this may differ from place to place. At the moment, we don't even know what these differences are - but we must certainly find out" (de Dombal and Gremy 1976, 174). In building computer-based decision tools, many authors have come up against this issue as a hindrance to their goals (see e.g. Barnett, 1968; Lusted, 1968). The following quote from de Dombal (1989) illustrates the persistence of this problem: as the first "obstacle to progress, [standing] in the way of widespread implementation," he mentions the lack of clear medical terminology. "Medical questions and answers have to be simplified, pre-defined, and agreed in advance by wide consensus" he argues (see also Gill et al. 1973).

7. See Bjerregaard et al. (1976); Horrocks et al. (1976); Fenyo et al. (1987). "The clinical spectrum of e.g. pancreatitis cases differs," Bjerregaard et al. note, "alcoholic male cases being seen more often in Copenhagen than in Leeds."

8. Similar problems have been reported for consensus reports. Kosecoff et al. (1987) found that compliance with consensus statements varied by hospital. One of these statements, regarding urgent performance of coronary angiography in patients with unstable angina, was (not surprisingly) found to be adhered to more in hospitals having a catheterization laboratory. See also Lomas (1989).

9. Several authors have reported on the problem that physicians often see consensus reports as being "biased." "The physicians seem to be aware of the sponsors' intent to produce a document that meets the sponsors' need to communicate a particular message about how medicine ought to be practiced" (Hill and Weisman 1991; see also Grol 1989).

10. Marks' account of the therapeutic research on pneumonia serum and arsphenamine in the beginning of this century provides a historical example of this phenomenon (forthcoming, Chapter 2).

11. Interview de Dombal, and see Ikonen et al. (1983); de Dombal (1988); Chatbanchai et al. (1989) and de Dombal et al. (1991).

12. *Cf.* Field and Lohr (1990); Eddy (1990e, 1990f); Society for Medical Decision Making Committee on Standardization of Clinical Algorithms (1992). Eddy's difficulties to use international data were due to an array of reasons resembling the "non-transferability of databases" discussed above.

13. Intimately related to localization in space is what might be separated as yet an additional type of localization: *localization in time*. A tool is always bound to a specific time, to a specific state of medical knowledge, a specific epidemiology: just as it is located in a space, it has a history and a projected future (*cf.* the notion of the trajectory of decision support tools in Timmermans and Berg, forthcoming). I have not opted for localization in time as a separate category however, since striving for a tool which has an infinite, universal usability through *time* is generally not part of the advocates' ideal-typed views.

14. See for similar self-diagnoses of the "current state of the art" in computer-aided decision tools e.g. Miller and Masarie (1990) and Schwartz et al. (1987). A typical example is MYCIN's successor, ONCOCIN, which explicitly "retreated" on a very small domain: it

is designed to control the usage of predetermined, oncological protocols (Shortliffe and Clancey 1984). A more detailed discussion of ONCOCIN can be found in Lipscombe (1991); see also Berg (1994).

15. Wilson et al. (1977), interview de Dombal. See for similar experiences e.g. Davis et al. (1977) and Engle (1992). If nothing more, it is an enormous amount of work just to expand the required data base and/or rule set: INTERNIST, designed to cover "all of internal medicine," had cost 15 (wo)man years of programming time. INTERNIST, however, was notorious for its strange interactions between rules, its many errors, its impracticality and an overall sacrifice of depth, precision and transparency (Miller et al. 1986; Miller and Masarie Jr. 1990; see also Sutherland 1986).

16. Similarly, Greenfield's upper respiratory tract infection protocol cut off some diagnostic routes since the health assistants were not trained to do an examination of the heart: it was not contemplated to have them listen to murmurs or gallops (interview Komaroff, November 9, 1993).

17. I am not arguing that no system could do so. Rather, the point is that *in this specific practice*, with this set of data items available, this local population of patients, this group of nurses, and so forth, it was not feasible.

18. "Scope" has a temporal dimension too: many tool builders have criticized their own tools for just taking "snapshot" pictures; focusing merely on a "'snapshot' of data about a patient at a fixed time." Most computer-based tools do not take the temporal development of illness, symptoms, reactions to therapy and so forth into account; "the past is treated in the same vein as the number of years you've smoked" (Buchanan and Shortliffe 1984, 393; the last quote is from a discussion with Buchanan, December 6, 1993).

19. Miller et al. (1986); Barnett et al. (1987); Miller and Masarie Jr. (1990) and Lipscombe (1991). On this phenomenon as a broader trend, see e.g. Perolle (1991).

20. Some commentators have argued that just about any tradeoff (say between the speed of the tool-in-use and its comprehensiveness) could be added as an additional type of "localization." I do not think that this is so. I focus on space, scope and rationale since these are dimensions where *universality* is the ultimate ideal-typed goal *and* the ultimate reason of existence of these tools. "Comprehensiveness" or "speed" is much less basic to what these tools stand for than e.g. "scope."

21. See also Star and Griesemer (1989), Fujimura (1987; 1992) and Timmermans et al. (forthcoming).

22. For a similar tale vis-à-vis policy-oriented decision analysis, see Ashmore, Mulkay et al. (1989).

23. Again, this is a phenomenon which can be seen time after time. For e.g. MYCIN, see the detailed book by Buchanan and Shortliffe (1984). On INTERNIST, see e.g. Miller et al. (1982); Pople (1982) and Miller et al. (1986).

24. The expression "light without heat" is derived from Zuboff (1988, 349). For a harsh insiders' critique of the success of expert systems, see Sutherland (1986).

25. *Cf.* Löwy (1993). As pointed at, joining the breast cancer protocol group had a distinctive advantage for oncological centers: they would get additional funds with which they could buy "shiny," "state-of-the-art" technology. On the status attached to working with (decision support) technologies, see e.g. Ashmore et al. (1989); Bloomfield (1991); Saetnan (1991).

5

Supporting Decision Support Techniques: Medical Work and Formal Tools

In the previous two chapters I looked at how tool builders reshape both practices and tools in the process of getting their decision techniques to work. This chapter changes the point of entry again: from looking at the work of tool builders, I shift to studying medical personnel at work. The focus turns and centers on the primary work tasks of what was so far just an element in the heterogeneous networks that had to be disciplined.[1] How do formal decision tools figure in the work of medical personnel? How do they transform this work? Should we conclude from the previous chapters that a functioning decision support tool implies that the medical staff is rendered docile? And if not, what does the resulting configuration look like?

To get a grasp on these issues, the chapter begins with investigating a rather basic question: what is it that medical personnel do?[2] I question the views of medical practice contained in the decision tools' discourses. I challenge, more specifically, those aspects that rhetorically allow for a smooth fit between the work of medical personnel and the functioning of the formal tools: the clear-cut nature of medical data and criteria; the separation of the "scientific" content of medical work from its social, potentially biasing context; the sequential, step-wise nature of medical practice; and (for the decision-focused tools) the prominent role of the physician's mind. By attributing these characteristics to medical work, the discourses discussed in the first two chapters create a seemingly perfect match between the formal and "scientific" nature of the tools and the nature of medical practice. In these ideal-typed views (modeled after positivist images of science and the developing tools themselves), the nature of medical practice inherently provides for the requirements of the formal tools.[3]

This now-severed "natural link" is (re)created in concrete practices through the disciplining of practices to specific formalisms. Studying medical personnel at work, however, brings the fragility of the achieved discipline (of *all* heterogeneous elements) in full view. It shows how this inevitable fragility leads to a situation in which disciplining efforts are moderated, that is, additional localization is allowed, in order to prevent total failure. In other words, this chapter will argue the seemingly paradoxical claim that formal tools can work only since medical personnel are *not* transformed into fully docile bodies. In and through their work, they can actively *produce* many of the tool's prerequisites - while reappropriating the

tool in the process. Only in this way, it will be shown, is the transformative potential of the tools realized.

This argument has consequences for some of the main claims of both critics and advocates of the tools, which will be discussed in the final sections of this chapter. Through "pre-analyzing" the decisions, through taking some responsibility away from the physician - whether only at the moment of decision or during a whole sequence of actions - the work of medical personnel was supposed to be made easier. But is this so? How does the need to actively produce the formalism's requirements affect the advocates' hope of "simplifying" a practice through decision support techniques?

On the other hand, many critics of decision support techniques have argued that these tools take control of the medical decision making process out of the hands of the physicians. In their view, this control should be returned to doctors - either through abandoning the tools or through fundamentally changing their design.[4] These critics, however, overlook that through the medical personnel's production of the formalism's requirements, these formalisms can become active agents in the practices - creating new dimensions in the personnel's work. Neither the tools nor the staff members can be said to "be in control." Rather, control is distributed among them in intriguing and as yet poorly understood ways - and it is in this distribution that the power of the resulting hybrid lies.

Medical work: the heterogeneous management of patients' trajectories[5]

I met Mr. Wood during my participatory research at the Oncology ward and outpatient clinic of a University Hospital in the Netherlands. Two years earlier, his internist had told him that he suffered from Hodgkin's Disease: cancer arising from the lymphoid tissues. He had visited his doctor because of his persistent coughing, his fatigue and the fact that his weight had dropped, in only a few months, from 120 to 100 kilos. The internist found tumorous tissue in the neck, in both armpits and in the space between the lungs, the mediastinum. Mr. Wood was given chemotherapy, followed by radiation of some of the larger tumors. His disease responded well: after this treatment, his symptoms were alleviated and the tumors had disappeared.

Some months later, however, Mr. Wood was back at his internist's office. He was gradually feeling tired again and his weight was dropping. What he had been afraid of appeared to be the case: his Hodgkin's Disease was back. The cancer, as all too often is the case, had not been eradicated completely. He was sent to the University Hospital to see whether he would be eligible for bone marrow transplantation. In this therapy, bone marrow is collected from the patient's pelvic bone and deepfrozen. Subsequently, the patient is treated with massive doses of chemotherapy. These doses are so high that they kill the bone marrow cells, responsible for the production of bloodcells. Without these cells, the patient is depleted of vital parts of the immune system and dies of infection. Also, the lack of blood platelets, crucial for hemostasis, may cause lethal internal and external bleedings. In bone marrow transplantation, however, the unfrozen bone marrow cells are given back to the patient after the highly toxic chemotherapy, which is hoped to affect the tumorcells equally deadly.[6]

Mr. Wood was found eligible for bone marrow transplantation and was hospitalized. When I met him, he had already had his treatment and his bone marrow reinfused. This stage is a crucial one, since the effects of the chemotherapy on the bone marrow are starting to show, and the "reinstalled" bone marrow is as yet only

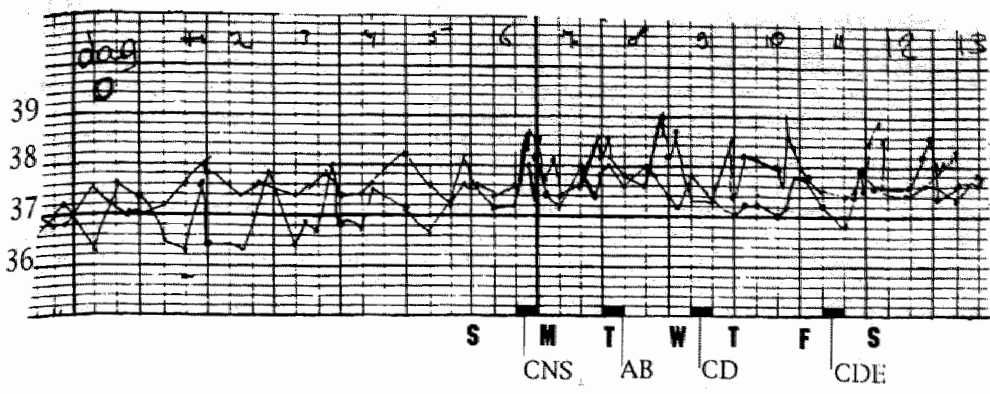


Figure 1. Mr. Wood's temperature curve. The relevant line is the one with the high peaks. The other is the pulse curve. The vertical lines indicate eight-hour periods, the letters under the graph indicate the days of the week. The row of numbers at the top indicate the number of days after bone marrow reinfusion. AB: start first antibiotic. CD: shift to second antibiotic. CDE: third antibiotic added. CNS: time of taking of bloodculture which was positive (see text for details).

starting to proliferate. Bloodcells are counted daily, and their decline is closely monitored. The patient is closely monitored for signs of infection or bleeding. To minimize the risk of infection, the patient is kept in isolation: (s)he stays in a special room, with an air lock and double doors. To enter, visitors and medical personnel have to wear special robes and masks. The isolation procedures are continued until the bloodcell counts have again risen above a certain threshold. Then the patient is declared "out of the dip" (referring to the dip in the curve of the bloodcell count graph in the patient's record), and the isolation measures are stopped.

It is Monday. Dr. Howard, head of the department, has been on duty this weekend. When he meets John, the senior resident who runs the day-to-day affairs of the oncology ward, he tells him to keep an eye on Wood's temperature. "They called me to say that it rose above 38.5 this weekend" says Howard. "I said that they should check again two hours later; and if it was above 38.5 again, we would have to start with antibiotic treatment. But it wasn't. You've got to check it, John, because he's got 0 granulocytes [type of white bloodcell] at the moment, and that means he's prone to get a blood-poisoning [infection of the blood itself], and then they're dead before you know it."

During the day Wood's temperature remains below 38.5 degrees. John investigates Wood, but cannot find a reason for the fever. That night, however, his temperature rises again: when it is measured at 10 p.m. it is 38.8. According to her routine, since the temperature is above 38.5, the nurse calls the resident who is on night duty and sends a sample of Wood's blood to the laboratory for a bloodculture. The resident, in his turn, calls Howard, who tells him to wait two hours and measure again. At midnight the temperature is 38.6. Howard is called again, and he states that this could imply that Wood is having a blood-poisoning. Howard orders to start a broad antibiotic therapy, consisting of two antibiotics which I will refer to as AB. When John arrives in the morning he reinvestigates Wood and finds a small anal fissure close to a pile. He jots this finding down in Wood's file and adds an exclamation mark: this could be the source of the infectious trouble! He orders an X ray of the lungs to search for other possible focuses, but the X ray, according to the report, is "clean."

The next day, Wednesday, John discusses Wood's case with Liston, the oncologist who supervises John's work on the ward. "He's been having a fever for some days," says John, "and lately he stabilized above 38.5 degrees, so we started with AB. And we've just received the bloodculture from Sunday, which says that in one

out of the four tubes of blood a coagulase-negative staphylococ has been found for which AB is effective. So we're on the right track." [a coagulase-negative staphylococ (CNS) is a type of bacteria]. "Yeah, well," replies Liston while she is pointing to the temperature curve, "but the temperature is still rising, so we're going to shift to CD [a different combination of antibiotics]." John is silent, while Liston writes this change in Wood's record. "CD hits that bacteria even harder," she adds.

Two days later, Friday, during the so-called "paper rounds," John recounts this story to Bear, another oncologist who has taken over supervision duty. That night the temperature had risen again, to 39.0 degrees. The resident on duty had not undertaken any specific action, and had not called Bear - it had been a single peak. "He is some thirteen days after transplantation" says John. "He has acquired a fever, since, actually, this weekend. We started with AB, and changed to CD on Wednesday." Matthew, a nurse attending these rounds, adds: "Yes, and his blood pressure is low, with a pulse of ...." Bear interrupts him, and continues on the topic of the antibiotics. He has not heard about this patient for a while, and he wonders about the quick change in therapy. He asks the bacteriologist, who is frequently present at these rounds, whether he has seen a positive bloodculture. "No." Bear then sees in the record that "one out of four tubes was positive." The bacteriologist shakes his head. "One out of four means nothing. I would define that as a contamination, as a skin bacteria which has accidentally entered the needle when the blood was drawn. I would see that as a negative bloodculture." John intervenes to defend the shift in antibiotics, which now looks somewhat dubious: "Since the temperature was high and not responding to AB, we changed to CD." The bacteriologist shakes his head again, and repeats: "one out of four tubes is nothing" - implying that there was no good reason to shift from AB to CD. Bear now comes to John and Liston's aid: "Yeah, well, we know that, but we had to shift the antibiotic treatment anyway, since the temperature wasn't reacting properly." John, shifting the discussion somewhat, adds: "His pile presents a similar problem. Do you have to gear your antibiotics towards that, or towards the positive tube?" The bacteriologist is confident: "In a situation like this the clinical situation should prevail. That's what I think." He explains that DE would have been a better choice: that combination would have been targeted more directly against bacteria which may be present in the pile. Bear joins in: "Yeah, that pile problem. Why haven't I been called last night? There was a high temperature, and we should have acted in some way. And what do we do now? Do we change from CD to DE or add something to CD?" The latter is not the most elegant solution: these three agents overlap each other significantly. "However," the bacteriologist remarks, "you have, in fact, already made the decision. I think it would be best, now, to add E." Which is what happens.

Some days later, Wood's temperature quiets down, and the isolation measures are stopped. In the discharge letter, one month later, this episode is summarized as follows:

On day +8 [eight days after bone marrow reinfusion] a fever develops. The focus appears to be a thrombosed [blood-clot containing] pile. Empirically [based on the clinical picture] AB is started. Bloodcultures show a [CNS], reason to treat patient further with CD. Because the temperature responds insufficiently, E is added. The temperature subsequently normalized only slowly. The anal fissure eventually quieted down. No other infectious problems have occurred. ... Patient was discharged in a good general condition [35 days after reinfusion] and went home.

***

Medical personnel are continually struggling to *make* a patient's case work: to keep a patient's trajectory "on track" (Strauss et al. 1985).[7] In this process, anamnestic information, examination results and medical criteria are not so much "uncovered" or "given" but are continuously *(re)constructed*.[8] In the next fragment, Daton and Beatty can be seen to construct their respective sets of medical data:

Dr. Daton, a general practitioner, is called in by her patient Mr. Porter, who, as Daton knows, suffers from chronic bronchitis and low back pain. This time Porter is complaining of "attacks of lung pain" (he rubs his hands against his chest) which come up unexpectedly and disappear promptly upon taking the "pills from the specialist."

D. When you've got that pain, do you want to lie still or rather move about?
P. No, doctor, I can't lie still.
D. Does the pain start when you exert yourself, or when you leave the house when it's cold out?
P. No.
D. Do you tire quickly when you exert yourself?
P. Yes.
D. Does the pain start in your low back?
P. Yes.
D. When you're in bed, or also at other times, for example when you make a wrong movement?
P. Well, doctor, I'm mostly in bed or in my chair when the pain comes, but yes, I do sometimes make a wrong movement.
D. And the pain stops when you take the pills I gave you? (Daton had given him Feldene, a simple pain killer)
P. Yes, doctor, but the pills of the specialist do work better, more rapidly.

These "pills" appear to be Nitrostat, an anti-angina pectoris drug which reduces the workload of the heart by dilating the vascular system. Driving back to her office, Daton remarks that Porter is actually suffering from atypical low back pains. Nitrostat, she claims, only "works" because of a placebo effect.

According to the letter from Dr. Beatty, a cardiologist to whom Porter had been referred by a fellow general practitioner, the electrocardiogram (ECG) and the cardiac enzymes were normal. According to Beatty the pain generally started in the left side of the chest. His diagnosis: "atypical angina pectoris."

Patients' symptoms take shape in and through the questions a physician asks and the examinations (s)he performs. Daton only pursues her questioning on items and remarks she considers to be relevant to differentiating between low back pain and angina pectoris ("Does the pain start when you exert yourself?" "Does the pain start in your lower back?"). In this way, she elicits information corresponding to the outcome she is considering: in these questions, the typical diagnostic patterns are already embodied. Similarly, cardiologist Beatty does not ask any questions concerning Porter's backaches. Not familiar with Porter's medical history, he does not consider this option.

In addition, questions like "when you've got that pain, do you want to lie still or move about?" already contain the preselected answers. The expected answer can be implied in the question: the reactions to the questions "the pain starts in the low back?" (Daton) and "the pain starts in the left side of the chest?" (Beatty) are *both* affirmative. The very dissimilar outcomes these two doctors are considering are both confirmed.[9]

Likewise, the physician selects certain examination procedures and omits others, thus producing the pathological reality (s)he will want to counteract.[10] In the case of Mr. Wood, John selects an X ray of the lungs but not of the abdomen. The piles and the fissure, also, would never have appeared in the records if John had not specifically looked for them.

Gathered data mutually elaborate each other (Whalen 1993): the pattern of examination data and anamnestic information lead Beatty to denote the attacks of "lung pain" as "angina pectoris," which is often felt outside the heart area. At the same time, these elaborations can lead in wholly dif-

ferent directions: to Daton, who draws upon different background information (she knows Porter as a "back pain patient") the vague designation of the pain indicates referred back pain. Similarly, the normal ECG and cardiac enzymes are mobilized by both physicians as sustaining their divergent proposed solutions. These findings can be seen as compatible with both angina pectoris and low back pain.

The acquired solidity of anamnestic information and examination results can continually be undone. When data conflict, coherence is often restored through *reconstructing* a piece of information: by rephrasing a question, for example, or by stating that the patient "did not seem very reliable" when the history was taken. John denotes the temperature peaks as "a temperature stabilizing above 38.5"; Liston, however, relabels the temperature as "rising" (see Figure 1).[11]

A prominent feature of these reconstructions is that there are no types of data which *always* prevail when they clash with others. It would be erroneous to conclude that examination data, for example, form the "rockbottom" to which anamnestic data are adjusted; or that laboratory-data "win" over physical examination results. *All* data can be and are reconstructed. Where the fast response to Nitrostat would have convinced Beatty of his diagnosis, and would have countered any patient's remarks about the pain starting in the lower back, Daton weighed these contradicting pieces of information in exactly the opposite way. She nullified the effect of Nitrostat by stating that it was just a placebo effect. There is no "foundation" to give us a final answer; no "basic ground" upon which these reconstruction processes rest.[12]

"Medical criteria" are (re)constructed in comparable ways. While the bacteriologist stated that "one positive tube is a negative," John referred to this one tube as a legitimation of the change in treatment. The patient had a fever, so *in this specific situation* this one tube was relevant. Similarly, in this ward it had become common practice to refrain from immediate intervention when a temperature spiked above the 38.5 limit a single time. More action than the routine blood culture would only be necessary when a repeated measurement, two hours later, would again show a temperature above this limit. In this specific situation, however, Bear disagrees. He should have been called, he argues, even though there was only one single spike.

Medical criteria should thus not be seen as fixed rules which merely have to be "applied." Like medical data, criteria are (re)adjusted to concrete situations: the rule "one tube is no tube" can continually be reconstructed. Here also, no "foundation" can be found from which the processes of construction and reconstruction could be derived. On the contrary. When work proceeds smoothly, when there are no conflicting data, "criteria" are seldom explicitly mentioned or referred to. In those moments, action is transparent; "what always takes place takes place." Rules or criteria are only made explicit when this smooth flow is interrupted; when it is no longer obvious what to do (Suchman 1987, 53-54). As ethnomethodology has taught us, explicit "rules" are not the foundation of action: they are post-hoc, contingent and situation-dependent *derivatives* of concrete action. They are codifications of "what is normally the case" - but since such a codification is always partial and tied to a particular situation, it is always

open for reinterpretation. Criteria are fluid *resources* for the actors involved, who can actively orient to them in order to accomplish the task at hand.[13] In the case of Mr. Wood, adding antibiotic E to C and D is "normally" not a very elegant combination. This "rule," however, ignores the possible event that the combination of C and D has already been started recently. *In this situation*, stopping with C and D seems like an even less elegant solution.

Medical practices, in addition, are not solely concerned with data and criteria: there are more "cross-cutting systems of relevance" which interfere (Bosk 1979; Strauss et al. 1985). Organizational limitations, the patient's needs and desires, or financial matters are seen as "secondary" or "bias" by the decision support tool's discourses - or are simply neglected.[14] In doctors' and nurses' work, however, such issues are inseparably interwoven with "medical" matters as decision criteria and examination data. For a physician confronted with a patient's problem, these elements play similar, equally essential roles. John, for example, has to gain Liston's backing for continuing the AB treatment. But John does not obtain her support: Liston changes the treatment. Similarly, Matthew does not get the physician's support for his remark on Wood's low blood pressure. He is simply ignored - an experience familiar to nurses.

Organizational issues such as the support of colleagues, then, may obstruct the smooth flow of medical work just like contradictory data can. As was the case with the data, such issues may lead to the reconstruction of other elements, as the following fragment illustrates:

> Mr. York had recently been admitted to the oncology ward for an intensive course of chemotherapy. He had just started this treatment and, according to the treatment protocol, he was supposed to remain in his room. However, the protocol also required some X rays and other investigations which had not been done yet. So the medical personnel wheeled him around anyway - and my questioning about this was seen as a flagrant example of not understanding "how these things work."

Here, medical criteria were "overruled" because of the practical exigencies of the situation - the nurses had not had time to do the investigations earlier. On the other hand, medical personnel often intervene in organizational routines in order to secure a specific outcome. Bear risked a fight with the hospital's dispensary by ordering a specific medication he wanted for Mr. Wood without asking the dispensary's chief pharmacist first. Following the organizational rules, Bear said, would have wasted too much time and trouble.

To summarize, medical work is a "moulding process in which the patient and his situation are reconstructed to render them manageable within existing agency routines."[15] A physician, confronted with a patient's problem[16], tries to *transform* this into a *manageable* problem: one which matches his/her work routines; which implies a limited set of actions (s)he perceives to be a sufficient answer (at this time and place). All heterogeneous elements mentioned (data, organizational issues, medical criteria, and so forth) reciprocally shape this transformation *and* are moulded in this process: John may try to obtain Liston's support to continue with the AB combination, and Bear may try to reshape the organizational routines so that he will be called. In this transformation, aspects from the patient's history take shape or are forgotten and the patients' hopes and desires are

transformed; the result is a patient having backaches necessitating Feldene or angina pectoris requiring Nitrostat. What counts as the solution of the patient's problem is a result of the *outcome* of the transformation; equally, what counts as the original problem is redefined during this process (*cf.* Davis 1986).

Medical work as Situated Work[17]

> There are so many more places in the world than a doctor's head.
> (Mol 1993)

Physicians are not collecting data in order to create a "true image" of their patients' bodies. For one, physicians are too aware of the rough and blunt character of the techniques and instruments available to want to strive for such an illusion. Diagnostic technologies have too many "ifs" and "buts" attached: John's X ray of Mr. Wood's lungs will miss a small focus of infection, a "rising temperature" can be an artefact of two slightly differently performed measurements, and so forth. Physicians, moreover, are not even primarily interested in creating such a "true picture" even if they could. Their work is directed at finding an answer to what Garfinkel has called the "practical problem par excellence: 'What to do next?'" (1967, 12). With the tools they have, they do not attempt to create "true" images of nature, but a *meaningful difference for the purpose at hand*: a result sufficient to direct the immediate course of action.[18]

The "manageable problems" thus constructed are always only *provisional:* they are "adequate-for-the-moment" (Heritage 1984, 149). They are but temporary "crystallization points" which start to dissolve when the purpose at hand is no longer immediate.[19] Liston orders a change of the antibiotic therapy to CD, but she does not prescribe what to do if the temperature should again fail to respond. Such detailed, prefixed plans are useless. The achieved "fit" between the heterogeneous elements is precarious; the rough and blunt techniques available for constructing, monitoring and maintaining the patient's trajectory are all too easily brought off course. At all times, new, contingent developments may occur, necessitating renewed articulation efforts. Organizational routines may break down, an unexpected result of a laboratory test can pop up, or a new supervisor may disagree with the policies of the previous one. In the midst of all this, medical personnel are engaged in a never ending process of ad hoc reconstructions. They are continuously working, managing with odds and ends, to keep the patient's trajectory on track - while, concurrently, reconstructing its course.[20] To paraphrase Latour: "The [physician] makes plans that are constantly drifting away; [s]he seizes upon opportunities in the midst of confusing circumstances" (1988, 259).

All this is reflected in the staff's *reactive, opportunistic* stand. Personnel react to topically salient events; they respond to "cues" that happen to be at hand.[21] Liston changes the antibiotics because the temperature curve tells her that AB is not working properly. Since John has just mentioned the CNS bacteria, Liston opts for antibiotics which affect that micro-organism best (CD). Medical work is making do with what one has got; the positive

bloodculture may be doubtful, but it is the first available cue to which Liston can orient herself. When, somewhat later, Bear worries about the anal fissure and the pile, he simply adds the antibiotic (E) brought up by the bacteriologist. A similar pragmatism is reflected in the habit to "wait two hours" when the temperature raises above 38.5 degrees. This is a practical, informal routine, resulting from the fact that isolated temperature peaks are recurrent and often innocent events. Adhering strictly to the rule would result in much meaningless effort: the physician has to be notified, blood has to be drawn and sent off to be cultured, and so forth. The "two hour" routine is an ad hoc habit to separate the wheat from the chaff. Anticipating on the need for continuous "maintenance work," on the always only provisional nature of action plans, is ubiquitous - "two hours later" might be a different world.[22]

In this perspective there is no longer place for a notion of medical work as consisting of single-moment, cognitive "decisions." "Decisions" often come about incrementally: "intentions" can become conclusions, previous steps can be undone, changed, or simply forgotten. Nurses redirect physicians, and laboratory results can start having a life of their own. Similarly, previous events can suddenly become "decisions" when the setting changes: since Mr. Wood had already had his antibiotics changed recently to CD, E was added to CD instead of selecting a less redundant combination. "You have already made the decision," the bacteriologist remarked, making a joke which played exactly upon this phenomenon: nothing as such *had* been "decided" at all.

"Decisions" are frequently only referred to *afterwards*; as is the case with medical criteria, the occurrence of a "decision" is often constructed *post hoc* to underscore the "rationality" of events. The fluidity which typifies the ongoing work reifies in its subsequent reconstructions: at each new patient round, the previous history of the case is briefly summarized and Mr. Woods' story is further stylized. The two peaks above 38.5 degrees become a "temperature stabilizing above 38.5," until, in the discharge letter, the fever simply "developed at day +8" (*cf.* Figure 1). These continuous reconstructions should not be seen as a falsification of history: they are "needed to produce an account ordered enough to enable action or to communicate what is going on" (Gooding 1992, 76). Ultimately, they create the type of report as exemplified in the fragment of the discharge letter, where *almost every sentence* reflects a history of repeated reconstructive work. As is apparent in the account, the causal role of the pile in the development of the fever was never very clear. Similarly, the anal fissure had never been described as "unquiet"; it was simply the only clue the physicians had.

The resulting trajectory, consequentially, is not so much a product of consciously developed plans as a result of the continuous reactions to contingent, salient features of the local settings staff members happen to be in. The actual itinerary of the trajectory, as Lynch phrases it, is an "in-course accomplishment" (1985, 247). It is continually reset "on the spot," in ongoing interaction with details from the work environment as recent reports, patient records in view, nurses' comments, and so forth. Similarly, medical "decisions" are the *outcome* of continuous, distributed interactions between physicians, nurses, laboratory results, the patient, and organizational set-

tings. In situations differing in time and place, small steps are taken which, in retrospect, appear to have led up to a "decision."[23]

Making patients' problems manageable thus implies much more than a cognitive reconceptualization of the patient's case.[24] The transformation of a patient's problem in a "doable" problem is a practical endeavour, situated within local, material settings. Medical work is an ongoing, active process of articulating a broad array of diverse elements (*cf.* Fujimura 1987). It is ad hoc, reactive, and draws upon opportunities immediately available. It is characterized by the smooth interweaving of "social" and "medical" issues, and by the (re)constructed nature of medical data and criteria. Finally, it is thoroughly *temporally* structured: all activity is located within a trajectory's projected "history" and "future," both of which are continuously reconstructed.[25]

Fragile Tools

This depiction of medical work challenges the smooth fit between "tool" and "practice" so matter-of-factly postulated in much of the writings of the tools' advocates. Contrary to how the decision techniques' discourses depict the practice of medicine, medical work is not characterized by clear-cut medical data and criteria, nor by a clear separation between "medical content" and "social/biasing context," nor by a fixed, step-wise sequence. What should be concluded from this? How is this line of reasoning related to the processes of disciplining and localization? Does disciplining a practice to a formalism imply the transformation of medical work so that it fits the discourses' depiction? Does this imply the erasing of contingencies, the complete obeyance of the heterogeneous elements involved to the tool's script? Do we then witness the replacement of situated work with the meticulous following of preset plans?

If we look at the few days from the trajectory of Mr. Wood recounted above, we have to conclude that these conclusions miss the point. Mr. Wood's trajectory was being shaped by *several* protocols: an international research protocol on the treatment scheme (including directives for post-reinfusion care), a local protocol on isolation procedures, a local protocol regulating his diet, and so forth. This was a highly disciplined practice: nurses were accustomed to working with detailed protocols, organizational patterns were designed accordingly, and so forth. At the same time, the ongoing flow of interrelated, contingent events constituting medical work had not disappeared. How can this be understood? Are these two observations not contradictory?

The contradiction disappears when it is seen that it is exactly through the situated work of medical personnel that they can actively *fill the gaps* between the tool's demands and the ever-occurring instances when the disciplining of the practice falls short. Wherever heterogeneous elements behave "illegally," whenever the tool's script is disregarded, medical personnel can intervene and rearticulate the tool with the ongoing course of events.

To elaborate this argument, I discuss some instances where the tool's requirements are *produced* in the ongoing work of medical personnel. I fo-

cus on some moments where the fit between tool and practice, construed in the tools' discourses, has broken down.[26]

The Active, (Re)constructive Process of Generating Medical Data

Formal tools require clear-cut, elementary data items as input. Medical data, however, do not come pre-defined: they are actively (re)constructed in the process of making a patient's problem manageable. Moreover, their "solidification" only goes as far as is required for the practical management of the problem at hand; at any new crystallization point, new articulations are readily forged. *Creating* the required, "digitized" data, then, is no easy task. Tool builders can labor to select only those elements whose behavior fits the tools' demands most closely, or to ensure standardized usage of proper investigative techniques. Such disciplining, however, is always inadequate. All too often, data present themselves in ways not anticipated by the tool builders, or data items deemed "trustworthy" suddenly behave unexpectedly.

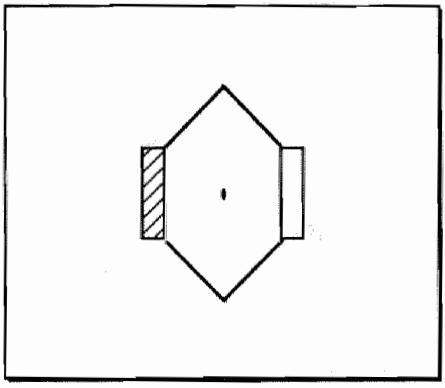


Figure 2. The representation of a patient's abdomen on the form used with de Dombal's tool (see for the complete form Chapter 2, Figure 1). The area crossed by the physician (see text) is shaded.

In a session with de Dombal's tool, for example, a physician ran into the problem that the patient had explicitly mentioned that his pain "radiated to the back." On the form used for registering the medical data, (see Figure 2), he could only mark some areas of the patient's abdomen. "We don't have that on the machine," he muttered, and checked the area representing the patient's flank. As another example, Hartland describes patients interacting with GLADYS (Glasgow Dyspepsia System): a Bayesian computer-based decision tool for diagnosing, broadly defined, problems of indigestion (1993). The system is designed to take a patient's medical history, analyse the symptoms and generate a list of possible diagnoses and a management advice. Hartland describes a patient confronted with a question about the benefits of a particular diet, to which he should answer "yes" or "no." The patient complains:

> at the start it did good, but after that it didn't matter. But I can't get that across by saying yes or no.

When questioned about whether she had had "heartburn," a patient remarked:

> well I don't know if it is heartburn. It's never been diagnosed as heartburn, I just assumed it was, I've had Brufen for it. The doctor doesn't tell me why he's given me Brufen he just gives it to me.

In such situations, Hartland continues, patients often called in the help of nurses, who reassured them or helped completing the questionnaire.[27]

Although both de Dombal's tool and GLADYS ask very basic questions, unanticipated answers immediately generate trouble. Medical data do not come readily packaged: even in recounting an elementary story of stomach troubles, much unexpected translations were required to get the story workable for a formal tool. Medical personnel have to be stand-by to *preprocess* the data whenever they escape the formats prescribed by the tools' script.

In addition, the tools discussed here themselves embed certain translations in their rules or formulas. Inevitably, they articulate certain signs and symptoms in preset ways. ACORN, for example, asks whether the patient has recently experienced an unusual cough, hemoptysis [the spitting of blood], or a fever. When the answer to these and some additional questions is "no," the system concludes (on its printout) that "no respiratory symptom is present." This is itself an example of a translation which is "good-for-all-practical-purposes," and meaningful only in the field of operation of this particular tool. Fever and hemoptysis can both be due to a myriad of causes having nothing whatsoever to do with the respiratory tract - but the system assumes that *if* hemoptysis is present for some other reason, the erroneous conclusion that "a respiratory symptom is present" may be a bit dumb, but not consequential. In these instances, the system assumes that these oddities will be prevented or corrected by the personnel operating the tool. When a patient has had a slight fever due to an infection of his left toe, for example, the nurse using ACORN should be aware that this is *not* what a "fever" means in this particular situation.

The fact that these translations are preset, however, *guarantees* a continuing tendency to stumble over contingent events not foreseen by the tool builders (Collins 1990, 93-105; Dreyfus 1992, 56-7).[28] The narrowly circumscribed shape and predefined relations of the data items in the tool contrast with the ongoing, smooth modifications and reconstructions we witness in the work of medical personnel. In the case of a research protocol for disseminated lung cancer patients, the protocol required some "routine" blood tests after each cycle of chemotherapy. These tests would indicate the general reaction of the patient's body to the therapy. A too sharp decline in the number of white blood cells, for example, would lead to an extension of the pause between the cycles, to give the body more time to recuperate. With one patient, however, who had just had her first intensive dose of chemotherapy, the nurses felt that the patient's "clinical condition" was deteriorating. She looked weaker, had severe diarrhoea, and

was increasingly complaining of pain. At this moment, the physician in charge ordered a CT scan, to see whether the size of her tumors was decreasing: maybe her deteriorating condition was due to the tumor being progressive under therapy. Now, the routine blood tests mentioned in the protocol had lost their significance. *In this situation*, the implied translation "normal bloodcounts means everything is OK" no longer made sense. The protocol builders had not foreseen the possibility of a tumor not even reacting to the first intensive dose - "that is a rare event," one of them later commented. The tool was not "programmed" to anticipate such early progression. It would have just continued its course; the data used as indicators by the protocol were not (yet) deteriorating. It is the medical personnel's burden to *repair* mismatches like this; without their interference, nobody or nothing notes that the preset articulations of sets of data have lost their relevancy.

In a variation of this theme, personnel can intervene to steer data to fit the preset translations. Doing so, they manipulate the tool's input to achieve the output they prefer. Harry, a gynaecology resident in a small hospital, described such a recurring situation. His account accentuates the preprocessing work of medical personnel with regard to medical data:

> For large infants, falling in the upper 10% range [i.e. weighing more than 90% of all newborns], we have a protocol stating that the blood sugar should be measured. We might, after all, be dealing with a missed gestational diabetes. [A large infant may be an indication of this, usually transient, type of diabetes. It is important to spot these cases because such newborns may suffer hypoglycemia, a sharp decline in their blood sugar, leading to flaccidness and even coma]. The annoying thing of this protocol is that once you're in it, you're stuck to it. I mean, Miss B.'s child weighed 4650 grams, which is just in the upper 10% range. But she had had a large infant before, and she's pretty big herself. She'd also been tested for gestational diabetes during her pregnancy - and all these tests were negative. So you just know that that lady doesn't have gestational diabetes, and that there is no real danger of hypoglycemia. But when you then get a blood sugar of 1.8 you're stuck to it - you can't send the baby home. The protocol tells you to keep the child hospitalized for observation: it figures that a blood sugar below 2.0 could indicate a developing hypoglycemia. They're irritating, at times, these protocols. But the thing here is that you can feed such an infant some additional glucose before the blood measurements are taken, assuring, generally, values above 2.0. And nurses know that too...

Fluid Criteria

Formal tools operate through processing preset, definite rules. In medical work, on the other hand, criteria are fluid: they are (re)adjusted to concrete situations, and are only explicated when the smooth flow of work is halted. Such explications, moreover, are always situation-specific: their relevancy for a current situation can and often has to be reassessed. In disciplining a practice to a formalism, through e.g. implementing bureaucratic hierarchies or materializing the tool's demands, tool builders may attempt to fix these assessments and broaden the range of unequivocal relevancy of a tool's rules. Here again, however, these attempts either fall short or are frustrated by fortuitous events. The codifications inscribed in decision support tools, then, may cause at least two types of problems. First, the tool may rigidly apply a criterion in a situation where medical personnel

would typically subordinate the rule to the practical exigencies of the actual situation. In the case of a protocol requiring a complex test at an inconvenient moment, for example, the physician in charge (at the top of this specific hierarchy) simply decided to request the test a day earlier: "I think we can adjust that a little," he remarked. Similarly, recall Mr. York, who was transported for X rays and other tests while the protocol involved already demanded "isolation measures." Medical personnel resolved the conflict, and thus kept the protocol functioning: the tests themselves were a key part of the protocol. Keeping the tool on track, here, necessitated violating it at the same time.[29]

Second, tools inevitably end up in situations where it just is not clear whether a rule applies or not. This is often the case when it has to be decided whether a patient is eligible or suitable for the tool. Most oncology protocols, for example, work with staging tables to determine the extent of disease. For lymphoma's, the so-called Ann Arbor staging has as stage II:

> Involvement of two or more lymph node regions on the same side of the diaphragm or localized involvement of an extralymphatic organ or site and of one or more lymph node regions on the same side of the diaphragm.[30]

In a lymphoma patient presenting with a "firm infiltrate" in the skin above the chestbone, the discussion arose whether this was a "localized" involvement of another organ (i.e. the skin), or whether this indicated "diffuse" involvement. In the latter case, the lymphoma would have to be called stage IV (i.e. "a worse case"), and different treatment protocols would apply. What "localized" or "diffuse" means, in these instances, has to be (re)negotiated in the light of the specifics of the situation at hand.[31]

Similarly, when the Coronary Care Unit would be full, ACORN's rules would suddenly become highly problematic because the context of their use would have been altered. Should patients then be sent to other hospitals? Or would maybe the general Intensive Care Unit, or even the ordinary Cardiology Department be sufficient if this need arises? How should ACORN's output be interpreted in this case?[32] The next fragment also illustrates how a written rule can suddenly become highly problematic:

> At a Tuesday patient meeting, Bear presented Mr. Pierce, who had had a bone marrow transplantation treatment for a "relapsed Hodgkin." "He has had a rapid recovery after having his graft back [his frozen marrow cells, reinfused three weeks ago]. Last Friday he had more than 500 granulos [the number of granulocytes, a type of white blood cell, is used as an indication for the recovery of the immune system; the patient is getting "out of the dip"]. So we stopped the isolation procedures. Now, however, the count is way down again. I don't have an explanation for that. There is no fever. He had an [infection of the scrotum], but that's improving. I stopped all antibiotics yesterday. But his granulos yesterday were around 250, so I'm a little bit scared."
>
> "More than 500 granulos," I had learned, was a rule used to end the isolation procedures. When this level was reached, the "door" would "open," as they said, since it meant that one could now freely enter and leave (including the patient him/herself). The time of the "closure" of the door, *before* the "dip," was simply set at a fixed time relative to the start of chemotherapy. I thus asked whether they would now close the door again; whether they would restart the isolation procedures.
>
> "I was hoping nobody would ask me that," Bear said, and smiled. "Don't you have anything else to do?." Everyone laughed, but the bacteriologist, present also,

was nodding approvingly; he agreed with the suggestion. "You propose to start isolation again...." Bear looked at me. "Give me reasons...." An intern joined in: "Well, there is an increased infection risk again, isn't there? I mean, that's the point of this rule in the first place?" John, the resident, started nodding too: "that's logical, that's logical." "But it is all an arbitrary limit, right?" the bacteriologist asked. "Yes," the intern replied, "but it's the rule we work with here." Bear joined in again: "We just use the granulos to open the door, not to close it." Everyone laughed again. He continued: "Maybe we can give him GMCSF [a drug stimulating blood cell recovery]." The discussion shifted to whether or not that would be of help. They decided to start precautionary antibiotic treatment again, but took no other action. "Let's wait for today's results," was John's final remark, before they went on to the next patient.

Pierce's bloodcells slowly rose again, but remained beneath the "500 granulos" level for at least two additional days. Nevertheless, the isolation measures were not restarted.

What the "500 granulos" rule *means* now that the granulocytes first have opened the door and now start to act unexpectedly is an open question. As the fragment shows, Bear seems inclined to keep the door open. To him, starting precautionary antibiotics is "good enough" for now; closing the door would imply reoccupying the special isolation room, both blocking the treatment of other patients and potentially upsetting Pierce. No explicit "decision" is made on this matter: as often happens, further action is postponed. John does not raise the issue again, and the door remains open.

Anticipating Futures

The fixed, step-by-step character of the tools often conflicts with the more fluid, ongoing flow of medical action. Whereas the tools demand the completion of a step before a next one can be made, staff members often anticipate next-steps-to-take, and adjust their current activities accordingly. In the account about the lung cancer patient receiving a CT scan (above, p. 128-29), the protocol called for a roentgenographic evaluation of the effects of the therapy only after the whole treatment schedule was finished. In the account, however, the physician anticipates a failure of the tumor to react. He performs the evaluation immediately - which, in the end, led to the patient being taken "off" the protocol. In another situation concerning the same protocol, the physician in charge had altered the treatment scheme: the second course of chemotherapy would now be given at "day 15" instead of "day 16." Upon my query, the physician said:

> This woman seems to have long dips [i.e., her immune system recovered slowly from earlier courses of chemotherapy]. If we'd keep the second course at day 16, the third course would fall right into the dip - and we don't want that.

Here, the protocol's sequencing is tinkered with since the physician anticipates problems with subsequent steps.

The following quote by de Dombal illustrates this same phenomenon:

> when the ward is in shambles, for instance, when it is overcrowded, or when the patient cannot communicate, or when a resuscitation is going on in the bed next to the patient, he or she is not entered [into the computer]. (interview July 16, 1993)

While medical personnel continually readjust the sequence, relative importance and priority of the different steps to the situation at hand, the tools neglect the fact that in a more hectic situation, a briefer investigation will have to suffice.[33] Whereas ACORN, similarly, does not proceed without an ECG being taken, medical personnel may improvise on the spot. When the machine is occupied, or gives unintelligible output, staff members can proceed with other parts of the examination; when need be, they can already call the Coronary Care Unit if they anticipate the ECG showing an infarction. Anticipating tensions between the tools' requirement that all data are elicited and the practicalities of the situation at hand, staff members can decide to circumvent the tool, or adapt it to the altered circumstances.

The Real-Time Work of Making Tools Work

A "sufficiently formalized" practice, thus, is not merely the accomplishment of the tool designers. Creating the conditions for the tool to work, is also a real-time, local accomplishment of the personnel working with the tool. Formal tools function because medical personnel fill in the gaps between the tools' prerequisites and the inevitable occurrences in the ongoing work where the disciplining falls short. They preprocess data, repair output where the tool has gone astray and re-interpret the preset rules whenever unforeseen situations occur. They omit items, deviate from the prescribed path, or even *play upon* the tool: purposefully manipulating the tool's input (the blood sugar values, in the account of the gynaecological intern) to satisfy *both* the contingent, practical exigencies of the situation at hand *and* the demands of the tool. They tinker with organizational settings if necessary, or ensure the patient's compliance - they are everywhere at once.[34]

Just what this means for the positioning of the tool in a practice, and of staff members vis-à-vis the tool will be explored in the final part of this chapter. First, I look at how the incorporation of a tool in a practice in fact leads to a certain increase in medical personnel's work instead of "simplifying" it, as the tool builders promised. Second, I discuss the localization this positioning implies. In the last section, I focus on the resulting arrangement of tool and staff members - an arrangement which cannot be captured by a simple model of "who is in control."

Additional Responsibilities

In one of the examples discussed, a physician translated pain "radiating to the back" into a category available on de Dombal's form: "pain in the flank." This seemingly trivial example reflects an issue which is of vital importance to an understanding of personnel's work with decision support techniques. This physician not only had to translate the patient's complaint into a symptom the computer understood, *he simultaneously had to judge the consequences of this translation*. Staff members have to estimate the effects of their actions in the light of the tool's subsequent performance.

Such judgments are highly consequential: mistranslating a data item might harm the patient by setting the tool off on a wrong track. A symptom like "a pain radiating to the back" has diagnostic significance, according to most textbooks: both acute pancreatitis and acute cholecystitis can be accompanied by such pain.[35] The physician has to judge, then, whether by translating this complaint as "a pain in the flank," he might end up feeding this tool information which would lead it to miss a crucial diagnosis.[36] Similarly, by intervening in the protocol's course of action the patient's well-being can be at stake. The nurses who were wheeling Mr. York around while the protocol demanded "isolation measures" had to judge whether this action would not lead to infections when the protocol's next step, the chemotherapy, would hit York's immune system. Likewise, by doing a complex test a day earlier on a more convenient time, exactly those effects which that test set out to screen might be missed. It was such potential consequences which upset the patients working with GLADYS: not being able to oversee the repercussions of the continuous translations they had to make was thoroughly disquieting. How could they know that *their* understanding of "heartburn" matched the machine's? Might they not, by accidentally misunderstanding a question, end up with a wrong diagnosis? For each and every repair, streamlining, deviation, and so forth, those working with the tool have to judge these interventions as to their adequacy in the light of how they might affect the subsequent diagnostic and/or therapeutic steps of the tool.

The consequences of tinkering with the tool are not restricted to the impact the tools' following actions might have on the *patient*. Since the tool's implementation in a given practice thoroughly restructures that practice, the consequences that have to be anticipated are as heterogeneous as the practices themselves. Interventions might generate troubles for medical personnel dealing with the tool later, or elsewhere. Advancing the date of a screening test in a protocol, for instance, should not lead to timing problems later on in its course. Also, consider the following remarks from one of the nurses taking care of the deteriorating cancer patient mentioned above (p. 128-29; the remarks were made just after the CT scan had been made, but before the decision came that the patient would be taken off protocol):

> Her clinical condition is worsening. Her diarrhoea is severe, she's got a fever, and her pain is intense. Now the protocol says that she needs to go to East Hospital [a half hour drive] tomorrow for radiation treatment. But I don't think we can do that to her. And it's a pain for us too: with her condition, we'll have to send along an additional nurse in the ambulance. We're short on staff here - and they might send her back anyway because in this condition she's probably not even suitable for radiation treatment. (interview Carl, January 6, 1993)

The decision to intervene or not intervene in the protocol's course, Carl makes clear, has consequences for both the well-being of the patient and for the organization of a days' work on the ward: both will have to be taken into account in figuring out how to tinker with the protocol.

Indicating "pain in the flank" on a form while the patient had complained of "pain radiating to the back," in addition, might be construed later as a medical error. Similarly, manipulating a protocol's input may

infuriate one's superior. Here, it is not so much the effect of the tool's further workings on the patient which has to be taken into account, nor an increased workload for colleagues. The tool is also often a source for continual and retrospect inspection and judgment of the adequacy of the staff's work. It can function as a document for "'making evident' ... the rational and standard character of their actions" to which superiors and colleagues (and maybe even lawyers or administrators) orient themselves (Whalen 1993; *cf.* Zuboff 1988, 315-61). Hence, in tinkering with a tool's input or repairing output, personnel frequently keep this potential role in mind as well.

In dealing with the tools, thus, personnel not only have to avoid detrimental consequences of the tool's rigidities on the situation at hand. In the articulation of the tool's demands to the medical work of managing a patient's trajectory, personnel cannot only be concerned with the immediate impact of the articulations construed for the present situation of the patient. In addition, *they have to continually gloss their activities in terms of how this intervention will have its repercussions within the formal system* (Suchman 1993b). In the words of Agre (forthcoming), the formal representation of the staff's workpractices is not merely something posited by tool builders and ascribed to staff's activities. Rather, this representation becomes a concrete, active element in the activity itself, whose reactions to all the tinkering mentioned continually have to be anticipated. Telling ACORN that no fever has occurred since an inflamed toe could never be significant for the tool's purpose requires a judgment as to that purpose - which requires knowledge of the inner workings of the technique.

The burden of this additional responsibility becomes even more salient when one realizes that glossing one's actions in terms of the tool's logic is no easy matter. The inner workings of the tool are more often than not (partially) opaque to the staff members: personnel are frequently not informed about the background of the tool, where it is coming from, or what it is trying to achieve. More fundamentally, I have argued that a certain opacity is the inevitable outcome of the negotiation processes of which the tool is the result. Since it results from an eclectic mix of logics - amongst which their expertise is only one - the tool's functioning is often incomprehensible for medical personnel. In a related vein, Weizenbaum has pointed at the fundamental opacity of modern computer programs. Since these programs are the end result of a protracted, collaborative process of interactively adding program-fragment to fragment, there is no one point from which an "oversight" of the whole program is realistically possible (1976). Nobody "knows," to the core, what the program does; even the designers of ACORN, with its relatively simple rule base, are sometimes surprised by the systems' behavior (interviews Wyatt, Emerson).

Due to its literal black box character, this issue is particularly salient for computer-based tools. A protocol is always completely explicated on a piece of paper. Any staff member with a knowledge of the protocol's domain can at least form some judgment as to the protocol's purposes and potential bottlenecks - something which, in fact, occurs all the time. Even so, complete transparency is generally only achievable (if at all) for the protocol designers themselves. *Was* it important to get Mr. Pierce back into isolation since his granulos have fallen beneath the "500" limit? Nobody

really knew; it was an unfamiliar situation for all those involved. *Should* the lymphoma patient with the patch of hardened skin have been included in a protocol for Stage I-II patients? To be able to truly judge this, one should have known all the intricacies, effects and side-effects of the protocol's treatment schedules, its relation to skin-involvement, and so forth.

Once part and parcel of a practice, thus, formal decision tools both create a need for ongoing repairing and preventing work and a pressing necessity to continually oversee the repercussions of this work. Due to the impossibility of fully comprehending the inner workings of such tools, a continuous, tentative "groping in the dark" as to the consequences of these deviations and repairs is inevitable. While tasks are redelegated from personnel to tools (see also further), this does not lead to an unequivocal simplifying or "streamlining" of personnel's work, or to a doing away with the "information overload." The situation is much more complex. While tasks are redelegated from personnel, the tool at the same time *adds to* their work by generating additional, intricate responsibilities. The fact that medical personnel have to continually achieve the tool's prerequisites nullifies any hope for a simple, unidirectional redistribution of tasks from burdened staff members to decision support tools.

Localization Revisited

Faking data or disobeying a tool's output, improvising to grasp the inner logics of the tool: these activities seem at odds with what the tool builders were trying to achieve. In a way, the tools seem to be saved by the same qualities that they set out to erase: the personnel's tinkering would certainly not appear to be "scientific" from the point of view of the tools' discourses. This would, rather, look more like the portrayed characteristics of the irrational, cognitively overloaded physician causing medical practice's problems. As Gasser phrases it, however, "far from acting irrationally, the informal practical actions of participants actually make systems *more* usable locally" (1986, 222). Seen from the perspective of medical personnel, the tool appears as an *additional element* in their ongoing, situated work - an element which is continually reappropriated for purposes at hand.[37]

Tools which "work" are those tools which have merged into the local routines; which have become part and parcel of the messy, ad hoc work of medical practice. Within limits, their actions can be anticipated by staff members: they have become partially "transparent" to the users.[38] Local theories come into being about how to deal with lymphoma patients with skin lesions, or how to get "radiating pain" on the form, or how to cope with a full CCU. Flexible interpretations of the tool's output arise: when the bloodcell counts seem to recover well, a next chemotherapy course is started even if the protocol's lower limit is not exactly achieved yet. These local theories are tied to local notions about the (non)functioning of the tool, about when it should be used and when it should not, and so forth. Such "custom fitting" is an invariable feature whenever a tool becomes embedded in concrete medical practices: the development of "procedural dialects" is ubiquitous given the incessant stream of conflicts between the

practices' concerns and contingencies and the tool's demands (Jordan and Lynch 1992).[39]

The tool does not become the ideal-typed, central decision maker or rational pathway it was proposed to be, to which the personnel merely have to feed data, and maybe fill in some details. The tool does not do away with the ad hoc, heterogeneous work of managing patients' trajectories. This is not a deplorable and preventable outcome of the "corrupting" processes of getting a tool to work: it is the only way for the tools to work in the first place. Getting a tool to work *requires* leaving medical personnel the leeway to digress from the tool's prescribed steps, to skip or skew input - or to sometimes just avoid the tool completely (*cf.* Lipscombe 1989; Star and Griesemer 1989). It requires allowing medical personnel to subordinate the tool to their ongoing work - managing patients trajectories. It requires that the tools become part and parcel of local work routines. It requires, thus, a further *localization* of the tool: a moving away from its ideal-typed universality and uniformity.[40]

A Distributed Locus of Control: The Transformation of Medical Work

For the same reasons that the tools do not unequivocally "simplify" medical work, the tools cannot be said to simply "guide" personnel in their work. The tools are reappropriated: worked around, circumvented, fooled, repaired - all of which require much (anticipatory) skill at the system (Gasser 1986).

This reappropriation, however, is not a matter of shifting the locus of control back from the tool to the staff members, as some critics would have it.[41] It is not a matter of reshifting all responsibility back from the tool to the doctor. There are two complications to this picture. First, as I have argued above, the idea of "total transparency" is an illusion: the tools' inner workings inevitably contain some opacity to the personnel's reappropriating efforts. Since the personnel's grasp of the tool's logics remains incomplete, since transparency is always partial, a simple delegation of control back to staff members is never achieved.

More fundamentally, arguing to shift control back to medical personnel ignores the tool's position as an active agent within a medical practice. We cannot conclude from the tool's inability to stand on its own that the work of personnel does or should unequivocally regain command. We cannot be content arguing that the formal representation is an impoverished version of "what really goes on," since *the formal representation has become a highly consequential part of "what really goes on."* The protocol, for instance, allows articulations of activities over different sites and times. Nurses find leads when to do which laboratory test, and when to shift from one chemotherapeutic drug to another. Likewise, through the protocol the radiotherapist can know when what is expected of her, and how her actions fit in the overall picture. Through the making of detailed "lists," actions and events in different spaces and times are brought together and coordinated (Goody 1977; Star 1989a; Callon 1991). The protocol functions as a focal point of reference, a common resource, to which different staff

members refer, can orient themselves, and can find clues on what to do next.

Formal tools, moreover, can fulfil these functions *because* they are "impoverished" versions of "what really goes on." Only through this "detachedness" can they hope to be transported from the situation in which they are produced in the first place; only through this packaging can they *be* in locales different in space and time. Only through simplification, through reduction of complexity, deletion of details, can workable tools emerge. They fix a history of negotiations and allow (or require) medical personnel to *work from there*. It is exactly through this fixed, black boxed character that they allow personnel to incorporate accumulated past work into the present (Wood 1992).

Through personnel's work to keep the formalism functioning, thus, the tools create a new world. Through the tool transforming, and becoming part and parcel of, a local network, a powerful hybrid emerges in which the interlocking of physicians' and nurses' work with the tool's functioning results in wholly novel workplaces. For one, by taking coordination tasks out of the hands of the medical staff, a protocol may increase the overall complexity of activities. The highly complicated oncological treatment schedules, for instance, are only *possible* through the protocol's core role: its *coordinating* function makes the elaborately sequenced chemotherapeutical combinations and the highly differentiated diagnostic schemes doable. In addition, such a tool can create comparability of activities over time and place - both a core function of the research protocol and of those protocols that are intended to "erase variations" between practices. A coordinating tool, in other words, is a means to *increase* the complexity of a local network, and/or to *extend* this network in time and space.

Other configurations are possible too: in the case of tools which attempt to create a single-moment intervention (as an expert system or a statistical tool), a different set of tasks is delegated to the tool. Here we witness more of a single-moment assimilation and processing of data resulting in a new piece of information, or advice, about the patient. The network can be extended here as well, but not so much in time and space as in *depth*: what used to be handled as a whole array of data is now available as a single statement.[42] Computer-based statistical tools or expert systems can fulfil elaborate functions in amassing, assimilating and continually monitoring the avalanche of medical data incessantly being produced. In this way, ACORN attempted to make it possible for nurses to decide on admittance without having to consult the physician: the tool allowed nurses to perform a task they could not perform before. As with coordinating tools, thus, *accumulating tools* can make complexity doable. Similarly, accumulating tools can be a means to render the processing of sets of data in different practices more alike. In the case of de Dombal's system, the tool drew on a standard set of rules to assimilate some thirty-some data items into one, prognostic statement.[43]

In becoming a participant in the practice it represents, thus, the representation transforms this practice: new decision criteria figure, more elaborate treatment schemes can be dealt with, more actions spread out in time and space are coordinated, and so forth. We see, thus, a reappropriation which is partial also in the sense that in reclaiming some of the control,

staff members make the tool's controlling them - transforming their world - possible in the first place.

Who, now, is controlling what? Or what is controlling whom? We end up in a situation where control is *distributed*, in a novel way, among tool *and* personnel. There is no one person or thing in control - there is, rather, a hybrid, a powerful amalgam of heterogeneous elements, through which control is dispersed in intricate ways. Only *through* the protocol can medical personnel achieve the orderly administration of oncological therapies; only *through* the personnel's work to achieve the formalism's demands, similarly, does the protocol function in the first place. Their is no one person or thing to be held "responsible": only *through* ACORN can nurses admit patients to the Coronary Care Unit, but only *through* the nurses repairing in- and output can the tool fulfil this role.

Attempts to shift this balance inevitably lead to a disruption of the flow of activities in the (transformed) workplace. Granting situational control to ACORN leads to chaos at best - to queues in front of the CCU when the latter is full; to patients torturing themselves to answer all the required questions in the required format; and to erroneous decisions. Arguing, on the other hand, that situational control belongs (or should belong) "wholly" to the physicians or nurses overlooks the fact that practices *are* changed by the tools, and that these changes only occur through the delegation (and subsequent transformation and loss of transparency) of tasks from staff to the tool.

Notes

1. My field experience in medical practices started when I worked as an intern. In addition, I spent two months on an oncological ward specifically studying protocols-in-use. For the other tools, see Chapter 3, note 2.

2. For a study of the nature of tool building (in this case, the construction of an expert system) see Suchman and Trigg (1993). The laboratory studies of e.g. Knorr-Cetina (1981); Lynch (1985; 1991); Latour and Woolgar (1986) and Fujimura (1987) have strongly shaped the development of the first few sections of this chapter.

3. On the links between cognitive psychology and a positivist epistemology, see Lave (1988) and Gigerenzer and Murray (1987) ; see also Chapter 2 above. See also the work of Forsythe, focusing on the inscription of positivist interpretations of "work," "information," "knowledge" and "explanation" in medical expert systems (1992; 1993b; 1993a). This critique has its analogies in many fields - see, e.g., the influential early criticisms of policy decision making models of Braybrooke and Lindblom (1963) and Allison (1971); the critical appraisal of economic decision analysis by Ashmore et al. (1989), the analysis of automated administrative systems in concrete work settings (e.g. Gasser 1986), criticism of automation in the military (see Edwards forthcoming, MS112-3, 150-1), and so forth.

4. See e.g. Lipscombe (1989), Hartland (1993a; 1993b), Dreyfus and Dreyfus (1986, Chapter 4), Gordon (1988).

5. Lave talks about "problem management" as opposed to "decision making" or "problem solving" - both of which have strong cognitivist connotations (1988).

6. Bone marrow transplantation can be performed with a donor's marrow, in which case it is called allogenic. In this case the marrow of Wood himself is used: a so-called autologous transplantation.

7. For Strauss et al., the notion "illness trajectory" includes the course of the illness, the total organisation of work done over that period and the impact of that work on those involved (1985, 8). See also Silverman's concept of "site" (1987, 21).

8. Anamnestic information is information obtained through questioning the patient, also called the patient's "history." Studies showing the constructed nature of anamestic data are e.g. Young (1981a); Cicourel (1986); Davis (1986); Silverman (1987); Helman (1988); A. Wynne (1988); Fisher and Todd (1983); Måseide (1983). For studies depicting medical practice as a locus of constructive *work* see e.g. Bloor (1976; 1978); Lynch (1984); Nyce and Graves III (1990), Hirschauer (1991), and the studies reported in Casper and Berg (1995), and Berg and Mol (forthcoming).

9. One could object that these physicians "do not listen well" or "are biased towards a presupposed diagnosis." Such an objection, however, still assumes that historical data are "there to be found" for the physician who listens properly. As stated, however, historical data are not "givens" but *constructions* which come about during the doctor-patient contact. In that setting, the historical data take shape for both physician *and* patient (see the studies quoted in the previous note). After the consultation, the pattern of symptoms has crystallized in a new form for both participants.

10. On the *production* of realities in medical work see e.g. Hirschauer (forthcoming), Cussins (forthcoming), and Mol (forthcoming).

11. See Berg (1992) for a more detailed exposition of this argument. See Lynch (1985, 202-273) for a superb account on how accounts of objects are continually modified in laboratory "shop talk."

12. See Chapter 3, note 26.

13. Garfinkel phrased it powerfully: "For the practical decider the 'actual occasion' as a phenomenon in its own right exercised overwhelming priority of relevance to which 'decision rules' of theories of decision -making were without exception subordinated in order to assess their rational features rather than vice versa" (1967, 13, 73-75). For similar accounts of rules and rule-use see e.g. Zimmerman (1970); Heritage (1984, 120-9); Knorr-Cetina (1981); Lynch et al. (1983); Harper and Hughes (1993). See also Wittgenstein (1958) on rules; *cf.* Baker and Hacker (1985) on Wittgenstein's views and Lynch (1993, 159-202) on how Wittgenstein's (and ethnomethodology's) views on rules should affect science and technology studies. On the metaphor of "fluids," see Mol and Law (1994).

14. *Cf.* Star (1989a); Forsythe (1993b). Clinical decision analysis, of course, does take the patient's preferences into account (see Chapter 2).

15. Rees (1981, 57), quoting G. Smith's unpublished PhD thesis *Ideologies, beliefs and patterns of administration in the organisation of social work practice*, University of Aberdeen, 1973. See also Atkinson (1981).

16. Whatever a person and/or his/her environment perceives to be a problem for which a doctor should be consulted.

17. On the notion of "situated action" see Suchman (1987). See on this term also Heritage (1984); Button (1993); Lave (1988).

18. On the illusion of the "true image" see e.g. Mol (forthcoming).

19. See Timmermans (1993) for the notion of "crystallization points" in the ongoing flow of medical work. See also Bowker (1994b, 151-2).

20. Nurses play a crucial and often subtle role in this process - see Hughes (1988). The continuing occurrence of contingent, interfering developments and the complexity of this articulation work is further enhanced by the fact that many patients have more than one affliction - and thus more than one illness trajectory, and all the organizational (and often crosscutting) work that goes with that.

21. These formulations are after Lynch (1985, 235) and Knorr-Cetina (1981, 34).

22. Arguing that this "ad hoc" character of medical work is a secondary or only a superficial "stain" on the fundamentally Rational character of medical decision making "is very much like complaining that if the walls of a building were only gotten out of the way one could see better what was keeping the roof up" (Garfinkel 1967, 22; Garfinkel and Sacks 1970, 322-44).

23. On the "in-course" accomplishment of a trajectory's itinerary, see e.g. Luff and Heath (1993); Knorr-Cetina and Amann (1990); Suchman (1993a). For similar comments on the

notion of "decision" see the studies of decision making in neonatal intensive care units of Frohock (1986) and Anspach (1993).

24. It would even be mistaken to locate this "cognitive" reconceptualization of the patient's case in the physicians' mind. On "cognition" as a social process see Woolgar (1989); Coulter (1983; 1989); Amann and Knorr (1989); Young (1981b; 1981a); Latour (1986).

25. See Latour (1988, 164-5). This point draws strongly on Garfinkel's discussion of the "documentary method" (1967, Chapter 3). On the temporal structure of medical work see Zerubavel (1979). See Hunter on doctors' narrative structuring of patients' cases (1991).

26. I focus on breakdowns since this is where the issues at stake are most prominently visible - *cf.* Garfinkels' methodological recommendations (1967, 36-8). Reinterpreting the moments portrayed as "anecdotal exceptions to the rule" misses the (methodological) point that the empirical case fragments are not what "proofs" the argument. They *illustrate* the argument, and *clarify* it - the "proof," however, is in the persuasiveness of the argument itself. See also note 22.

27. Misunderstanding (in both directions: the patient misinterpreting the computer and vice versa) easily occurred - and this was upsetting to patients who have no clue of what the consequences of such misunderstanding might be. Patients could also become upset when the computer would *not* ask a certain question. A woman became distressed after the interview was over, saying that "it doesn't ask you if you have had gallstones. I have had my gallbladder out, but it didn't ask me that, so this interview doesn't show what I am really like" (Hartland 1993). See Suchman for an analysis of similar troubles in human-machine communication (1987, 118-177).

28. See Collins (1990, 236) for a hilarious but illustrative thought experiment about an expert system dealing with acute pain complaints (say ACORN or de Dombal's tool) confronted with a patient with a pointed weapon embedded in his/her body. The tools are not programmed to ask "Is there anything stuck in your back"? - they will just run their preset questions and generate one of their preset diagnoses. It is the medical personnel's task to "repair" errors like this.

29. "Protocol violation" is an expression used by the medical staff. Decisions whether an event *is* a violation, and, if so, whether it is "major" or "minor," are continually interactively established (*cf.* Bosk 1979; Lynch 1985).

30. These staging definitions are generally included as an appendix in oncological research protocols.

31. In such situations, Suchman states, these rules "pose problems of interpretation that are solved in and through the objects and actions to which the instructions refer" (1987, 142). See also note 13. The focus on "negotiated interpretations" is a core tenet of the "controversy" studies in the sociology of science, as exemplified by e.g. Collins (1985) and Mulkay (1988). See also Woolgar (1988).

32. Likewise, personnel working with a system evaluating measurements in patients with peripheral vascular disease (Talmon et al. 1987) ran into troubles. Ordinarily, they would deem blood pressure measurements unreliable whenever they were "too high"; something, they explained, which could be due to the bloodvessels being arteriosclerotic. Since the system had no specific provision for such situations, entering these measurements into the system would obviously set it off into a wrong direction. Not entering anything, however, would be interpreted as entering "zero," which would be equally wrong. The staff involved was at a loss: they had no clue how to deal with the tool in such situations; how to reinterpret its demands.

33. A similar phenomenon would have plagued MYCIN. Buchanan and Shortliffe (1984, 505) state that "MYCIN's control structure is not concerned with resource allocation; it assumes that there is time to gather all available information that is relevant [it would ask 20-70 questions, MB] and time to process it."

34. In Star's terms, they are "tall thin people" (1989a; Star attributes the metaphor to one of her respondents). "Playing upon the tool" includes feeding in fake data, manipulating data in some indirect way (as in the story of the gynaecological intern), or accepting error and repairing it later on, and so forth. Gasser's study of automated administrative work (1986), Collins' study of expert systems (1990) and the work of Star (1989a; 1991) have been influential in the formulation of these ideas.

35. See for example Boey and Dunphy (1985, 392).

36. As an obvious but illustrative example, consider the fictitious patient mentioned in note 28 entering the emergency department. The tool's diagnostic output is limited to seven non-traumatic categories (see Chapter 2), so filling in the form and translating "pain in the back" as "pain in the flank" will here ensure the generation of a meaningless diagnosis. On the effort involved in man-machine communication to "align" oneself to the tool, see Suchman (1987) and Wooffit and Fraser (1993). This is an instance of systems "falling of the knowledge cliff" (Forsythe 1993a): the problem of limited scope referred to in Chapter 4 .

37. In oncology wards, for example, research protocols often figure as treatment options to choose from instead of as strict guiding principles which constrain physicians' choices. Rather than judging the "fitness" of a patient for a protocol, then, physicians are considering the fitness of the protocol for the patient. *Cf.* Timmermans and Berg (forthcoming); Star (1989a); Luff and Heath (1993).

38. To paraphrase Suchman: "Just as it would seem absurd to claim that a map in some strong sense controlled the traveler's movements through the world, it is wrong to imagine [formal representations] as controlling [personnel's] actions. ... In the last analysis, it is in the interaction of representation and represented where, so to speak, the action is" (1987, 189).

39. See also Jordan and Lynch (1993); B. Wynne (1988); Barley (1988); Callon (1980).

40. Whether in space, scope or rationale (or any combination of these) is an empirical question.

41. See note 4.

42. In actor-network terms: a part of the network is collapsed into one node (see e.g. Callon 1987). On hybrids and networks, see also Latour (1993) and Law (1994).

43. In addition to making complexity doable and/or enhancing comparability, a coordinating tool may be intended simply to standardize sequences of activities to some preset ideal (see the chronic lung disease protocol in Chapter 2, Figure 8). Likewise, an accumulating tool may ensure a standardized processing of a set of data so that similar sets always result in the same output statement. In the case of de Dombal's tool, for example, the tool articulates a certain group of input items by means of a standard process which, it is hoped, yield optimal results.

6

Producing Tools and Practices

In 1970, Schwartz predicted in the *New England Journal of Medicine* that computer-based decision tools would "fundamentally alter the role of the physician, and profoundly change the nature of medical manpower recruitment and medical education." He foresaw a revolution: in the year 2000, many a patient's history would be taken by the computer.

> ... [I]n the not too distant future the physician and the computer will engage in frequent dialogue, the computer continuously taking note of history, physical findings, laboratory data, and the like, alerting the physician to the most probable diagnoses and suggesting the appropriate, safest course of action. (Schwartz 1970)

This glorious future has not come about. Seventeen years later, Schwartz and others look back and have to conclude that "the revolution has not occurred" (1987): computer-based decision tools in routine use are surprisingly rare. Many Bayesian or rule-based systems have been made in these seventeen years, but few have actually been implemented successfully. A list of "Artificial Intelligence Systems in Routine Clinical Use," compiled in 1993, mentions only 18 systems. Most of these, moreover, are used in only a small number of practices; many have never left their site of origin.[1] MYCIN, for example, has attracted much attention in the (medical) press - but it has never been put to practical use. Nor has INTERNIST, another famous rule-based system, which was designed to handle the broad domain of internal medicine.[2] One of the systems included in the 1993 list is ACORN, which, in fact, is currently leading a dormant existence. Emerson has taken the system home: amongst others, nurses were not using the system frequently. "The impact of information technology in the diagnostic aspects of medical decision making [has been] limited," Barnett and others argue: they mention de Dombal's system as the only exception (1987). As one computer scientist put it, most systems are "dead."

Decision analytic techniques have met with a similar fate. Here as well, expectations were high:

> In the 1970s, there was optimism that the application of decision analysis to clinical problems would rapidly disseminate. Some envisioned computer terminals in every clinician's office that could tap databanks filled with estimates of key variables used in decision-analysis models, and decision trees were envisioned that could be applied to the care of individual patients. Even television commercials depicted clinicians in white coats moving from the bedside to the computer to take advantage of this electronic network of information. (Detsky 1987)

Unfortunately, this author continues, "this optimistic vision has yet to be realized." As a tool for making decisions on individual patients, decision analysis is hardly ever used. In a nevertheless enthusiastic appraisal of the technical state of the art, Kassirer and others mourn that:

> House officers rarely use [decision analysis] or even request it, even for tough clinical problems; practicing physicians more or less ignore it; and many medical educators are frustrated because their efforts to teach it to students have been wasted by the lack of exposure in the postgraduate years. (1987)[3]

Three years later, an editorial in the *Journal of the American Medical Association* notes that there is still an urgent need for ways to "free decision theory from its restrictive academic base and provide a place for it in clinical reality" (Flanagin and Lundberg 1990). In fact, when I started looking for clinical applications of decision analysis in 1992, the situation seemed even more gloomy: several informants pointed at the Clinical Decision Making Unit in Boston as the *only* place in the world where clinical decision analysis, in its original form, was being used.[4]

Protocols, finally, form a more complex story. Research protocols, for one, are ubiquitous in many specialized fields (especially oncology). As such, however, they are rarely studied: they seem to be a mundane, taken-for-granted part of the huge enterprise of clinical research.[5] This relative neglect, taken together with their ubiquitous presence, was one of the reasons I looked at some of these protocols in more detail.

The impact of national consensus-based guidelines, on the other hand, has remained doubtful. Despite the fact that a whole industry has grown up around the creation and diffusion of such protocols, several studies have shown that physicians are often not even aware of their existence - and if they are, they often report to disregard them (Kanouse et al. 1989; Van Dijk and Bomhof 1991). Moreover, even when physicians report that they have changed their practice in accordance with the guideline, actual practice patterns often show that they have not. Lomas and others, for example, studied the impact of the Canadian consensus guideline on the indications for cesarean section (1989). This guideline tried to reduce the rate of cesarean section by, amongst others, propagating a reduction of operations in cases of breech presentation and in patients with a previous cesarean section. In a survey, 33% of the obstetricians said that they had changed their practices; they reported that, among women with a previous cesarean section, the rate of cesarean sections had significantly declined (for breech presentation, no significant change was reported). Interestingly, when actual practice was measured[6], the number of cesarean sections appeared to be 15 to 49% higher than reported. These measures showed only a tiny decline of cesarean sections for the indications mentioned here - while a decline in operations for previous cesarean section patients had already been present before the guidelines had emerged.[7] "Those who pay for and review medical care," Lomas concludes in a later paper, "bemoan the lack of impact on clinical behavior of practice guidelines" (1993; *cf.* Eisenberg 1986, 103).

Whereas much effort has been spent on evaluating the effects of national consensus guidelines, considerably less is known about *local* protocols (ranging from regional interdisciplinary protocols prescribing the ap-

proach to suspect skin lesions to ward-specific protocols on the proper way to administer a certain drug). What is clear is that, contrary to computer-based tools and clinical decision analysis, local protocols are a common feature of current medical practices. Due to the differences in their "goals, methods, formats, and degrees of precision," however, their overall impact is impossible to estimate (Pearson 1992, *cf.* Chapter 4, p. 102). Moreover, they widely differ in the way they are being adhered to: some protocols are followed by the rule, while others "are available in case somebody does not know what to do," as a surgical resident expressed it to me. All in all, studies of protocols-in-use tend to show much deviations from the prescribed path. In their study of a sore throat protocol, Grimm's team found that physicians often prescribed different types of antibiotics, or did not record data perceived to be "crucial to sound clinical decisions" (1975; *cf.* Sox Jr. et al. 1973). Similarly, Wachtel and others studied protocols that tried to reduce health care costs through limiting unnecessary laboratory testing, drug use and length of stay. They note that although some effect was measured, "the average utilization of services in the intervention group was higher than proposed in the protocols." It was found that "protocols were frequently broken," and that "participating physicians varied in their ability or willingness to enforce strictly the use of the protocols with the housestaff" (1986).

Analyzing the Lack of Success: Of Right Things and Right Places

Why has the computer-aids revolution failed to occur? What explains the lack of success of so many decision support techniques? Both the tools' advocates and the critics pose these questions, but answer them in wholly different ways. After briefly depicting their positions, I will argue that they are *both* problematic. Recapitulating the main themes of this book, I will attempt to point out that the way both critics and advocates frame their analyses leads them to miss crucial opportunities to understand and intervene in the processes described.

Obstacles to the Tools: The Advocates' Analysis

According to its advocates, we have seen, decision analysis is a universal technology: in principle, the technology can be used everywhere, by anyone, on any medical problem.[8] Why, then, has the "permeation of decision analysis into routine clinical practice" been "sluggish" (Kassirer et al. 1987)? There is no single answer:

> Several factors appear to be at work: a large number of highly competent teachers of the method are not available; few practitioners are skilled in its clinical use; computer programs to support decision analysis that are as simple and convenient to use as word-processing programs are not available; time required to complete a satisfactory analysis is costly; and efforts to enhance the "marketability" of the method by targeting specific audiences have not evolved. (ibid.)

What is notable about this list is that most of these issues concern the *setting* of which the tool has to become a part: too few teachers, too little time, too little funds are available. Likewise, other authors argue that psychological obstacles hinder physicians' enthusiasm, as "the traditional reverence for intuitive judgments" (Detsky 1987; *cf.* Böckenholt and Weber), and the fact that decision analysis outcomes can "feel wrong" because the analysis does not share the physician's biases (Balla et al. 1989). And, it is added, more attention should be paid to generating the data needed for decision analysis - data which are often not available (Cebul 1984). These are the obstacles that prohibit widespread deployment of the technique. The techniques themselves are not to blame: the places where the tool is to work are at fault. Although some problems still remain with the measurement of utility, for example, to the advocates these are not fundamental. "[I]mprovements are needed," Balla et al. argue, but this need does not invalidate the current potential of the tool (1989; Holmes-Rovner 1992). Through educating doctors, obtaining more funds, and smart marketing (through e.g. targeting the message to "role-model" physicians), clinical decision analysis might still become a success.[9]

When one looks at computer-based decision tools a similar picture emerges. The overall tone is somewhat more subdued; most articles combine a critique of the stubbornness of their tool's setting with a (mild) self-critique. Again, the tool is universal in principle: current limitations result from obstacles which, with some additional effort, can be overcome.[10] Shortliffe, one of MYCIN's creators, argues that perhaps nothing "accounts more fully for the impracticality of early clinical decision tools than their failure to deal adequately with logistical, mechanical, and psychological aspects of system use" - the so-called "human factors issues." These, he states, are problems of both design and "resolution of inadequacies at the institutional level" - emphasizing both the immaturity of the tool and the setting's shortcomings (1987). Programs should not be too slow, they should not require lengthy typing, and they should use graphical displays to optimize the user-system interface. At the same time, the compatibility of departmental computer systems in hospitals leaves much to be desired, and physicians should start to accept that they do need decision support tools (ibid.; Buchanan and Shortliffe 1984, 599-601).

De Dombal similarly spreads the blame. He criticizes the poor "doctor-computer interface" which limits the tools' efficiency in real-time use (1989). But he also immediately turns around and accuses the practices in which the tools are supposed to work. Even a major improvement in interface design would not help much, de Dombal argues, because the progress of decision support systems is obstructed by the often "totally unstructured" state of medical knowledge. "In many areas there is no agreement about definitions of symptoms and signs, or what symptoms are to be collected - or even what disease categories there are, and how they are defined." Until these problems are addressed by the medical profession, "all the computer scientist [can] do is wait." De Dombal is harsh in his judgment on the absence of initiatives to address this conceptual chaos:

> In part this waiting is due to simple lack of co-ordination and bureaucratic inertia ... This lack of co-ordination makes it impossible to devise systems which will have

> wide acceptance; and is doubly surprising for often it is coupled with a plea from the relevant bureaucracy for more "efficiency in medicine."

Notwithstanding these obstacles, his analysis also ends in an optimistic tone. The challenges can be overcome, he argues: "it seems only a matter of time before the use of decision support systems becomes more widespread" (ibid.).[11]

The small impact of national consensus-based guidelines, finally, is likewise accounted for. In their evaluation of the NIH consensus development program, Kanouse et al. state that some guidelines suffered from excessive vagueness and "left important concepts undefined" - like how "high risk patients" are defined (1989, xxvi, xxxi).

> We found that we often needed expert medical consultants to help us interpret what the panel's recommendation would mean in clinical practice, and that the interpretation we arrived at was sometimes not the only one that might reasonably be made. (ibid.; *cf.* Kahan et al. 1988)

Here again, however, the main reason for the disappointing effects were the several "administrative, educational, patient-centered, or economic barriers" to the implementation of guidelines (Lomas et al. 1989). Through educational channels, popular press coverage, publications in major journals and direct mailings, the guidelines should be much more widely distributed, Kanouse et al. argue (1989, 65, 244-7). Simultaneously, additional strategies are necessary: to get physicians to change their behaviour in accordance to the guideline, Lomas argues, drawing upon tactics as "peer pressure, marketing of information and exploitation of informal communication channels" is indispensable (1990, 183).[12]

Tools Which Cannot Work: The Critics' Analysis

The tools' advocates explain the lack of diffusion of the tools primarily by pointing at resistances in the medical practices of which the tool should become part. The present limitations are a matter of fixing some remaining obstacles: in principle, the tool has a universal reach. It is just blocked, here and there, by practicalities such as insufficient funds, stubborn physicians and bureaucratic obstinacy.[13] The technologies themselves have their problems too, but these can be remedied by advancing on the current path of development.

The tools' critics turn this argument upside down. They chastise the tools for fundamentally misunderstanding the world in which they are to function. It is not a matter of "fixing" the settings' shortcomings: according to many philosophers, sociologists, and anthropologists, the tools "are all on the shelf" because they are *essentially* misconceived.[14] According to them, we saw in the Introduction, "formal models fail to capture human expertise": decision support tools attempt the impossible. Intuitive, involved, skilled "knowing how" can never be replaced by detached, objective "knowing that" (Dreyfus and Dreyfus 1986, 184, 189); human beings can "act" (i.e., they have intentions, which can be understood by other social beings), where machines can only "behave" (i.e., machine activities

can be exhaustively described by their physical coordinates or by a set of rules) (Collins 1990, 30-61). Where the advocates emphasize the (in principle) universal character of computer-based decision techniques, critics point out that these tools can only operate in very specific, confined areas. Tools can only be "successful" whenever the tasks they are to perform do not *require* skilled "knowing how"; whenever the tool "takes over from us" those actions that we already do "through machine-like actions" (ibid., 61). Moreover, one can try to "capture" social practices in formalisms, but the ceteris paribus condition assures that all such codifications are *necessarily* tied to the specific context in which they were generated (see p. 10-11). The "typical" acute abdominal pain patient of de Dombal's tool is only "typical" of the specific population which visited the emergency ward of that hospital, in that particular time - since that is the population on which the tool's formulas are based. The only "universal" element in the discussion, according to these critics, is that "expertise" is always context-bound. No matter how thoroughly spelled out, knowledge is linked to specific practices, and is therefore always local.

The Production of Working Tools: Re-analyzing "Success" and "Failure"

The advocates have a hard time explaining just why computer-based tools are so good if so few are in actual use. Dreyfus' sneers that these claims sound like people who, having climbed into a tree, say that the moon is near (1992) - and doesn't he have a point? On the other hand, research protocols are ubiquitous in many wards all over the (Western) world. From a critic's point of view, these highly detailed formalisms should not work - but they do. On several crucial points of critique, research protocols are no different from computer-based knowledge systems or decision analysis: the basic, formal structure of the tools is equivalent (*cf.* Reggia and Tuhrim 1985, 3-45). Research protocols attempt to eradicate the contingent flow of reactive activities and replace these with clear-cut, interinstitutionally comparable, standardized actions. They are explicitly intended to be in widespread use, in many hospitals at the same time. Moreover, they are to take the locus of control out of the staff's hands, and to determine by means of a set of explicit rules what should happen given a certain situation.

How, then, is it possible that research protocols have become so widespread? Are the protagonists right after all? Are the critics missing the point? If so, how can we understand the differences in "success" between tools? Or are both analyses on a wrong track?

Critics and advocates phrase their analyses in contrasting terms. Whereas advocates blame medical practices for resisting the tool, critics argue that the problem lies with the tools' misconception of the nature of these settings. Where the former stress the tools' universal problem solving power, the latter see no truly universal feature of these technologies but their impossibility. Underlying this opposition, however, lie some strong assumptions shared by both critics and advocates. The structure of their analyses, the central questions they ask, follow a similar pattern. In attempting to explain the success or failure of decision support techniques,

both critics and advocates debate the *nature* of tool and practice. This mode of arguing renders "Tool" and "Practice" into two fixed categories, with fundamental properties which are to be listed and compared. Both critics and advocates want to judge the actions of formal tools in comparison with human expert action, and query whether the tool is an adequate *representation* of the structure of medical practice, of the physician's decision making, or of an optimal doctor's performance. Both analyses, then, *contrast* tool and practice by reifying them as separate categories, comparing them, and distributing blame for failure, or praise for success.

Critics, for example, state that the computer cannot but fail as a decision support technique since it cannot operate/behave/think like a physician - it lacks the agency, the "intentions," the "tacit knowledge" humans have. For authors such as Dreyfus and Collins, human practice is endowed with fundamental qualities that machines simply do not have: every attempt to "capture" skilled performance is doomed to falter. "Human Practice" is pitted against "Machine Action," and the argument attempts to demonstrate the ontological and epistemological gulf that forever separates Us (humans) from Them (techniques). As Robinson (1994) nicely phrases it, critics "put intimacy and abstraction at opposite ends of the same continuum": in criticizing the formalisms, they emphasize the poverty of these representations vis-à-vis the richness of the actions represented. Whether we are dealing with expert systems or statistical formula which would represent the physician's mental processes, or protocols which would represent a sequence of practical activities: critics expose the a priori failure of a map to capture the intricacies of the domain the map represents.

Advocates also debate the nature of "Human Practice" and "Machine Action." Whereas critics argue that these categories are separated by a deep gap, advocates argue that they share crucial similarities. According to some advocates, for example, both tools and medical personnel belong to the species of information-processing agents; according to others, both tools and practices are structured according to a scientific, step-by-step approach to clinical problems (*cf.* Chapter 2). Here as well, the issue is framed in terms of just how much these categories are alike, and just how well the tools represent the decision making or actions they embody. For example, advocates often attempt to evaluate computer-based tools by setting up the computer versus the human expert and comparing how each deals with a similar set of (paper) cases.[15] Instead of putting intimacy and abstraction at opposite ends, advocates would rather argue that the former can be rendered into the latter without danger; that the tools can adequately represent the physician's decision making or a practice's management strategies. Here, Tool and Practice are not contrasted to demonstrate their difference, but to argue their similarity. Since the reason for failure, then, cannot be found in fundamental properties of either tool or practice, it must lie elsewhere: in current imperfections of the tools, and in contingent, psychological and socioeconomical barriers in the practices.

However, contrasting human practice and tool to compare the two, or to distribute blame and praise, is not a fruitful way to frame the debate. Asking what properties of tools or practices can explain success or failure, is posing the wrong question. In the next two subsections, I argue that starting from this question hampers the understanding of the events de-

picted in this book; in the last subsection, I argue that this framing also reduces the potential to intervene in them.

Producing Tools and Practices: Changes and Convergences

> Nothing is, by itself, either reducible or irreducible to anything else ... there is no equivalence without the work of making equivalent.
> (Latour 1988, 158, 183)

Debating success or failure by pointing to fundamental qualities of Human Practice and Tool overlooks that these qualities *themselves* are not pre-given. By basing their critique on the unique properties of human practices, or, respectively, by positing the existence of the universal category of information-processing species, critics and advocates assume a *foundation* from which to argue Difference (the critics) or Similarity (the advocates).[16] Such foundationalism, however, overlooks that what is taken as "foundation," as "basic category," is *itself* already a product of the historical intertwinement of humans and tools. What is our understanding of "human-ness" *else* than a conglomerate in which a whole range of metaphors (the steam engine, the telephone switch board, and yes, formal technologies) play a role? How can we hope to disentangle "human practices" from the technologies which structure them, the artefacts upon which they feed? And do we not describe, develop and work with computers in terms directly derived from human "properties" and activities such as "memory," "writing," "processing," "messages," "hierarchies," "commands," and indeed "intelligence"?[17] Equivalence between Tool and Practice is (or is not) achieved in and through the transformation of discourses and the disciplining of practices: in and through the set-up of psychological experiments (including the use of paper cases to compare Man and Machine), the explosion of computing metaphors for the "mental," and the transformation of diagnostic and therapeutic tasks. The *nature* of, say, "medical practice," "expert systems," or "physicians," is the outcome, the *effect* of these intertwined transformations (*cf.* Law 1993).

In the early sixties, for example, statistical techniques similar to those used for the first statistical diagnostic tools were employed as an attempt to ameliorate the disordered state of medical taxonomy: to eliminate the "logical fallacies which abound in the subjective, orthodox methods" of disease classification (Sokal 1964, 51; *cf.* Jacquez 1964; Overall 1972). If these approaches had had more impact, if medical taxonomy had been restructured in this way, the statistical tools would have transformed medical practices more to their liking - and they could, then, have been more "successful." Also, through the proliferation of biophysical technologies, monitoring instruments, and so forth, qualitative data are increasingly being replaced by "digital" data.[18] Here again, practices alter and become more amenable to formal tools. On the other hand, more and more computer-based tool builders are adjusting their tools to the resistances experienced in earlier attempts. Tools are built which function more in the background (as "checklist"), for example, or tools which are integrated more into other, larger information technology products (and so built upon infrastructure

already in place).[19] In the processes which have led to these current states of affairs, then, both practices and tools have been transformed.

If this is so, the very question *why* a tool is "(un)successful" becomes problematic. "What accounts for failure or success" tends to be read as a *general* question, requiring a general answer, and resulting in a list of "factors" or "aspects" which, by their general nature, are easily read as qualities of tool and/or medical practice. For example, the widespread use of the research protocol can be "explained" by referring to its rational, scientific character. Or the "lack of success" of expert systems can be denoted as a "matter of time" by pointing at the fundamental equality of the nature of the tool and the nature of medical thinking. Alternatively, the research protocol's ubiquitousness can be accounted for by pointing to the research careers that are tied to the usage of these protocols, to the existence of regional coordinating offices that aid in the proper application of protocols and process and check collected data, and to the many (international) medical research funds that are centered around the development and implementation of research protocols.

To turn "(non-) successfulness" into a property of either tool or practice, however, to explain "success" or "failure" through contrasting the two, is to isolate the tool from its contingent history. It is not *because* the research protocol is so rational that it has become so widespread: the specific rationality and shape of decision techniques, we saw, emerged together with the development of these tools and the specific configurations of their use.[20] Likewise, the tool's setting cannot explain its "success." To the contrary, I would conjecture that protocols have played a significant role in the shaping of both content and context of current oncology: it is probably *because* the emerging protocol has been an ubiquitous element of this field since its early days that the diagnostic categories and therapeutic interventions are as complex and elaborate as they are now.[21] As a final example, arguing that the expert system is a "failure" because medical practices resist it, overlooks how the cognitivist perspective has been central in the very redrawing of the nature and flaws of medical practice in the first place. The questions of critics and advocates pry apart Tool and Practice instead of interrelating them; they freeze the two into an artificial, snapshot-opposition instead of following their intertwined trajectories. There is no meaningful way to talk about a tool without at the same time speaking of the practice with which it co-evolved: the research protocol is part and parcel of the network which *is* "current oncology." One cannot cut away either one or the other without loosing sight of the constituent features of their current "essences." The "properties" of both tools and practices *themselves* are in constant motion. And when nothing has remained the same, when tools and practices have become so thoroughly intertwined, blame or praise for a tool's fate cannot be distributed unequivocally. A story of heroes or villains has lost its ground. "The tool" or "the practice" can no longer figure as explanatory category: an understanding of the current situation requires an understanding of the way these categories are intimately involved in each other's production.

This book, then, focused on how specific tools and practices have evolved together. Where critics and proponents primarily center on the tools as

product, this book looked at the *process* of the construction and implementation of these tools, and of the discourses that came into being with them.[22] The tools and the worlds in which they become embedded thoroughly transform each other - and these mutual transformations are key to an understanding of their (non)functioning. A working tool, I argue, is the outcome of these mutual transformations: of the *convergence* of tools and settings into a network in which heterogeneous elements are interconnected - and are transformed (Bowker 1994b).

There are many interrelated layers to this process. I focused, first, on *medical literature* and described how with the coming of the tools new discourses came into being. These new discourses redefined what medical practice is, what its problems are, and how these problems should be addressed - and did so in a way which *made* the tools "the right ones for the job" (*cf.* Clarke and Fujimura 1992). By redescribing medical practice in the image of the tools, the discourses achieved a smooth fit between tool and practice. In discussions on medical practice and its problems, the tools became obligatory points of passage for many attempts to rationalize medical work.[23]

In the *system builders'* construction and implementation of individual tools, medical practices and tools are made to fit each other. Practices are disciplined to ensure appropriate input and adherence to output. Concurrently, the technique is transformed: in inscribing locality into the core of these tools, they are redescribed in the image of the practices.

In the work of *medical personnel*, finally, the same mutual transformations return. Where the disciplining of a practice falls short, medical personnel can achieve the formal tools' prerequisites in their work. In doing so, they help the tool transform the practice, extend and/or complexify the local network - and the tool is further localized.

All these practices intermesh: I have been describing a complex network of medical personnel, blood cells, journals, decision techniques, system builders, patients, and so forth. This one network can at the same time be seen as the interlocking of several networks, all exerting their own, sometimes conflicting, demands: physicians and nurses are monitoring and shaping their patients' trajectories *and* they are themselves an element in the network built by the tool builders; system builders and physicians are influenced by journal articles *and* they themselves participate in the very activity described in these texts. Shifting the point of entry from medical literature to system builders' work to the work of medical personnel, then, was a means to get a grasp on the different layers that are at stake in the development of decision support tools, and on their interrelations. We can never hope to comprehend these multilayered processes if we attempt to stick to one point of entry for our analysis: the network is not set in motion by one central actor. To get a hold on this multidimensionality, it is a prerequisite to shift positions throughout the network - to attempt to grasp the way the different layers can be tightly interconnected yet have distinguishable dynamics of change.[24]

"Convergence," then, points at the way technique and setting mutually transform each other to each other's image - whether the focus is on medical discourses, the work of system builders or the work of medical personnel. It points, also, to the way these different activities and occur-

rences (including those I have only tangentially mentioned) mesh together into niches for the co-evolving decision support techniques. Only when new discourses have made tool builders' interventions possible; only when medical personnel in real-time maintain the niche tool builders have construed; only when the tool builders' attempts interlock with coinciding practices of standardization - only then can a working tool evolve.

"Convergence," also, points at the simultaneous and similar transformations of "social" and "natural" worlds (Bowker 1994b). The work of medical personnel is rewritten in the light of the tool - and vice versa. Similarly, blood tests have to fit the tool's requirements as much as nurses' behavior. The thoroughly heterogeneous practices are disciplined through equally heterogeneous means - through forms, training, bureaucratic hierarchies, pre-packaged medication, and so forth. Localization, simultaneously, is inevitable since both "social" and "natural" elements can elude control.

"Convergence," finally, points at the fact that characteristics of "tool" and "practice" are not pre-given but rather *emerge* in and through the development and intertwinement of the networks.[25] "Universality," for example, is not a static attribute to be debated as to its actuality or its absurdity, as the tools' protagonists and critics tend to do. It is not so much a characteristic of the tool but a possible *consequence* of extending the tool's reach through disciplining a whole array of local practices, and achieving the formalism's prerequisites. Universality is not some ethereal quality of a superior technology, but an emerging feature of the networks of which that technology has become a part. It is a *dynamic* characteristic which, in the fighting of the different types of localization, is continually in the process of becoming - and continually in the process of slipping away.

Producing Tools and Practices: Different Questions, Different Issues

Instead of contrasting tool to practice, then, I have focused on how tools and practices have developed together and transformed each other. Instead of trying to explain success or failure by opposing two pre-fixed entities, with pre-fixed characteristics, I have gained some understanding of the current (non)usage of decision support techniques by looking at how the networks in which tool and practice co-evolve have converged - or not.

In refocusing the questions in this way, some issues come into view that are neglected by critics and advocates. To advocates, for example, the formal, "objective" nature of their tool guarantees diffusion: it is only the *resistances* that have to be accounted for - and dealt with. To them, the tools are like faithful, objective maps: representations of optimal practice, which lead medical personnel towards supreme performance. They are windows on a perfect world. How could they *not* be universally used, except for a few lingering obstacles (Latour 1987; Wood 1992)? Advocates, however, overlook the *work* which makes their tools feasible - the ad hoc management of (tools' and patients') trajectories which characterize both the work of system builders and of medical personnel. "Disciplining practices to a formalism" is not the taking away of obstacles: it is the *building* of the roads on which the tools can travel in the first place. "The outside world is fit for an

application of the map only when all its relevant features have themselves been written and marked by beacons, landmarks, boards, arrows, street names and so on" - that is, when the landscape is *itself* rewritten in the *same* language as the map (Latour 1987, 254). Once through this (ongoing) work a network has been made stable enough to have a formal tool function, the advocates' views retrospectively *become* true (Bowker 1994b). Once through the co-evolution of tool and practice settings are sufficiently transformed to the tools' liking, institutions ensure uniform terminology, meetings ensure identical laboratory tests and criteria, and technologies are distributed uniformly - then a tool's "universality" is accomplished. At the same time, however, this accomplishment is dependent upon the endless work of committees to standardize tests, secretaries to fill in and check forms, and personnel to fix the breaches between their ongoing work and the tool's demands. In the advocates' view, this prior and ongoing work "disappears into the doneness": it becomes invisible (Star 1989a; 1992). The technology seems to function on its own, superb and universal power - while the work of meticulously creating and repairing the social and material infrastructure which makes this functioning possible (and in which the tool itself has taken shape) disappears from sight.[26]

Similarly, advocates overlook how tools are always *located*; how a local context and reflections of past negotiations are built into the heart of the Rational Tool.[27] There is no breaking away from localization, no transcending of impurity: the advocates' illusion of unqualified universality is the illusion of tools having no history. It is the illusion of the one transparent, objective, true, optimal documentation of the world - of the "unmarked category" which represents while itself "escaping representation" (Haraway 1991, 188). Every decision support tool has its past of struggles, failures, contested choices and omissions carved into its core - every map has an author, a subject, a theme (Wood 1992, 22). Every tool silences some voices and amplifies others; every tool helps to strengthen some knowledges and helps to forget others; every tool is coming from somewhere and going someplace else. And since the tools' actions can be so consequential, it is important for both participants and analysts not to delete its past - not to forget its locatedness.

Both critics and advocates, to continue, remain stuck in *comparisons* between tool and practice. They compare the terrain with the map: does the latter adequately represent the former? Is the image objective or biased, the performance (as measured through e.g. paper cases) comprehensive or impoverished? Framing the debate in this way, however, ignores the crucial matter how the representations are *used* in the practices. Many critics remain stuck in stating that the tools cannot or should not function, or that they only function through human help[28] - but they pay little attention to *how* this changes both content and context of personnel's work. The intriguing feature of these systems is that they alter the work which allows them to exist - but critics do not focus on the coordinating and accumulating functions of these tools. In criticizing the formal tools' poor mapping of human practices, to paraphrase Robinson (1994), they make the mistake to want to *substitute* informality for formality - to demonstrate the powerlessness of the formal tool in the light of the sheer dexterity and skilfulness typifying human action. When critics do address actual tools-in-practice,

they tend to demonstrate how immersing a formalism into a practice's routines results in a complete subordination of the tool - a return, as it were, to things as they were before. Grasping the intricacies of the domain and the crudeness of the map, however, does *not* render the latter useless. On the contrary, it is the map's crudeness (or, less disapprovingly, "selectivity") which allows it to *work* like a map (Wood 1992, 1): to allow overseeing large areas at a glance, or navigating one's way across long distances. In and through the intersecting processes of achieving a tool's prerequisites, a world is created in which hybrids of techniques and personnel coordinate tasks and accumulate data in wholly new ways: across larger distances; incorporating more complexity. Through the *intertwinement* of informality and formality, thus, through the interlocking of representation and represented, worlds shift into a new plane.[29]

In addition, comparing a medical staff member's actions to the actions of a tool overlooks the fact that a tool does not "fill in" the slot in the practice previously "filled" by the medical staff member - the very process of constructing and implementing a tool *alters* this slot. As the end result of the negotiations in the implementation of ACORN or de Dombal's tool, a practice had emerged which embedded decision strategies unfathomable to physicians. Comparing the performance of a physician to that of the tool, here, ignores that physicians and accumulating tools more often than not measure different things, combine different data, or decide on different questions: in getting a tool to work, we have seen, these issues are invariably transformed. Arguing that computer-based decision aids are not able to mimic the expert physician side-steps the fact that most implemented decision support techniques are not mimicking physicians' actions in the first place: part and parcel of a changed practice, they do things nobody else does or can do. Arguing, similarly, that diagnostic tools can make diagnosis more accurate overlooks the fact that more often than not, what "diagnosis" *means* is changed. Self-perceptively, co-workers of de Dombal remarked about the tool's greater success in diagnosing acute appendicitis:

> It must be borne in mind, however, that the clinician is taken as having diagnosed appendicitis whenever he chooses to perform an appendectomy, whereas the computer is trying to predict whether the appendix is inflamed or not. (Bjerregaard et al. 1976)[30]

Stretching this last argument a little might make the point more clear. It is never argued whether the CT scan does its job properly, or whether humans should be left to the job. This technology has created a *new* activity: accumulating a myriad of "simple" X rays into one composite cross-cut image. It offers a new piece of data - and the way it arrives to this conclusion, the many "decisions" and negotiations incorporated in the tool[31], remain mostly opaque to the medical personnel who act upon its output. The decision support tools discussed, I argue, are not so dissimilar from the CT scan as it might seem: they perform tasks that did not exist before, in practices that have changed in major ways. Since the practices "before" and "after" the tool's implementation are so different, since there are no "similar tasks" which can be compared, evaluating a tool's performance through comparing tool and staff member misses the point.

Investigating the intermeshing of formal tools and human practices, then, and studying the transformations in logics and decision criteria, holds an important challenge. Of course, the idea that decision processes or action sequences can be "faithfully represented" is as chimerical as the dream of a map without a theme: every representation is at the same time a transformation. Yet, this conclusion is only a beginning: in the end, what matters is how this representation *in* a practice alters the work of nurses, physicians, the lives of patients, and so forth. What matters is how the production and usage of the map *transforms* the terrain. Where we can go now and could not before - and what is hidden from view. The complex interactions between tool and medical practice have hardly been investigated. Just how responsibilities are diffused over collectives of doctors, nurses and protocols, for example, or just how tools like ACORN alter the decision making strategies embedded in a practice, has received little attention. Yet the central legal and ethical notions of "responsibility" and "decision" subtly shift meaning when control, as I argued in Chapter 5, is distributed among tool and personnel - when asking for the one (or what) who is finally "in control" has become an unanswerable question (see further).[32] Similarly, decision criteria subtly change, affecting who is deemed "acute," who is deemed "incurable," and so forth. When one is too preoccupied with artificially comparing performances on "similar" tasks, with artificially isolating a representation from the activities in which it is composed and used,[33] with opposing categories that rather evolve together and transform each other, these crucial issues are all too easily overlooked.

There is no longer much sense, then, in asking what properties explain success or failure. This question reifies the categories whose development and intertwinement needs study. Rather, focusing on the local, heterogeneous work performed, the ways tools are located, the interlocking of formal procedures with the ongoing flow of ad hoc activities and events (with its distribution of responsibilities and shifting of decision criteria) yields a new array of questions and issues - issues crucial for an understanding of the current state and practical (im)possibilities of decision support techniques. The unit of analysis has shifted: instead of focusing on either tool or medical practice (or opposing them), it is their very *interrelation* (their historical co-evolution and their interlocking in current practices) which is put central.

Refiguring the Critical Stance: The Power of Tools-in-Practice[34]

> The point is to get at how worlds are made and unmade, in order to participate in the processes, in order to foster some form of life and not others. ... The point is not just to read the webs of knowledge production; the point is to reconfigure what counts as knowledge in the interest of reconstituting the generative forces of embodiment. ... The point is, in short, to make a difference.
> (Haraway 1994)

The final issue concerning the questions critics and advocates pose is that their modes of analysis significantly reduce their critical potential. This may look like a contradiction: the critics have their devastating appraisal of

decision support techniques, and the advocates have their criticism of the current, unscientific state of medical practices. I, on the other hand, have been critical of several of the critics' and advocates' assumptions - but in doing so I may seem to have drowned any critical potential vis-à-vis tools or medical practices in the networks I have sketched. I can no longer draw upon the tool's fundamental characteristics to criticize medical practice - or vice versa. By arguing that we should follow the development of the intertwined trajectories of tool and practice rather than opposing them, I seem to loose the possibility of critique by denying myself a firm ground to criticize *from*.

In fact, however, the critical potential that is gained through refiguring the debate in this way is both much more comprehensive and much more nuanced than the often rather black and white views of both critics and advocates. The shift in analysis proposed in the previous subsections leads to concurrent shifts in the *terms* in which practices of tools and medical personnel can be judged. Moreover, leaving behind both medical practice and tool as the final ground from which to judge also implies leaving behind any pre-given notion as to the potential benefit of a tool, or its proper (non-) usage. The tool can no longer be perceived as the carrier of Rationality, but neither can it be seen as essentially powerless. Tools nor practices can any longer be criticized from the comfortable position of the outside observer. Refiguring the critical stance means abandoning *all* unmarked positions: it implies immersing oneself in the networks described and searching for what is or can be achieved by new interlockings of tools and practices. In short, refiguring the critical stance implies a new focus on design: a focus on *design as critique*. To explicate these issues, I will elaborate two interrelated routes to new questions which can and must now be posed.

- Different Rationalities

First, advocates and critics turn the debate on decision support techniques into a monolithic argument. Either the tool's universal, scientific Rationality is applauded, or the stifling effect the tool will have on the skillful and informal work routines of medical personnel is deplored. Depending on the sides authors are on, the Machine stands either for the final coming of Rational Medicine, or for the threat of a soulless, empty formality.[35]

According to many advocates, for example, decision support techniques are the optimal tools to fight medical practice variations: they may streamline and rationalize medical practice according to scientific principles. The urge to uniformize medical work, however, epitomizes the *illusion of the singular answer*: the illusion that, in principle, for each "given" medical problem there is one optimal intervention which can be scientifically established. It is in this view that the existence of "practice variations" is an embarrassment. And indeed, within a cognitivist perspective, medical practice variations become a problem of the limited information-storing and processing capacities of physicians. Drawn into a discourse prioritizing standardization, the variations become a problem of a lack of uniform terminology and criteria. In both cases, however, a broad array of elements medical personnel constantly deal with in managing patients' trajectories

are being overlooked - or are being looked at as "bias" of a potentially Rational process. Organizational considerations, desires and hopes of patients, and the availability of certain technologies all shape patients' trajectories in important and inevitable ways. These considerations vary from patient to patient, from region to region: the specifics of the patient populations a physician is dealing with vary, as do her relations with her colleagues, the ways of referral, and so forth. When all these elements are taken into account, and when the need to continually (re-)articulate them is taken seriously, the illusion of the singular answer vanishes, and the problem of "medical practice variations" changes shape as well. Instead of a sheer threat to the latent scientific character of medical practice it can be reconceptualized as the reflection of the situated work of physicians.[36] Striving to eliminate variation in the name of Rational Medicine is an utterly useless enterprise as long as the differences between the configurations of medical practices - in a broad, heterogeneous sense - are as variable as they currently are.

More fundamentally, the idea of a singular answer is immediately connected to the assumption of the existence of the *one*, optimal Rationality. In medical work, however, many rationalities co-exist. There is not one way to typify the actions of a "good doctor" (or nurse, or paramedic) (de Bont 1994; Berg and Mol forthcoming). Sometimes being "rational" is stated to imply that the patients' desires should be primary: the doctor, after all, is the patient's "advocate." Sometimes being "rational" is taken to mean that the doctor should take into consideration that many medical resources are scarce in contemporary society - health workers, after all, are themselves part of that society. Sometimes, as a last example, it is argued that data obtained through face-to-face interaction with the patient should outweigh data gathered through "additional" laboratory tests - since the former are the "core" of the health worker's professional expertise.

We can argue in favor of some of these statements, or against them - but they can never be unified. They represent different logics, different rationalities, some of which may be preferred to others - in certain contexts, in certain situations. There is no foundation *outside* of these rationalities which could provide an ultimate ground for comparison - let alone that they could be "united" into one "overall" Rationality (ibid.). Pretending that there is one optimal answer to every medical problem, to subordinate one logic to another, is to deny that in different situations, different answers may be more "rational" than others. Pleading for the One Scientific Way is to give up the leverage that is gained when we *stop* depicting biomedicine as a singular unity.

The illusion of the singular answer, then, is the illusion that it would be meaningful or possible to search for the ultimate yardstick with which to measure medical work. In addition, if we turn to these yardsticks, not even the decision support tools themselves constitute one Rationality. As "universality" is not about being without location, the "coming of the Machine" should not be equated with homogeneity, or with the gradual advent of the one true voice. Both advocates and critics overlook the fundamental *differences* between the rationalities embedded in the tools. Formality should not be opposed to content, as the tools' critics often seem to do: formal tools can carry clinical or statistical logics in more or less impure

mixtures, they can and will address similar problems in wholly different ways[37], they contain different assumptions about the roles of patients, and so forth. There is not one map per terrain - there can be many maps, yielding different navigations, different foci of attention, different themes. Disciplining practices to a formalism results in as many different practices as there are tools. The opposition, then, is no longer between the domination of the sole reductionist voice and the rich multivocality of the reduced. Rather, the *different* rationalities, embodied in the different tools and practices, can now be contrasted and weighed. One rationality might be preferable to another; the politics of one formalization might be favored more than the next.

What could this look like? For one, tools' scripts could be judged as to the roles of medical personnel they embed. The strengthening of hierarchies, and/or the reduction of influence of personnel on the shape of their work, I argued, is an inevitable consequence of the disciplining of a practice to a decision support tool. Yet, within the many rationales that can be found in the realm of the formal, some may imply hierarchies that are less desirable than others. Many attempts to formalize nursing work, for example, often can be seen to end up with a strong management-imperative. Such tools contain categories that are of primary interest to management (what type of interventions are being made, who is doing what in how much time?), implement a centralized mode of decision making, and/or build in direct mechanisms of control. Rather than supporting the exigencies of nursing work, then, these tools strengthen the hierarchical control *over* their work (Wagner, forthcoming; Timmermans et al., forthcoming). The ability to scrutinize the rationalities of (proposed) tools with regard to such imperatives seems to be an important gain.

As another example, the pros and cons of statistical versus clinical, and decision-focused versus coordinating logics can be debated. The different logics' notions of rational medical practice and its obstacles (embedded in the tools, the texts that accompany them, the implementation-strategies that are associated with them, the way they are (supposed to be) used, the way they are materially configured) prestructure the way in which specific practices, "problems" and "needs" are approached, conceived, and structured (*cf.* Agre 1994). Going beyond the debate *whether* tools can be of use, the (much more fruitful) question arises *which* logics will yield *what* consequences. General claims, again, are difficult: judgment depends on the specific uses to which the different tools are put, and the specific circumstances of the medical practices involved. In addition, I argued that during the process of constructing and implementing a tool, its logic becomes mingled with ad hoc considerations, logics pre-existing in the practices, and so forth. Nevertheless, for example, the clinical logic that typified the expert system and Weed's and Feinstein's protocol is *itself* already (and wittingly) a mixture of pathophysiological, anatomical, experiential and indeed statistical knowledges; it itself already acknowledges some of the messiness (and, to some degree, the cross-cutting relevances) of ongoing medical work. It is far less obsessed with "purity," with the one optimal answer, than the statistical logic is; it far less attempts to search Rationality *elsewhere* - as in the clean smooth numbers of a formula.[38] Notwithstanding the caveat against general statements, then, a clinical

logic may prove a more fruitful and desirable point of departure than Bayes' Theorem or the calculations of decision analysis.

The decision-focused logic, in conceptualizing medical practice as a scientific and mental activity, narrows its focus down to the cognitive functioning of individual staff members and overlooks the interactive and thoroughly heterogeneous nature of current medical work. Doing so, for one, runs the danger of solving the wrong problems. One of the reasons that ACORN was not a great success, according to Wyatt, was that the Coronary Care Unit was indeed often full. This organizational fact caused delays in the emergency ward, and resulted in patients being sent home who would otherwise have been admitted. Focused on ameliorating the limited decision making capacities of medical personnel, ACORN had set out to reduce delays and prevent erroneous decisions by improving the decision making process. As Wyatt stated in hindsight, "we identified a non-problem" (interview; *cf.* Wyatt 1991a, 180). Organizational contingencies of their hospital caused the phenomena ACORN had set out to erase - and to which it formed no solution at all.[39]

In addition, the decision-focused logic is problematic since it channels attention away of the social and material *organization* of medical work. The medical practice variations I mentioned above are a case in point: in attributing these to "not-so-optimal decision making capacities," the ways intellectual capacities are embedded and constituted in concrete practices threaten to disappear from view. Conceptualizing the individual health care worker as a central source of trouble in current medical practice distributes blame in a highly specific manner - foregoing scrutiny of insurance-arrangements, the role of pharmaceutical companies in structuring the therapeutic options available, and so forth. It was a President of the Medical Decision Making Society itself who stated that the contributions of "inept doctors" to medicine's failings are much smaller than the contributions of "inadequate systems" (Berwick 1988). As a modest move in this much larger debate, then, my criticism on the decision-focused logic is an argument in favor of a logic more attuned to the social and cooperative character of current medical work.

A focus on different rationalities is also important for yet another reason. As argued above, comparing what the tool does to what medical personnel do, or attempting to determine who or what should have final decision making control, is meaningless when "decisions" are made by hybrids of tools and medical personnel and "responsibilities" are diffused over heterogeneous elements. In such situations, all grip on the shifting decision criteria and the specific redelegations of decision power is lost if some purified notion of "human decision making" or of the blueprint of the tool is hold on to. In debating and judging *rationalities*, however, we are no longer judging either physicians or tools. The theoretical unit of attention is no longer what a medical staff member decides, or what a decision support technique does. Different rationalities can be discerned and discussed irrespective of in what or who these are embedded. It is not that rationalities, or logics, are not performed and embodied in concrete practices: I am not trying to reintroduce a category which can stand apart from the networks in which it is constituted (see Chapter 1 and 2). I argue, however, that these rationalities constitute discernible patterns, or "modes of orden-

ing," which cut across the categories of "tool" and "human practice," and which can be debated *as such* (Law 1993, Mol forthcoming). When it is exactly the interaction between formal tool and human work which yields new realities, we require a vocabulary which does not require a "sorting out" of categories *before* it obtains the leverage to judge. Rather than recreating dichotomies, then, a focus on "rationalities" or "logics" is a much more fruitful way to start. Rather than setting up a procedural argument which has lost all relevant anchoring points, rather than forfeiting critical leverage, a focus on the *content* of the shifting decision strategies and working patterns (including the specific ways responsibilities are distributed) offers a more promising path to take.

- Multiple Positions

In addition to focusing on different rationalities, the analysis developed in this book moves beyond a "for" or "against" position by centering on the manifold transformations that occur in the process of constructing and implementing a tool. The network of texts, artefacts, personnel, patients, institutions, I argued, is at the same time a coalescing of a multitude of networks - physicians dealing with patients' problems, patients managing their lives, nurses accomplishing their work-schedules, tool builders construing their tools, hospitals increasing their efficiency. There are *many* "central" actors, each having its own goals, needs, wants, idiosyncrasies. To travel through the network, then, to acknowledge the many positions, experiences and perspectives, is vital to any judgment of a specific tool-in-practice. The contingent requirements of the one network might reverberate in unwanted and unforeseen ways in the other: to judge a working tool, one needs to get at the clashes that occur where networks intersect, or the resistances that come up due to crosscutting needs and goals.

This issue is another reason why "success" or "failure" are such problematic terms. When we follow the transformations that take place in and through the emergence of new tools and new practices, it becomes clear how "success" or "failure" are always success or failure *for* somebody, *for* something. Speaking about "success" without such qualifications is itself already a characteristic of the view that the one true Rational voice is at stake - the privileged view that is a common good *naturally*. Given the analyses presented in this book, however, it is clear that "success" in terms of "a working tool" always also implies that some voices will have to be silenced, and that some knowledges are neglected. "Success" in terms of a working tool is success *for the tool builders* - but maybe not for patients who may be faced with more, painful investigations, or nurses who may have to do more data-gathering chores, or hospitals who may be confronted with higher costs.[40]

Rationalities and positions form crosscutting dimensions: different modes of ordening juxtapose heterogeneous elements in particular ways, and have different consequences for different positions in the network. A strong management imperative, for example, will be valued differently by the hospital administration than by medical staff members; a statistical logic may embody patients' preferences in a principled manner, but work

against work routines of medical personnel. Moving beyond either being for or against decision tools, then, I have complexified the issue considerably - surely raising more questions than answering them. How to weigh different consequences, how to deal with the fact that most of these will be fundamentally unpredictable? But these are questions that need to be confronted - not run away from.

First and foremost, these considerations lead toward the issue of *design*.[41] When Tool and Practice are seen as essentially structured in the same way, when the tool is depicted as simply filling an already existing niche in a practice, the process of design is downplayed to the technical problem of construing the physical artefact. When the work of *achieving* a niche is seen, however, when the recursive relation between tool and practice is taken seriously, "design" is put in a different light. When the illusion of diffusion without transformation is left behind, it becomes clear that traveling through the network is a prerequisite for construing a working tool. To get a grasp on the possible instances where different networks and network building activities either converge and mutually strengthen each other, or rather clash and obstruct the emergence of a working tool, it is utterly necessary to acknowledge the presence of different logics and multiple positions. This need springs both from the drive to construct a tool that might actually be used and from the (intimately related) requirement to ask just *for who* or *what* it would then be "a success." When the breadth and depth of the occurring transformations are taken into account, it becomes clear how the politics involved are just as multilayered.

In addition, design comes to the fore when the illusion of tools which are *essentially* defective is left behind: looking for (potentially fruitful) consequences of interlockings of tool and practice *is* nothing more nor less than design. Reappropriating decision techniques as political agents, then, is reconfiguring *design as critique*: reconceptualizing design as a means to intervene in the constellations of tools and practices. In Suchman's terms, this would imply a shift from "viewing design as the creation of discrete devices ... to a view of systems development as entry into the networks of working relations - including both contests and alliances - that make technical systems possible" (1994; *cf.* Markussen 1994). Finding a place for one's tool within "specific ecologies of devices and working practices" implies crossing and reconstructing the boundaries that may exist between designers and users. Rather than sustaining the illusion of "design from nowhere," it implies acknowledging the locatedness of the tool and the rationalities it embeds (ibid.). As is the case for the management of patients' trajectories, "success" is tinkered to; not preplanned from elsewhere. This notion of design, moreover, implies abandoning the illusion of the singular answer, and considering the knowledges one's tool would strengthen or, rather, erase. It implies leaving behind the illusion that tools could be construed which would *not* hurt somewhere, which would not discipline, or not curtail options or ranges of activity at some place in the network. It means, finally, acknowledging and dealing with the intangibilities that responsibilities for "decisions," the "agency" directing the patient's management, will be distributed over hybrids of humans, tools, regulations, and so forth even *further*.

What could this mean? Looking at specific instances, what could "taking the manifold of positions and rationalities seriously" come down to? Focusing on the position of medical personnel, for example, we could look again at the tendency of formal tools to strengthen hierarchies. An important point to consider in the construction or judgment of decision support techniques, then, is that these effects should be counterbalanced: that disciplining does not occur without gaining something as well. In other words, the protocol's coordinating function, or an expert system's accumulating potential should lead to personnel acquiring new *competences* (new skills, new responsibilities, new drugs to deal with, and so forth (*cf.* Robinson 1991b)).[42] The coordinating role of the protocol, for instance, can be put to use to allow regional hospitals to handle complex treatment regimes that used to be reserved for academic centers. Arguing for and construing decision support techniques that use their potential for "providing a wider perspective and for creating new terrains across professional and departmental boundaries," is a crucial strategy in the attempt to intervene in the development of this class of tools (Wagner 1994).

Recreating "specific ecologies of devices and working practices" also involves attention to the additional responsibilities that are introduced with the coming of a decision support tool (see Chapter 5). When a decision technique has been implemented in current medical practices, personnel have to achieve a tool's demands at the same time as they have to manage a patient's trajectory. Looking at different types of tools, the workings-in-practice of one tool can facilitate or hinder the interlocking of these two activities more than that of another. Having shifted the terms in which tools and practices are judged can then also mean preferring one type of tool to another. So, I argued that the unintelligibility of the tool's decision making processes is much more acute in the case of the decision-focused, accumulating tools than in the case of the coordinating protocol. What a computer-based decision tool does with the input data is hidden in a literally opaque box: the firing of the many rules, or the unfolding of the mathematical formulas, are intended to create a meaningful advice - not to be understandable in their precise interactions. Protocols' decision points, on the other hand, can always be traced in the tool's tables or text. The more opaque the tool, the more precarious the work to achieve the formalism's prerequisites: what the tool intends with a certain question or advice, or how an unexpected contingency should be dealt with, is then even harder to anticipate.[43]

A related difference between current accumulating and coordinating tools-in-practice is that the former require explicit moments of input of data and acknowledgement of output. The accumulating tool forces a "stop" in the ongoing flow of medical work, in which it creates an advice and informs the user. The protocol, obviously, has to be followed as well, and equally gives rise to sessions where data are entered and decision criteria are looked up. But because the internal logic of the protocol is generally more transparent and less complex, routine users often work "according to the protocol" without explicit reference to it. They do not have to explicitly "go to" the tool to use it. Data collected are often not written into the protocol until much later, and subsequent steps of the protocol are regularly anticipated beforehand.[44]

The intertwining of decision-focused tools with the ongoing work of medical personnel, then, creates more frictions than protocols do: the demands of accumulating tools generally clash harder with the demands of managing patients' trajectories in current medical practices than those of coordinating tools.[45] With regard to this issue, thus, and with regard to medical personnel, coordinating tools may generally be preferable to accumulating tools.[46]

As a last example of how a focus on the different rationalities and positions can increase the potential to intervene, we could focus on the consequences different knowledges might have for patients. Decision support techniques can and are used as means to "transport" knowledges deemed "better" than others to a range of different practices. Of course, formal tools are not inert media at all, and every transportation implies a transformation of the knowledges and decision criteria involved. However, there is no reason why this usage cannot be reappropriated. The awareness of the many options the formal leaves open, and of the multiple rationalities that abound in medicine can lead one to create decision support techniques that embody specific medical criteria and knowledges that are held important. As Feinstein is attempting to save "soft data" from oblivion, and insurance companies might be attempting to construe protocols to minimize costs, decision support techniques could also be construed which attempt to mitigate, for example, oncology's often "aggressive" approach in palliative treatments (i.e., when there is no more hope for cure). (At the very least, the fact that so many rationalities coexist is a strong argument for participation of a broad range of potentially affected groups (such as patient organizations) in the construction of new decision support techniques - including, importantly, the research protocol.)

The debate, then, is no longer either "for" or "against" the tools. The specific locations of (individual) tools can now be thematized - and the illusion held by some critics that "all control has to be delegated back to the users" can be left behind. Just as tools without a history are a fantasy, so it is an illusion to call for tools that do *not* require a disciplining of practices; that do *not* interfere in the "informal" work processes critics hold so dear.[47] Seeing that different tools can carry different rationalities, seeing how different tools reshape practices in different ways, opens the way to a much more fruitful strategy. By breaking away from either having to embrace formal tools or denounce them, from the debates on the "essential" properties of the tools or the "final" reasons for failure, new space has been created, new leverage, new potential to intervene, to compare, to prefer - and maybe even to choose. Refiguring the critical stance, then, implies acknowledging both the dangers and opportunities that decision support tools might entail. It implies broadening the critic's scope of critique by going beyond the realm of proclaiming one's dissent to actually reappropriating the tools themselves: moving beyond *deconstructing* the tools' claims to Rationality and Universality towards actively trying to *transform* their development and use (*cf.* Haraway 1991, 149-82; 1994). No longer denouncing tool or practice, it means searching for ways in which decision support tools may become familiar yet never totally transparent, powerful yet fragile instruments of change.

Notes

1. This number excludes four "educational systems" and some statistical tools (including de Dombal's tool). It includes, on the other hand, several tools which are not primarily designed for usage in medical practice (such as a system which reviews physicians' prescribing patterns for Medicaid patients' drug utilization). The list (version V1.5, July 21 1993) is compiled by Enrico Coiera, and appears on the AI-Medicine mailing list (e-mail; contact ai-medicine-REQUEST@med.stanford.edu). The systems in operation in more than one location are: two laboratory systems for analysis of test results, PUFF (a system which interprets pulmonary function data in lung function laboratories; *cf.* Aikins et al. 1983), APACHE III (which predicts the risk of dying in the hospital for admitted patients; *cf.* Knaus et al. 1991), and a "managed second surgical opinion system."

2. On MYCIN, see Buchanan and Shortliffe (1984) and Lipscombe (1991). On the "demise" of INTERNIST, and its transformation into QMR, see Miller et al. (1986) and Miller and Masarie Jr. (1990).

3. *Cf.* Balla et al. (1989). See also the poor results of a "clinical trial of clinical decision analysis" performed by Clancy et al. (1988). In this trial, it was attempted to rationalize physicians' vaccination decisions (for themselves) through decision analysis.

4. On the limited practical success in the Netherlands see e.g. Knottnerus (1987) and Warndorff et al. (1988). In health policy, clinical decision analysis appears to be more successful - but see the critical appraisal of Ashmore et al. (1989).

5. On the explosion of research protocols in postwar medical practice, see e.g. Löwy (1993) and Patterson (1987). See Löwy (ibid.), Richards (1991), Epstein (1995) and Marks (forthcoming) for studies of the negotiation processes typifying the construction and evaluation of clinical trials.

6. By analyzing discharge data from Ontario hospitals over a six year period, starting four years before and ending two years after the consensus statement was published.

7. Similarly, in an evaluation of the effect of four NIH consensus development guidelines, the authors had to conclude that "taken as a whole, these four consensus conferences had no effect on physicians' hospital practice" (Kosecoff et al. 1987). And see Hill and Weisman (1991) who draw similar conclusions in their prospective study of Maryland physicians. They show that the consensus reports' influence is minimal: "the strongest predictor of congruent practice behavior 1 year after [the report] is congruent practice behavior just prior to the report's release."

8. Few advocates would claim that clinical decision analysis has to be applied always: most state that it should be reserved for "difficult cases" (*cf.* Schwartz et al. 1973).

9. According to Stephen Pauker, one of the founders of clinical decision analysis, the main problem was that there is no niche in the financial organization of medical care: "It is an activity which is not reimbursable by insurance companies or whoever pays for medical care. ... We are only reimbursed for seeing a patient. For contact time, or for procedures. I am a cardiologist. If I do a cardiac catheterization, and spend an hour and a half doing it, the insurance company would pay me $1500. If I do eight hours on the analysis of a patient, if they would pay me at all, they'd pay me for medical consultation, which could pay out at $100" (interview October 29, 1993). *Cf.* Littenberg and Sox Jr. (1988); Detsky et al. (1987); Dolan (1990).

10. In principle, advocates argue, wherever humans apply general knowledge to a particular problem, computer-based decision support systems should be feasible. Nevertheless, the practical state of e.g. expert system research might force a restriction to relatively well-defined domains - as medicine (*cf.* Duda and Shortliffe 1983).

11. See for similar analyses of the limited success of computer-based decision tools also de Dombal (1987b), Duda and Shortliffe (1983) and the early remarks by Barnett (1968).

12. In Lomas's terms, it is important to "facilitate the policy's incorporation into local practice environments by exploiting enough of the non-scientific influences on decision making for actual behavioral change towards the policy to be enabled and reinforced. There is a touch of Machiavelli in this process, since the underlying philosophy is that one cares not *why* the policy change occurs, only that it does occur. In an ideal world, ... scientific information [should be enough]. In reality, however, the behaviour change will [require tactics] learnt long ago by the pharmaceutical companies" (1990, 183). So, Lomas' team selected "educationally influential local physicians" through sociograms, who were trained and "sent back into the community with a request to maintain their

usual network of contacts and encourage implementation of the statement." They were also asked to distribute materials to their colleagues which were based on the "same principles used by pharmaceutical companies": "visually attractive and compelling, focused on specific concerns ..., and addressing 'bite-sized' chunks of information" (ibid. 186). Further, Lomas and his colleagues attempted to set up audit-rounds in local hospitals, using the consensus statement as gold standard. *Cf.* Kanouse and Jacoby (1988: 244-47); Lomas et al. (1991); Zwaard et al. (1989).

13. See Böckenholt and Weber (1992) for such an analysis of the general category of "formal methods in medical decision making." This view is shared by "diffusion" sociologists, who argue that e.g. computer-based decision tools have failed to catch on due to "too little information to doctors," "negative attitudes of medical personnel," and so forth. See e.g. Anderson and Jay (1987, 3-7); Lundsgaarde (1987); Weaver (1991). See also e.g. Evans (1990, 127), who argues that to understand the success or failure of expert systems one should look at the medical profession's "successful suppression of the nurse practitioner": it is all a matter of "control over the social and economic context of practice." On this mode of explaining "failure" see Kling and Iacono (1984); Kaplan (1987; 1995); Star (1989a; 1989b); Forsythe (1992); Latour (1987, 135-6).

14. Quote is from Forsythe (1992).

15. See e.g. Yu et al. (1979). On the evaluation of expert systems in medicine see e.g. Lundsgaarde (1987); Wyatt (1991a); Miller (1986).

16. Fuller criticizes the "practice-mysticism" of authors like Collins (1985; 1990), who argue that there is some "tacit dimension" in most domains of human action (yet another foundation) which will forever elude all effort to be formalized (1993a, 179-85).

17. In Fuller's terms, debating whether Man is fundamentally equivalent or fundamentally different from Formal Tools mistakes the practical problem of *achieving* equivalence for the metaphysical problem of the *existence* of transcendental equivalence (*cf.* Fuller 1991). Woolgar (1985; 1987), Turkle (1984) and P. Edwards (forthcoming) all have addressed the issue how new technologies transform our notions of what it means to be "human." See also the excellent collection of papers in Ashmore, Wooffitt and Harding (1994). *Contrasting* Human Practice and Formal Tool, comparing their "performance" on "similar tasks," *itself* already contributes to the active process of rendering them equivalent. Debating the superiority of human expert decision making over formal tools' decision making (or vice versa), for instance, implies a highly individualistic and mentalistic understanding of what "expert action" (in this case, medical work) is (*cf.* Fuller 1991; 1993a, 179-85; 1993b). The question of superiority or equivalence is to be settled as if a game of chess, one lone individual posited versus another, is paradigmatic of what medical work comes down to (see also D. Edwards 1994).

18. On the digitilization of medical practices, see e.g. Anon. (1964) and Reiser (1978); see Mesman (1993) for an account of the highly digitalized practice of current neonatology. This tendency to quantify information is obviously tied to the current high status of the statistical logic discussed in Chapter 2.

19. See e.g. MYCIN's successor, ONCOCIN, which was explicitly intended as a tool to be implemented in a practice already "predisposed" to formal tools. ONCOCIN is supposed to help oncologists follow research protocols. *Cf.* Berg (1994); Lipscombe (1991).

20. See Marks (1988), on research protocols, for an excellent illustration of this point.

21. Research on the treatment of cancer was one of the main spearheads of the coordinated medical research effort that boomed after WW II. See e.g. Löwy (1995); Marks (1988); Zubrod (1984).

22. This rendering is not completely fair to at least one of those I have labeled "critics." Collins, for example, can be said to look at the *process* of getting a tool to work. His position is a social constructivist's inversion of Dreyfus. He argues against Dreyfus that whether a domain is formalizable or not is not a given feature of the natural world-out-there. Even domains like mathematics and physics, Collins argues, are only formalizable because in the social practices constituting these domains we have *opted* to behave in a formalizable way. Collins calls this "behaviour specific action," which he sees as a specific, wilfully created subcategory of human, "regular action." Regular action is itself fundamentally non-formalizable. My problem with Collins' argument is that, in turning Dreyfus upside down, he ends up with a position which is similarly troublesome. Now, the formalizability of a domain is no longer a given property of the world out there, but

has become a property entirely situated in the domain of social phenomena. Translated into the vocabulary deployed here, Collins points at the need to discipline a practice to a formalism - but he focuses solely at the need of *people* to conform themselves to the tools' demands. Overlooking the sheer heterogeneity of these practices, the assembly of bodies, artefacts, machines, people, and so forth, which *all* need to be disciplined, Collins in fact does not bring Dreyfus' analysis all that much further. In addition, by making the a priori distinction between "regular" and "behaviour specific" action, Collins' argument is just as foundationalist as Dreyfus' position.

23. For the term "obligatory passage point" see Latour (1987).

24. See Timmermans and Berg (forthcoming) on how actor-network theory may sometimes seem to be unduly focused on the mythical and omnipotent Central Actor.

25. My characterization of formal tools as requiring constraints on input, output, and the relation between data entered and output-statements (Chapter 3), as the distinction between accumulating and coordinating tools (Chapter 5) both already *contain* this interrelation of "tool" and "context": these definitions are useful as far as the events described in this book go (i.e., they are not simply an alternative set of essential properties), and they describe the tool *in terms of* its relation with the network in which it is (to be) embedded.

26. The term "infrastructure" comes from Bowker (1994b). On how local instabilities can create global stability see Singleton (forthcoming).

27. *Location* was the topic of the inspiring 1994 Oksnoen Symposium on *Locating Design, Development and Use*, May 13-18, Norway.

28. The requirement of human intervention to overcome the ineptness of machines in the social order is the theme of Collins (see note 22).

29. Robinson's notion of *double-level languages* gets at just this interlocking of the formal and informal (1991b).

30. The two need not fully overlap: clinicians might think there is a high chance of appendicitis but - for a variety of reasons - delay surgery anyway. On the other hand, they might estimate the chance of appendicitis as rather low, but operate anyway - when, for instance, they are hesitating between "appendicitis" as a probable diagnoses and another affliction that would also require a similar, urgent surgical intervention.

31. The processing of its raw data involves many "decisions" on the way: the tool draws upon a specific mathematical technique in reconstructing the composite image from the individual shots taken and not on others; it makes so many readings and not more or less; it represents its calculations pictorially drawing on specific algorithms (Blume 1992).

32. Of course, this issue is not merely limited to situations where decision support tools are at stake - see e.g. de Vries (1993). The matter just comes up more saliently here.

33. This latter expression is after Lynch (1993, 192).

34. After Wood's insightful *The Power of Maps* (1992) - an important work for any sociology of the formal.

35. See Latour (1993, 65-7) for a critique of the notion that Technique is devoid of Being.

36. See Cummins (1980) for a perceptive analysis as to the reasons why physicians might still want to order a given test even if, according to an "indication list," no indication is present. Among these, many "contextual" reasons are mentioned.

37. This is and can only be a conjectural claim: there is no way this could be empirically "validated." Even if two systems could be found that dealt with *exactly* the same clinical problem (and given their different views of what medical practice's problems are, this is rare), it always remains possible to argue that differences in advice found are due to differences in assumptions embedded in the rules and formulas which are *not* related to the different rationalities investigated.

38. Advocates of a clinical logic do tend to overstress the unity in "clinical thought" that would characterize clinicians' work: the plethora of different logics - statistical, clinical, pathological, and so forth - is much larger than e.g. Feinstein would argue. The confrontation with a statistical logic, thus, is a much less alienating experience than he and others would claim (Mol and Berg 1994). See Lawrence (1985) for a historical account of how British cardiologists incorporated experimental physiology into their "clinical logics."

39. The evolution of de Dombal's tool leads to a similar conclusion. This tool started out in the early seventies as a diagnostic machine; as a tool which would offer the highest probable diagnosis given a case of acute abdominal pain. De Dombal's team produced

many papers showing how bad physicians performed on this task, and how much care improved when their tool was used. In the course of the years, however, de Dombal and his team started to realize that "at least half of the improvement [in clinical performance] was because of the adoption of preagreed, predefined medical terminology, coupled with discipline in data collection" (de Dombal 1987a). Some years later, a study demonstrated that the results witnessed were *entirely* due to the structured data collection ensured by the form - whether the physicians ever saw the computer's advice or not did not make any difference (de Dombal et al. 1991). Again, it seems that the cognitivist discourse, of which de Dombal's statistical tool was a paradigmatic exemplar, provided a wrong tool for the job: the paper form would have yielded the same results *by itself*. Again, a solution was provided for a problem which did not exist at the level of the work of medical personnel. The problem was not so much a suboptimal mental functioning of physicians, but rather a matter of having too little time for a patient, and not being comprehensive enough in one's physical examination and history taking.

40. *Cf.* Star (1991) and Haraway (1991, 183-202).

41. This is not the place to dwell in detail on how "design" would have to be refigured. For inspiring and elaborate accounts of alternative design methodologies, see e.g. Zuboff (1988); Greenbaum and Kyng (1991); Robinson (1991a; 1991b) and Agre (1994).

42. See also Zuboff (1988). Focusing on this aim would be a wholly different (and potentially more fruitful) approach to enhancing a profession's status as well.

43. The fact that a protocol-in-practice appears less opaque than a Bayesian tool has more to do with its material form (i.e., it is of limited complexity, and written down on some pieces of paper) than with its clinical rather than statistical nature - see above.

44. When observing, the protocols' frequent invisibility was a continuing source of frustration to me. The days from Mr. Wood's trajectory I described were highly structured by the running together of several protocols - yet to me this was not an overt, explicit feature of the activities that took place. It should be noted that this feature of the expert system may be circumvented by letting the system draw upon data that are already stored in electronic form (so-called "critiquing" systems may be set up in this fashion).

45. A corollary of the fact that protocols can "sink" into the work routines is that, once in use, they often become even more localized than accumulating tools. Protocols can become so routinized that at a given point nobody really knows whether certain actions were or were not part of the protocol (see Star 1994 on the "naturalization" of objects and artefacts in social practices). Asking physicians if a certain test was "according to the protocol" would yield answers like "I think so," or "it should be." For example, when I later checked the described days of Mr. Wood's trajectory with some of the protocols used I found that many small violations had occurred. Many blood tests had not been done that should have been done, prophylactic antibiotics (to prevent infections by the bacteria that patients carry themselves) were not started ten days prior to the high-dose chemotherapy but only seven days, and so forth. These "violations" were never discussed, never mentioned - nobody probably ever realized that they had occurred.

46. As argued in Chapter 5, coordinating tools may extend the reach of activities of medical personnel over time and place and/or increase the complexity of the tasks they may handle. Accumulating tools similarly may make complexity doable - yet they not so much enlarge the reach of personnel over time and place, but in depth. By preprocessing sets of data, they may increase personnel's grasp over a broader array of information. The different tools, then, can allow for different redelegations of competences. In line with the decision-focused discourses, however, accumulating tools are generally conceived and produced to "do the tasks that medical personnel do": to stand in the heart of the ongoing activity of managing patients trajectories. If accumulating tools would be decentered more from this position, their accumulating potential might be employed in a way which would less interfere with the configuration of current medical practices. If, in other words, they would attempt less to "replace" or "support" the doctors' or nurses' decision making, and try more to pre-articulate, for example, complex sets of laboratory data (think of tools - already widely used - as computers which analyze multiple blood tests in a single "batch," or the CT-scan), the criticisms I described would carry much less weight.

47. Sometimes, discussions in the CSCW (Computer Supported Cooperative Work) literature seem to aim towards this: towards a desire for technologies that are like blank sheets of paper - that are totally structured and shaped by those that use them. For a sympathetic critique, see e.g. Robinson (1994) and Berg (1993).

Het Rationaliseren van Medisch Werk: Beslissingsondersteunende Technieken en Medisch Handelen

Samenvatting

Beslissingsondersteunende technieken (zoals protocollen, expert systemen en besliskundige modellen) worden alom bediscussieerd als ideale technieken om het medisch handelen te rationaliseren: om medische praktijken wetenschappelijker, rationeler en homogener te maken. Voorstanders van deze technieken stellen dat hiermee medisch handelen eindelijk het stadium van de "kunst" kan verlaten en een "wetenschap" kan worden. Critici betogen daarentegen dat deze technieken niet kunnen werken zonder tot "deskilling", tot kwaliteitsverlies en tot starheid te leiden. Filosofen als Hubert Dreyfus, wetenschapsonderzoekers als Harry Collins en verschillende medici vinden elkaar in het argument dat formele technieken fundamenteel niet in staat zijn om mensen te "vervangen" in het uitvoeren van inhoudelijke taken.

In deze studie onderzoek ik wat beslissingsondersteunende technieken doen en hoe ze al dan niet functioneren in concrete medische praktijken. Het neemt de tegenover elkaar staande stellingname van voor- en tegenstanders als uitgangspunten voor onderzoek. Uitgaande van inzichten en methodieken uit het recente wetenschaps- en techniekonderzoek wordt onderzocht wat rationalisering nu eigenlijk betekent in concrete situaties. De centrale these is dat een "werkzame beslissingsondersteunende techniek" een *eindresultaat* is van een convergentie van diverse ontwikkelingen die op verschillende niveaus plaatsvinden: het ontstaan van nieuwe discoursen over (de problemen van het) medisch handelen in de medische literatuur, het veranderen van medische criteria, en het transformeren van de heterogene elementen die tezamen een specifieke medische praktijk vormen.

De eerste twee hoofdstukken richten zich op ontwikkelingen in de (hoofdzakelijk Amerikaanse) medische literatuur, vanaf 1945 tot heden. *Hoofdstuk 1* baseert zich voornamelijk op redactionele artikelen in de *Journal of the American Medical Association* en de *New England Journal of Medicine*, en laat zien dat wat onder "medisch handelen" en "wetenschappelijkheid" verstaan wordt niet eenduidig is. Stellen dat beslissingsondersteunende technieken het medisch handelen "wetenschappelijker" maken, zoals voorstanders doen, en dat deze technieken de "oplossing" vormen voor "de problemen" van medisch handelen, gaat voorbij aan het feit dat er *verschillende* definities van "wetenschappelijkheid" aan te wijzen zijn. Er is

geen sprake van een unilineaire, graduele "verwetenschappelijking" van een praktijk die voorheen "onwetenschappelijk" was. Er is ook geen sprake van één centrale set "problemen" die door de jaren heen worden besproken. Er zijn verschillende discoursen te onderscheiden waarbinnen "medisch handelen", "wetenschappelijkheid", "problemen" en "oplossingen" steeds anders worden ingevuld.

In het eerste naoorlogse discours dat beschreven wordt, wordt een duidelijk onderscheid gemaakt tussen "wetenschap" en "medisch handelen". "Wetenschappelijke geneeskunde" duidt hier op het bestaan van een groeiende basis van wetenschappelijke kennis, waaruit artsen kunnen putten bij het uitoefenen van hun vak. Dit vak zelf wordt niet omschreven in termen van "wetenschappelijkheid": integendeel, vaak wordt gewaarschuwd tegen een teveel aan wetenschap, dat de "kunst" van het "zorgen" voor patiënten in de weg zou kunnen staan. Waar problemen van het medisch handelen worden besproken gaat het steeds om zaken die te maken hebben met "socio-economische factoren" waar artsen geen invloed op hebben: het tekort aan (overheids)geld om voldoende artsen op te leiden, vooroordelen van patiënten, enzovoort.

Vanaf de jaren zestig ontstaan er discoursen waarbinnen "wetenschappelijkheid" wel een kenmerk is van het medisch handelen zelf. Zo wordt gesteld dat medisch handelen dezelfde *structuur* heeft als wetenschappelijk handelen: in het behandelen van hun patiënten voeren artsen eigenlijk continu *experimenten* uit. Met deze typering verandert meteen de omschrijving van de problemen van het medisch handelen: binnen dit discours wordt vooral gesproken over het feit dat artsen, in tegenstelling tot wetenschappers, niet over gestandaardiseerde procedures en terminologieën beschikken.

Binnen een ander, meer recent discours wordt de wetenschappelijkheid van het medisch handelen gelokaliseerd in het wetenschappelijke karakter van het *denkproces* van artsen: het stellen van een diagnose en het instellen van een therapie, bijvoorbeeld, is als het opstellen en toetsen van een wetenschappelijke hypothese. Medisch handelen wordt in dit discours geconceptualiseerd als een cognitief proces - en de feilen van medische praktijken worden hier dan ook gezien als tekortkomingen van de cognitieve capaciteiten van individuele artsen.

In *hoofdstuk 2* wordt dieper ingegaan op de verschillen tussen en de specifieke achtergronden van de discoursen die medisch handelen als een wetenschappelijke activiteit typeren. Beargumenteerd wordt dat de verschillende discoursen *tegelijkertijd met* de ontwikkeling van de verschillende beslissingsondersteunende technieken ontstaan zijn. Deze technieken zijn dus niet zozeer oplossingen voor al langer bestaande problemen, maar zijn mede bepalend geweest in de conceptualisering van deze problemen. De verschillende vormen van "wetenschappelijkheid" en "rationaliteit" die in de discoursen figureren zijn in hun specifieke vorm pas ontstaan met de komst van specifieke beslissingsondersteunende technieken.

Achtereenvolgens worden de verschillen en overeenkomsten tussen (de discoursen van) de statistische technieken (waaronder diagnostische technieken gebaseerd op het Theorema van Bayes, en de klinische besliskunde), protocollen ("algoritmen", "richtlijnen" of "standaarden") en

expert systemen belicht aan de hand van discussies in de wetenschappelijke literatuur. Er zijn twee dimensies in deze verschillen en overeenkomsten te onderscheiden. Ten eerste verhouden alle technieken zich op de een of andere manier tot het idee dat een volledig statistische, kwantitatieve benadering van medische beslissingen in principe de meest rationele en optimale aanpak is. De statistische technieken belichamen dit ideaal - maar zowel de eerste generatie expert systeem ontwerpers als voorvechters van protocollen als Feinstein bekritiseren juist de "rationaliteit" die in het statistische ideaal besloten zit. Volgens de laatsten is een "wetenschappelijke" medische praktijk juist een praktijk die geënt is op de *klinische* werkelijkheid van medisch handelen, waar men niet geïnteresseerd is in getallen maar in inhoudelijke relaties, en waar klinisch pragmatisme geen probleem is maar juist een deugd.

Een tweede dimensie verdeelt de technieken op een andere manier. Zowel expert systemen als statistische technieken zijn gericht op het moment van de *beslissing:* het moment waarop een diagnose gesteld dient te worden, of een therapie dient te worden gekozen. Auteurs als Feinstein, daarentegen, bekritiseren deze individuele benadering, en stellen dat ondersteuning noodzakelijk is voor het hele *proces* van dataverzameling, intercollegiale consultatie, therapie-evaluatie, enzovoort. Het protocol, zoals zij dat voorstaan, intervenieert niet alleen op het moment van een beslissing, maar neemt de gebruiker als het ware bij de hand gedurende een hele serie van handelingen.

In de volgende twee hoofdstukken verschuift de aandacht van de medische literatuur en de ontwikkeling van discoursen naar het werk van de ontwerpers van specifieke beslissingsondersteunende technieken. *Hoofdstuk 3* bekritiseert het idee dat beslissingsondersteunende technieken eenvoudigweg in bestaande praktijken kunnen worden ingevoerd, zoals voorstanders veelal claimen - maar het bekritiseert ook de stellingname van critici dat formele technieken gedoemd zijn te falen. Het centrale punt van dit hoofdstuk is dat het "te werk stellen" van beslissingsondersteunende technieken een grondige transformatie van medische praktijken vereist: de praktijken waarin dergelijke technieken worden geïmplementeerd dienen te worden *gedisciplineerd* zodat ze voldoen aan de eisen die het functioneren van een beslissingsondersteunende techniek stelt.

De precieze vorm van deze eisen resulteert ten dele uit het feit dat beslissingsondersteunende technieken *formele technieken* zijn: technieken die aan de hand van expliciete regels of formules input gegevens (medische data) omzetten in output (een "advies", of een "diagnose"). Voor dergelijke technieken geldt dat de input welomschreven en gepreciseerd moet zijn, dat de output tot een beperkte reeks opties gelimiteerd is, en dat de relatie tussen input gegevens en output stabiel dient te zijn. Het "implementeren" van een beslissingsondersteunende techniek, nu, behelst het "onderhandelen" met de heterogene elementen van een medische praktijk (verpleegkundigen, artsen, maar ook medische gegevens en organisationele procedures) opdat aan deze voorwaarden voldaan kan worden. Of een beslissingsondersteunende techniek kan functioneren of niet is dan ook niet een vraag die a priori kan worden beantwoord: of een specifieke praktijk "afdoende" gedisciplineerd kan worden is een empirische kwestie.

Het disciplineren van een medische praktijk om een bepaalde beslissingsondersteunende techniek te laten functioneren gaat vaak gepaard met een aantal specifieke ontwikkelingen. Hiërarchische verhoudingen tussen medewerkers worden veelal versterkt, en de eisen die een techniek stelt aan bijvoorbeeld specifieke procedures worden vaak "gematerialiseerd" in formulieren of verpakkingen zodat het volgen van de procedure geen kwestie van een "keuze" meer is. Tenslotte wordt veelal gekozen voor input gegevens die zo min mogelijk van menselijke interpretatie afhankelijk zijn - omdat deze veelal meer voorspelbaar zijn, preciezer, en dus beter passen bij de eisen van een formele techniek. Dit alles leidt ertoe dat ook *medische criteria* veranderen tijdens het invoeren van een beslissingsondersteunende techniek. De claim van voorstanders, dat beslissingsondersteunende technieken simpelweg optimale kennis transporteren, is onjuist: de specifieke aard van beslissingsondersteunende technieken transformeert alles dat wordt getransporteerd.

Hoofdstuk 4 bespreekt de andere kant van deze medaille: de wijze waarop een beslissingsondersteunende techniek verandert gedurende het proces van ontwikkeling en invoering. De onderhandelingen die worden gevoerd laten ook hun sporen na in de techniek zelf: daar waar een praktijk wordt gedisciplineerd, wordt de techniek *gelokaliseerd*. Beslissingsondersteunende technieken die daadwerkelijk functioneren hebben altijd veel van de ideaal-typische "universaliteit" en rationaliteit moeten inleveren. Drie dimensies van lokalisering worden bediscussieerd: lokalisering in plaats, in reikwijdte, en in rationale.

"Lokalisering in plaats" duidt op het fenomeen dat individuele technieken immer de specifieke context waarin ze zijn ontwikkeld met zich mee dragen. De meeste functionerende expert systemen, bijvoorbeeld, zijn alleen operationeel binnen de praktijk waar ze zijn ontworpen - ze zijn eenvoudigweg zozeer op deze lokale omstandigheden toegesneden dat ze hierbuiten niet kunnen functioneren. Het overwinnen van "lokalisering in plaats" impliceert het op elkaar afstemmen van medische terminologieën, criteria, gebruiken, en organisationele procedures tussen meerdere locaties: verdergaande disciplinering kan lokalisering terugdringen, en, andersom, tekortschietende disciplinering kan leiden tot verdere lokalisering.

Een tweede vorm van lokalisering is "lokalisering in reikwijdte": aan de taken die een beslissingsondersteunende techniek zou kunnen verrichten worden steeds meer beperkingen opgelegd. Dit kan betekenen dat een techniek zich alleen richt op één bepaalde aandoening, of dat de "adviezen" die een techniek kan doen zich beperken tot algemeenheden. "Lokalisering in rationale", tenslotte, is de laatste vorm. Hier gaat het om het niet realiseren van de ideaal-typische rationaliteit, precisie en accuratesse die in de beslissingsondersteunende techniek besloten ligt.

Aandacht voor lokalisering leidt tot twee conclusies. Zoals gezegd resulteert lokalisering in technieken die gesitueerd zijn in een specifieke plaats en tijd, en die op hun ideaal-typische rationaliteit hebben moeten inleveren teneinde te kunnen functioneren. De continue onderhandelingen die met het ontwerpen en invoeren van een beslissingsondersteunende techniek gepaard gaan resulteren in technieken die gekenmerkt worden door compromissen en door contingente, ad hoc keuzen. In plaats van Dragers van één Universele Rationaliteit blijken beslissingsondersteunende

technieken veeleer ad hoc mengsels van verschillende overwegingen en (on)mogelijkheden met zich mee te dragen. In plaats van het uitbannen van "variatie" dragen beslissingsondersteunende technieken, paradoxaal genoeg, bij aan het introduceren van weer *nieuwe* idiosyncratische beslissings- en handelingswijzen.

In *hoofdstuk 5* wordt onderzocht hoe beslissingsondersteunende technieken functioneren in het werk van *medisch personeel:* de verpleegkundigen, artsen en anderen die verantwoordelijk zijn voor de medische zorg. Allereerst wordt een typering gegeven van dit werk in huidige medische praktijken. Anders dan de eerder beschreven discoursen stellen, blijkt medisch werk een in concrete materiële en sociale contexten gesitueerde activiteit te zijn. Medisch handelen wordt niet gekenmerkt door omschreven beslisregels en een stapsgewijze aanpak van problemen, bijvoorbeeld, maar door uiterst fluïde criteria, "gegevens" die continue worden ge(re)construeerd, en "beslissingen" die slecht post-hoc als zodanig kunnen worden aangewezen.

Paradoxaal genoeg is deze aard van het medisch handelen zowel een probleem als de "redding" voor beslissingsondersteunende technieken. Daar waar disciplinering niet adequaat is, waar de vereisten van een formele techniek niet dreigen te worden gerealiseerd, is het juist het medisch personeel dat op dezelfde ad hoc en reactieve wijze de dreigende gaten dicht. Juist doordat de disciplinering van medisch personeel niet "totaal" is kunnen de technieken dus functioneren.

Dit fenomeen creëert echter wel een groot probleem voor een andere claim van de voorstanders. Daar waar gesteld wordt dat beslissingsondersteunende technieken artsen "ondersteunen", en het werk van medisch personeel kunnen "verlichten", blijken deze technieken juist *extra* verantwoordelijkheden te introduceren: het is het medisch personeel dat continu klaar moet staan om de vereisten van de technologie te produceren. Het functioneren van de technieken, met alle consequenties die daaraan vast zitten, wordt ten dele de verantwoordelijkheid van diegenen wiens werk hierdoor "vereenvoudigd" zou worden.

Ook de positie van de critici is hier weer in het geding. Beslissingsondersteunende technieken kunnen wel degelijk in een praktijk functioneren, zo blijkt. Met het actief produceren van de vereisten van formele technieken (door zowel systeem bouwers als door medisch personeel) worden de mogelijkheidsvoorwaarden gecreëerd waardoor deze technieken hun *coördinerende* (in het geval van protocollen) of *accumulerende* (expert systemen, statistische technieken) rollen kunnen vervullen. *Door* deze technieken kunnen medewerkers bijvoorbeeld meer complexe taken vervullen.

Tegelijkertijd wordt het door het op de in dit hoofdstuk beschreven wijze van functioneren van beslissingsondersteunende technieken ook volstrekt onmogelijk om nog op een eenduidige manier de "eindverantwoordelijkheid" voor medische beslissingen óf aan de techniek óf aan de arts toe te wijzen. Het is juist in de interactie tussen medisch personeel en techniek dat de nieuwe beslissings- of handelings-strategieën tot leven komen; "controle" en "verantwoordelijkheid" worden begrippen die alleen aan *hybrides* van mensen en dingen zijn toe te wijzen.

Hoofdstuk 6 gaat tenslotte in op de discussie tussen voor- en tegenstanders over het gebrek aan succes van de meeste beslissingsondersteu-

nende technieken. In dit hoofdstuk worden de vooronderstellingen van voor- en tegenstanders nog eens kritisch op een rij gezet en bekritiseerd. Beide gaan uit van vaststaande ideeën over wat beslissingsondersteunende technieken en wat medische praktijken zijn; beide zetten Techniek en Praktijk *tegenover* elkaar in plaats van ze (zowel historisch als in actuele praktijken) aan elkaar te relateren. Door middel van een recapitulatie van de centrale punten van dit boek wordt betoogd dat een dergelijke benadering *de verkeerde vragen* genereert. In plaats van trachten een *verklaring* van het "succes" of "falen" te zoeken door de essentiële eigenschappen van "techniek" of "praktijk" tegenover elkaar te stellen is het vruchtbaarder te onderzoeken hoe praktijken worden getransformeerd en waar de vervlechting van formele technieken en "informele" handelingspraktijken in resulteert. In plaats van het veroordelen van "medische praktijken" als "irrationeel" of "onwetenschappelijk", tenslotte, of van beslissingsondersteunende technieken als "onwerkbaar", is het zinvoller te onderzoeken op welke manier beslissingsondersteunende technieken interveniëren in de praktijken waar ze deel van uitmaken, welke rationaliteiten worden versterkt en welke verzwakt, waar welke consequenties voelbaar zijn, en welke consequenties wenselijker zijn dan andere. Beslissingsondersteunende technieken zijn noch a priori goed noch a priori slecht; het zijn technieken met grote potenties maar met minstens even grote problemen. Het onderkennen van deze potenties en problemen opent de weg naar een (letterlijk) meer constructieve kritiek: het creëren van beslissingsondersteunende technieken welke daadwerkelijk bijdragen aan noodzakelijk geachte ontwikkelingen.

References

Abbott, A. 1988. *The System of Professions. An Essay on the Division of Expert Labor.* Chicago: University of Chicago Press.

Adams, I. D. et al. 1986. Computer Aided Diagnosis of Acute Abdominal Pain: A Multicentre Study. *British Medical Journal* 293: 800-4.

Agre, P. E. forthcoming. From High Tech to Human Tech: on the Sudden Market for Social Studies of Technology. In *Social Science, Technical Systems, and Cooperative Work,* edited by G. Bowker, L. Gasser, L. Star, and B. Turner, MIT Press.

Agre, P. E. 1994. Design for Democracy. San Diego: Department of Communication, University of California. Unpublished manuscript.

Aikins, J. S., J. C. Kunz, E. H. Shortliffe, and R. J. Fallat. 1983. PUFF: An Expert System for Interpetation of Pulmonary Function Data. *Computers in Biomedical Research* 16: 199-208.

Akrich, M. 1992. The De-scription of Technical Objects. In *Shaping Technology, Building Society,* edited by W. E. Bijker, and J. Law, 205-224.

Akrich, M., and B. Latour. 1992. A Summary of a Convenient Vocabulary for the Semiotics of Human and Non-Human Assemblies. In *Shaping Technology, Building Society,* edited by W. E. Bijker, and J. Law, 259-264.

Allison, G. T. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis.* Boston: Little, Brown & Co.

Alvarez, W. 1953. On Disregarding Findings that Cannot Explain the Syndrome. *New England Journal of Medicine* 249: 184-6.

Amann, K., and K. Knorr-Cetina. 1989. Thinking Through Talk: An Ethnographic Study of a Molecular Biology Laboratory. *Knowledge and Society* 8: 3-26.

American Board of Internal Medicine. 1979. Clinical Competence in Internal Medicine. *Annals of Internal Medicine* 90: 402-11.

Andersen, T. F., and G. Mooney, ed. 1990. *The Challenges of Medical Practice Variations.* London: Macmillan Press.

Anderson, J. G., and S. J. Jay, ed. 1987. *Use and Impact of Computers in Clinical Medicine.* New York: Springer-Verlag.

Anderson, W. 1992. The Reasoning of the Strongest: the Polemics of Skill and Science in Medical Diagnosis. *Social Studies of Science* 22: 653-684.

Anonymous. 1922. Medical English as She is Wrote. *Journal of the American Medical Association* 78: 731-2.

Anonymous. 1926. Filling the Physician's Prescription. *Journal of the American Medical Association* 87: 35-6.

Anonymous. 1938. Master in the House of Medicine. *Journal of the American Medical Association* 111: 327.

Anonymous. 1945. A.M.A. Program. *New England Journal of Medicine* 233: 223.

Anonymous. 1946. The Inadequacies of Medical Care. I. General Considerations. *New England Journal of Medicine* 234: 515-6.

Anonymous. 1947a. Basic Science Instruction in Hospital Residencies. *Journal of the American Medical Association* 133: 697-8.

Anonymous. 1947b. Diagnostic Medical Care and Commercial Laboratories. *Journal of the American Medical Association* 134: 457-8.

Anonymous. 1948a. "Unnecessary Operations". *New England Journal of Medicine* 238: 339-40.

Anonymous. 1948b. Congress Looks at Health. *New England Journal of Medicine* 239: 102.

Anonymous. 1948c. The Providence Evening Bulletin Looks at Medicine. *New England Journal of Medicine* 239: 208-9.

Anonymous. 1948d. The Public Wants a Doctor When They Want Him. *Journal of the American Medical Association* 136: 695-6.

Anonymous. 1950a. Enrolments in Medical Schools. *Journal of the American Medical Association* 142: 420.

Anonymous. 1950b. The Next Fifty Years. *Journal of the American Medical Association* 142: 34-5.
Anonymous. 1950c. Quality of Medical Care. *New England Journal of Medicine* 242: 381-2.
Anonymous. 1950d. Roadblocks to Clinical Research. *New England Journal of Medicine* 243: 677.
Anonymous. 1951a. "C.P.C." - Clinical Analysis or Guessing Game? *New England Journal of Medicine* 245: 829-30.
Anonymous. 1951b. Physician-patient Relationship. *Journal of the American Medical Association* 147: 1054.
Anonymous. 1952a. "Quality of Medical Care". *New England Journal of Medicine* 247: 34-5.
Anonymous. 1952b. Clinical Accountancy. *New England Journal of Medicine* 247: 142-3.
Anonymous. 1952c. The Physicians Responsibility. *Journal of the American Medical Association* 148: 471.
Anonymous. 1954a. A Confusion of Tongues. *Journal of the American Medical Association* 154: 1093.
Anonymous. 1954b. Uses of Standard Nomenclature. *Journal of the American Medical Association* 154: 586-7.
Anonymous. 1956a. "Medical Science and Society". *New England Journal of Medicine* 254: 1190-1.
Anonymous. 1956b. The Epidemiologic Method. *New England Journal of Medicine* 254: 1044-5.
Anonymous. 1956c. Panorama of Medical Research. *Journal of the American Medical Association* 162: 209.
Anonymous. 1957a. Sequential Probabilities. *New England Journal of Medicine* 256: 524.
Anonymous. 1957b. Standard Nomenclature of Disease and Operations. *Journal of the American Medical Association* 163: 552.
Anonymous. 1958. From Guesswork to Guideline. *Journal of the American Medical Association* 166: 781.
Anonymous. 1959. President Eisenhower's Atlantic City Adress. *Journal of the American Medical Association* 170: 1072.
Anonymous. 1960. Medical Research. *Journal of the American Medical Association* 173: 685.
Anonymous. 1961a. Computers and the Practice of Medicine. *Journal of the American Medical Association* 177: 205-6.
Anonymous. 1961b. Development of an Industrial Medical-Records System. *Journal of the American Medical Association* 176: 805.
Anonymous. 1962. "Doctors' Dilemmas". *New England Journal of Medicine* 266: 1335.
Anonymous. 1963a. Electronic Data Processing Apparatus. *Journal of the American Medical Association* 186: 146-7.
Anonymous. 1963b. Judgment Difficult. *New England Journal of Medicine* 269: 1383-4.
Anonymous. 1964. Physical Diagnosis. *New England Journal of Medicine* 270: 476-7.
Anonymous. 1965a. Physical Diagnosis: Requim or Renaissance? *Journal of the American Medical Association* 192: 164.
Anonymous. 1965b. Rending the Veil of Mystery. *Journal of the American Medical Association* 192: 128.
Anonymous. 1967. Report of the Bethesda Conferences of the Committee on Standardized Terminology of the American College of Cardiology. Glossary of Cardiologic Terms Related to Physical Diagnosis and History. I. Heart Murmurs. *Journal of the American Medical Association* 200: 1041-2.
Anonymous. 1968. The Cost of Life. *Journal of the American Medical Association* 204: 923-4.
Anonymous. 1982. Algorithms for the Clinician. *Lancet* 528-9.
Anspach, R. R. 1993. *Deciding who lives. Fateful Choices in the Intensive Care Nursery.* Berkeley: University of California Press.
Aring, C. D. 1970. The Patient's Welfare and the Medical Record. *Journal of the American Medical Association* 214: 1317-9.
Arkes, H. R., and K. R. Hammond, ed. 1986. *Judgment and Decision Making.* Cambridge: Cambridge University Press.
Armstrong, D. 1983. *Political Anatomy of the Body. Medical Knowledge in Britain in the Twentieth Century.* Cambridge: Cambridge University Press.
Arney, W., and B. Bergen. 1984. *Medicine and the Management of Living. Taming the Last Great Beast.* Chicago: University of Chicago Press.

Ashmore, M., M. Mulkay, and T. Pinch. 1989. *Health and Efficiency: A Sociology of Health Economics*. Milton Keynes: Open University Press.
Ashmore, M., R. Wooffitt, and S. Harding. 1994. *Humans and Others, Agents and Things*. Special Issue of *American Behavioral Scientist* 37 (6).
Atkinson, P. 1981. *The Clinical Experience: The Construction of Medical Reality*. Farnborough: Gower.
Audet, A. M., S. Greenfield, and M. Field. 1990. Medical Practice Guidelines: Current Activities and Future Directions. *Annals of Internal Medicine* 113: 709-14.
Baars, B. J. 1986. *The Cognitive Revolution in Psychology*. New York: Guilford Press.
Baker, G. P., and P. M. S. Hacker. 1985. *Wittgenstein, Rules, Grammar, and Necessity*. Oxford: Basil Blackwell.
Balla, J. I., A. S. Elstein, and C. Christensen. 1989. Obstacles to Acceptance of Clinical Decision Analysis. *British Medical Journal* 4: 579-82.
Barley, S. R. 1988. The Social Construction of a Machine: Ritual, Superstition, Magical Thinking and Other Pragmatic Responses to Running a CT Scanner. In *Biomedicine Examined*, edited by M. Lock, and D. R. Gordon, 497-539.
Barnett, G. 1968. Computers in Patient Care. *New England Journal of Medicine* 279: 1321-7.
Barnett, G. O. 1982. The Computer and Clinical Judgement. *New England Journal of Medicine* 307: 493-4.
Barnett, G. O., J. J. Cimono, J. A. Hupp, and E. P. Hoffer. 1987. DXplain. An Evolving Diagnostic Decision Support System. *Journal of the American Medical Association* 258: 67-74.
Beecher, H. K. 1953. Clinical Impression and Clinical Investigation. *Journal of the American Medical Association* 151: 44-5.
Beeson, P. B., W. M. Dermott, and J. B. Wijngaarden, ed. *Cecil Textbook of Medicine*. Philadelphia: W B Saunders.
Bell, D. E., H. Raiffa, and A. Tversky. 1988. Descriptive, Normative and Prescriptive Interactions in Decision Making. In *Decision Making. Descriptive, Normative, and Prescriptive Interactions*, edited by ibid., 1-30. Cambridge: Cambridge University Press.
Beniger, J. R. 1986. *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge: Harvard University Press.
Berg, M. 1992. The Construction of Medical Disposals. Medical Sociology and Medical Problem Solving in Clinical Practice. *Sociology of Health and Illness* 14: 151-180.
Berg, M. 1993. A Latourian Outsider Looks At CSCW. On Protocols: Tools Which Shouldn't Work But Do. Paper presented at *1993 Oksnoen Symposium on Understanding Work Practice*, 21-26 May, Oeksnoen, Norway.
Berg, M. 1994. 'Ze zijn allemaal dood...'. Over kennissystemen in de medische praktijk, voorvechters en critici, en het mysterieuze succes van onderzoeksprotocollen. *Kennis en Methode* 18: 361-91.
Berg, M., and A. Mol, ed. forthcoming. *Differences in Medicine. Unraveling Practices, Techniques and Bodies*.
Berwick, D. M. 1988. Presidential address. The Society for Medical Decision Making: the right place at the right time. *Medical Decision Making* 8: 77-80.
Bijker, W. E., T. P. Hughes, and T. J. Pinch, ed. 1987. *The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology*. Cambridge: MIT Press.
Bijker, W. E., and J. Law, ed. 1992. *Shaping Technology - Building Society. Studies in Sociotechnical Change*. Cambridge: MIT Press.
Bishop, L. F. 1900. A Plea for Greater Simplicity in Therapeutics. *Journal of the American Medical Association* 35: 1332-3.
Bjerregaard, B. et al. 1976. Computer-Aided Diagnosis of the Acute Abdomen: A System from Leeds used on Copenhagen Patients. In *Decision Making and Medical Care*, edited by F. T. de Dombal, and F. Gremy, 165-171.
Bleich, H. L. 1971. Prognosis by Calculation. *New England Journal of Medicine* 285: 1533-4.
Bleich, H. L. 1972. Computer-Based Consultation. Electrolyte and Acid-Base Disorders. *American Journal of Medicine* 53: 285-91.
Blois, M. S. 1980. Clinical Judgment and Computers. *New England Journal of Medicine* 303: 192-7.

Bloomfield, B. P. 1991. The Role of Information Systems in the UK National Health Service: Action at a Distance and the Fetish of Calculation. *Social Studies of Science* 21: 701-34.
Bloor, M. 1976. Bishop Berkeley and the Adeno-Tonsillectomy Enigma. *Sociology* 10: 43-61.
Bloor, M. 1978. On the Routinised Nature of Work in People-Processing Agencies: The Case of Adenotonsillectomy Assessments in ENT Out-Patient Clinics. In *Relationships between Doctors and Patients*, edited by A. Davis, 29-47. Farnborough: Gower.
Blume, S. S. 1992. *Insight and Industry: On the Dynamics of Technological Change in Medicine*. Cambridge, MA: MIT Press.
Böckenholt, U., and E. U. Weber. 1992. Use of Formal Methods in Medical Decision Making. *Medical Decision Making* 12: 298-306.
Boon, L. 1982. *Geschiedenis van de psychologie*. Meppel: Boom.
Borak, J., and S. Veilleux. 1982. Errors of Intuitive Logic Among Physicians. *Social Science and Medicine* 16: 1939-1947.
Bosk, C. L. 1979. *Forgive and Remember: Managing Medical Failure*. Chicago: University of Chicago Press.
Bowers, J. 1992. The Politics of Formalism. In *Contexts of Computer Mediated Communication*, edited by M. Lea, Hassocks: Harvester.
Bowker, G. 1994a. Dismembering and Remembering: Classification and Organizational Memory. Paper presented at Conference *Locating Design, Development and Use*, Oksnoen, May 13-18.
Bowker, G. 1994b. *Science on the Run. Information Management and Industrial Geophysics at Schlumberger, 1920-1940*. Cambridge, MA: MIT Press.
Bowker, G., and S. L. Star. 1994. Knowledge and Infrastructure in International Information Management: Problems of Classification and Coding. In *Information Acumen: the Understanding and Use of Knowledge in Modern Business*, edited by L. Bud, 187-213. London: Routledge.
Braybrooke, D., and C. E. Lindblom. 1963. *A Strategy of Decision. Policy Evaluation as a Social Process*. New York: Free Press.
Brook, R. H. et al. 1988. Diagnosis and Treatment of Coronary Disease: Comparison of Doctors' Attitudes in the USA and the UK. *Lancet* April 2: 750-753.
Buchanan, B. G., and E. H. Shortliffe, ed. 1984. *Rule-Based Expert Systems. The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley.
Bursztajn, H. et al. 1981. *Medical Choices, Medical Chances. How Patients, Families and Physicians can Cope with Uncertainty*. New York: Dell.
Button, G., ed. 1993. *Technology in Working Order. Studies of Work, Interaction, and Technology*. London: Routledge.
Callon, M. 1980. Struggles and Negotiations to Decide What Is Problematic and What Is Not: The Sociology of Translation. In *The Social Process of Scientific Investigation*, edited by K. Knorr, R. Krohn, and R. Whitley, 197-220. Dordrecht: Reidel.
Callon, M. 1987. Society in the Making: The Study of Technology as a Tool for Sociological Analysis. In *The Social Construction of Technological Systems*, edited by W. E. Bijker, T. P. Hughes, and T. Pinch, 83-103.
Callon, M. 1991. Techno-economic Networks and Irreversibility. In *A Sociology of Monsters*, edited by J. Law, 132-161.
Casparie, A. F., and J. J. E. Everdingen. 1989. Maatstaven voor goede zorg: vijf jaar later. *Medisch Contact* 44: 1654-6.
Casper, M., and M. Berg, ed. 1995. *Constructivist Perspectives on Medical Practices*. Special Issue of *Science, Technology and Human Values* 20 (3).
Cebul, R. D. 1984. A Look at the Chief Complaints Revisited: Current Obstacles and Opportunities for Decision Analysis. *Medical Decision Making* 4: 271-83.
Cebul, R. D. 1988. Decision Making Research at the Interface between Descriptive and Prescriptive Studies. *Medical Decision Making* 8: 231-2.
Chatbanchai, W. et al. 1989. Acute Abdominal Pain and Appendicitis in North East Thailand. *Paediatric and Perinatal Epidemiology* 3: 448-59.
Cicourel, A. V. 1986. The Reproduction of Objective Knowledge: Common Sense Reasoning in Medical Decision Making. In *The Knowledge Society*, edited by G. Böhme, and N. Stehr, 87-122. Dordrecht: Reidel.
Clancey, W. J., and E. H. Shortliffe, ed. 1984. *Readings in Medical Artificial Intelligence: The First Decade*. Reading, MA: Addison-Wesley.

Clancy, C. M., R. D. Cebul, and S. V. Williams. 1988. Guiding Individual Decisions: A Randomized, Controlled Trial of Decision Analysis. *American Journal of Medicine* 84: 283-8.
Clarke, A. E., and J. H. Fujimura, ed. 1992. *The Right Tools for the Job. At Work in Twentieth-century Life Sciences.* Princeton, N. J.: Princeton University Press.
Clarke, J. R. 1989. Appendicitis. The Computer as a Diagnostic Tool. *International Journal of Technology Assessment in Health Care* 5: 371-9.
Clinton, J. J. 1992. Improving Clinical Practice. *Journal of the American Medical Association* 267: 2652-3.
Clouser, K. D. 1985. Approaching the Logic of Diagnosis. In *Logic of Discovery and Diagnosis in Medicine,* edited by K. F. Schaffner, 35-55. Berkeley: University of California Press.
Collings, J. S., and D. M. Clark. 1953. General Practice, Today and Tomorrow. *New England Journal of Medicine* 248: 141-8,183-94.
Collins, H. M. 1985. *Changing Order: Replication and Induction in Scientific Practice.* London: Sage.
Collins, H. M. 1990. *Artificial Experts. Social Knowledge and Intelligent Machines.* Cambridge: MIT Press.
Cornfield, J. 1972. Statistical Classification Methods. In *Computer Diagnosis and Diagnostic Methods,* edited by J. A. Jacquez, 108-130.
Coulter, J. 1983. *Rethinking Cognitive Theory.* New York: St. Martin's Press.
Coulter, J. 1989. *Mind in Action.* Cambridge: Polity Press.
Crosby, W. H. 1977. Chess and Combat: The Algorithm in Medicine. *Journal of the American Medical Association* 238: 2721.
Cummins, R. O. 1980. Clinicians' Reasons for Overuse of Skull Radiographs. *American Journal of Roentgenology* 135: 549-52.
Cummins, R. O. 1990. Decision Analysis, the Journal of General Internal Medicine, and the General Internist. *Journal of General Internal Medicine* 5: 375-8.
Cussins, C. forthcoming. Ontological Choreography. Agency for Women Patients in an Infertility Clinic. In *Differences in Medicine,* edited by M. Berg, and A. Mol.
Danilevicius, Z. 1975. The Power of Unbiased Observation. *Journal of the American Medical Association* 231: 966.
Davis, K. 1986. The Process of Problem (Re)Formulation in Psychotherapy. *Sociology of Health and Illness* 8: 44-74.
Davis, R., B. G. Buchanan, and E. H. Shortliffe. 1977. Production Rules as a Representation for a Knowledge-Based Consultation Program. *Artificial Intelligence* 8: 15-45.
de Bont, A. 1994. De goede dokter. Vijf variaties op een misverstand. *Gezondheid: Theorie in Praktijk* 2: 249-264.
de Dombal, F. T. 1987a. Back to the Future; Or Forward to the Past? *Gut* 28: 373-7.
de Dombal, F. T. 1987b. Computer-aided Decision Support - The Obstacles to Progress. *Methods of Information in Medicine* 26: 183-4.
de Dombal, F. T. 1988. The OMGE Acute Abdominal Pain Survey. Progress Report, 1986. *Scandinavian Journal of Gastroenterology* 23 (supplement 144): 35-42.
de Dombal, F. T. 1989. Computer-aided Decision Support in Clinical Medicine. *International Journal of Biomedicine and Computing* 24: 9-16.
de Dombal, F. T. 1990. Computer-aided Decision Support - Glittering Prospects, Practical Problems, and Pandora's Box. *Baillière's Clinical Obstetrics and Gynaecology* 4: 841-9.
de Dombal, F. T. 1991. The Diagnosis of Acute Abdominal Pain with Computer Assistance: Worldwide Perspective. *Annales de Chirurgie* 45: 273-7.
de Dombal, F. T., V. Dallos, and W. A. McAdam. 1991. Can Computer Aided Teaching Packages Improve Clinical Care in Patients with Acute Abdominal Pain? *British Medical Journal* 302: 1495-7.
de Dombal, F. T., and F. Gremy, ed. 1976. *Decision Making and Medical Care.* Amsterdam: North Holland Publishing Co.
de Dombal, F. T., D. J. Leaper, J. R. Staniland, A. P. McCann, and J. C. Horrocks. 1972. Computer-Aided Diagnosis of Acute Abdominal Pain. *British Medical Journal* 2: 9-13.
de Vries, G. 1993. *Gerede Twijfel: Over de Rol van de Medische Ethiek in Nederland.* Amsterdam: De Balie.
de Vries, P. H., and P. F. de Vries-Robbe. 1985. An Overview of Medical Expert Systems. *Methods of Information in Medicine* 24: 57-64.

Dent, M. 1990. Organisation and Change in Renal Work: A Study of the Impact of a Computer System within Two Hospitals. *Sociology of Health and Illness* 12: 413-431.
Detsky, A. S. 1987. Decision Analysis: What's the Prognosis? *Annals of Internal Medicine* 106: 321-322.
Detsky, A. S., D. Redelmeier, and H. B. Abrams. 1987. What's Wrong with Decision Analysis? Can the Left Brain Influence the Right? *Journal of Chronic Diseases* 40: 831-8.
Dodier, N. forthcoming. Clinical Practice and Procedures in Occupational Medicine. A Study of the Framing of Individuals. In *Differences in Medicine*, edited by M. Berg, and A. Mol.
Dolan, J. G. 1990. Can Decision Analysis Adequately Represent Clinical Problems? *Journal of Clinical Epidemiology* 43: 277-84.
Doubilet, P., and B. J. McNeil. 1985. Clinical Decisionmaking. *Medical Care* 23: 648-62.
Doyle, J. C. 1952. Unnecessary Ovariectomies. Study Based on Removal of 704 Normal Ovaries from 546 Patients. *Journal of the American Medical Association* 148: 1105-11.
Dreyfus, H. 1992. *What Computers Still Can't Do: a Critique of Artificial Reason*. Cambridge: MIT Press.
Dreyfus, H., and S. E. Dreyfus. 1986. *Mind over Machine. The Power of Human Intuition and Expertise in the Era of the Computer*. Oxford: Blackwell.
Duda, R. O., and E. H. Shortliffe. 1983. Expert Systems Research. *Science* 220: 261-8.
Eagle, K. E. 1991. Medical Decision Making in Patients with Chest Pain. *New England Journal of Medicine* 324: 1282-3.
Ebbesen, E. B., and V. J. Konecni. 1980. On the External Validity of Decision-Making Research: What Do we Know about Decisions in the Real World? In *Cognitive Processes, Choice and Decision Behavior*, edited by T. S. Wallsten, 691-7. Hillsdale, NJ: Lawrence Erlbaum Ass.
Eddy, D. M. 1982. Probabilistic Reasoning in Clinical Medicine: Problems and Opportunities. In *Judgment under Uncertainty*, edited by D. Kahneman, P. Slovic, and A. Tversky, 249-67.
Eddy, D. M. 1990a. Anatomy of a Decision. *Journal of the American Medical Association* 263: 441-3.
Eddy, D. M. 1990b. The Challenge. *Journal of the American Medical Association* 263: 287-90.
Eddy, D. M. 1990c. Practice Policies - What Are They? *Journal of the American Medical Association* 263: 877-8,880.
Eddy, D. M. 1990d. Practice Policies: Where Do They Come From? *Journal of the American Medical Association* 263: 1265-9,1272,1275.
Eddy, D. M. 1990e. Practice Policies - Guidelines for Methods. *Journal of the American Medical Association* 263: 1839-41.
Eddy, D. M. 1990f. Guidelines for Policy Statements: the Explicit Approach. *Journal of the American Medical Association* 263: 2239-40, 2243.
Eddy, D. M. 1992. Applying Cost-Effectiveness Analysis. The Inside Story. *Journal of the American Medical Association* 268: 2575-82.
Edwards, D. 1994. Imitations and Artifice in Apes, Humans, and Machines. *American Behavioral Scientist* 37: 754-71.
Edwards, P. N. Forthcoming. *The Closed World. Computers and the Politics of Discourse in Cold War America*. Cambridge: MIT Press.
Eisenberg, J. M. 1986. *Doctors' Decisions and the Cost of Medical Care*. Ann Arbor, MI: Health Administration Press.
Elstein, A. S. et al. 1986. Comparison of Physicians' Decisions Regarding Estrogen Replacement Therapy for Menopausal Women and Decisions Derived from a Decision Analytic Model. *American Journal of Medicine* 80: 246-58.
Elstein, A. S., L. S. Schulman, and S. A. Sprafka. 1978. *Medical Problem Solving*. Cambridge: Harvard University Press.
Elstein, A. S., L. S. Shulman, and S. A. Sprafka. 1990. Medical Problem Solving: A Ten-Year Retrospective. *Evaluation and the Health Professions Special Issue: Reflections on Research in Medical Problem Solving* 13: 5-36.
Emerson, P. A., J. Wyatt, L. Dillistone, N. Crichton, and N. J. Russell. 1988. The Development of ACORN, an Expert System Enabling Nurses to Make Admission Decisions about Patients with Chest Pain in an Accident & Emergency Department. In *Proceedings of Medical Informatics: Computers in Clinical Medicine*, 37-40. London: British Medical Informatics Society.

Engle, R. L. 1992. Attempts to Use Computers as Diagnostic Aids in Medical Decision Making: A Thirty-Year Experience. *Perspectives in Biology and Medicine* 35: 207-19.
Epstein, S. 1995. The critique of pure science: AIDS activism and the struggle for credibility in the reform of clinical trials. In *Constructivist Perspectives on Medical Practices*, edited by M. Casper, and M. Berg.
Evans, R. G. 1990. The Dog in the Night-time: Medical Practice Variations and Health Policy. In *The Challenges of Medical Practice Variations*, edited by T. F. Anderson, and G. Mooney, 402-11.
Everdingen, J. J. E. 1988. *Consensusontwikkeling in de geneeskunde*. Utrecht: Bohn, Scheltema en Holkema.
Feinstein, A. R. 1967. *Clinical Judgment*. Huntington, NY: R.E. Krieger Publishing Co.
Feinstein, A. R. 1973a. An Analysis of Diagnostic Reasoning. I. The Domains and Disorders of Clinical Macrobiology. *Yale Journal of Biology and Medicine* 46: 212-232.
Feinstein, A. R. 1973b. An Analysis of Diagnostic Reasoning. II. The Strategy of Intermediate Decisions. *Yale Journal of Biology and Medicine* 46: 264-83.
Feinstein, A. R. 1973c. The Problems of the "Problem-Oriented Medical Record". *Annals of Internal Medicine* 78: 751-62.
Feinstein, A. R. 1974. An Analysis of Diagnostic Reasoning. III. The Construction of Clinical Algorithms. *Yale Journal of Biology and Medicine* 1: 5-32.
Feinstein, A. R. 1977. Clinical Biostatistics XXXIX. The Haze of Bayes, the Aerial Palaces of Decision Analysis, and the Computerized Ouija Board. *Clinical Pharmacology and Therapeutics* 21: 482-96.
Feinstein, A. R. 1979. Science, Clinical Medicine, and the Spectrum of Disease. In *Cecil Textbook of Medicine*, edited by P. B. Beeson, W. M. Dermott, and J. B. Wijngaarden, 3-6.
Feinstein, A. R. 1985. *Clinical Epidemiology. The Architecture of Clinical Research*. Philadelphia: W B Saunders.
Feinstein, A. R. 1987a. *Clinimetrics*. New Haven: Yale University Press.
Feinstein, A. R. 1987b. The Intellectual Crisis in Clinical Science: Medaled Models and Muddled Mettle. *Perspectives in Biology and Medicine* 30: 215-30.
Feinstein, A. R. 1994. Clinical Judgment Revisited: The Distraction of Quantitative Models. *Annals of Internal Medicine* 120: 799-805.
Fenyo, G. et al. 1987. Computer-aided Diagnosis of 233 Acute Abdominal Cases at Nacka Hospital Sweden. *Scandinavian Journal of Gastroenterology* Supplement 128: 178.
Field, M. J., and K. N. Lohr, ed. 1990. *Clinical Practice Guidelines: Directions for a New Program*. Washington, DC: National Academy Press.
Fischoff, B. 1988. Clinical Decision Making. In *Professional Judgment: A Reader in Clinical Decision Making*, edited by J. Dowie, and A. Elstein, 409-24. Cambridge: Cambridge University Press.
Fisher, S., and A. D. Todd, ed. 1983. *The Social Organization of Doctor-Patient Communication*. Washington D.C.: Center for Applied Linguistics.
Flanagin, A., and G. D. Lundberg. 1990. Clinical Decision Making: Promoting the Jump from Theory to Practice. *Journal of the American Medical Association* 263: 279-80.
Flexner, A. 1910. *Medical Education in the United States and Canada*. New York: Carnegie Foundation for the Advancement of Teaching.
Forsythe, D. E. 1992. Blaming the User in Medical Informatics: The Cultural Nature of Scientific Practice. *Knowledge and Society:* 9: 95-111.
Forsythe, D. E. 1993a. The Construction of Work in Artificial Intelligence. *Science, Technology and Human Values* 18: 460-479.
Forsythe, D. E. 1993b. Engineering Knowledge: the Construction of Knowledge in Artificial Intelligence. *Social Studies of Science* 23: 445-477.
Friedman, A., and D. S. Comford. 1989. *Computer Systems Development. History, Organization and Implementation*. New York: Wiley.
Frohock, F. M. 1986. *Special Care. Medical Decisions at the Beginning of Life*. Chicago: University of Chicago Press.
Froom, J. 1975. International Classification of Health Problems in Primary Care. *Journal of the American Medical Association* 234: 1257.
Fujimura, J. H. 1987. Constructing 'Do-Able' Problems in Cancer-Research: Articulating Alignment. *Social Studies of Science* 17: 257-293.
Fujimura, J. H. 1988. The Molecular Biological Bandwagon in Cancer Research: Where Social Worlds Meet. *Social Problems* 35: 261-283.

Fujimura, J. H. 1992. Crafting Science: Standardized Packages, Boundary Objects, and 'Translation'. In *Science as Practice and Culture*, edited by A. Pickering, 168-211.
Fuller, S. 1991. Harry the Apostate (or PC-Squared) (Response to Collins' *Artificial Experts*). Paper presented at *Annual Meeting of the Society of the Social Studies of Science*, Boston, November 1991.
Fuller, S. 1993a. *Philosophy, Rhetoric, and the End of Knowledge: the Coming of Science and Technology Studies*. Madison: The University of Wisconsin Press.
Fuller, S. 1993b. What Dreyfus Still Can't See (Review of Dreyfus' *What Computers Still Can't Do*). *EAST-Newsletter* 12: 11-15.
Garfinkel, H. 1967. *Studies in Ethnomethodology*. Englewood-Cliffs, N.Y.: Prentice-Hall.
Garfinkel, H., and H. Sacks. 1970. On Formal Structures of Practical Actions. In *Theoretical Sociology. Perspectives and Developments*, edited by J. C. Mckinney, and E. A. Tiryakian, 338-366. New York: Appleton Century Crofts.
Garland, L. H. 1959. Studies on the Accuracy of Diagnostic Procedures. *American Journal of Roentgenology* 82: 25-38.
Gasser, L. 1986. The Integration of Computing and Routine Work. *ACM Transactions on Office Information Systems* 4: 205-225.
Gigerenzer, G., and D. J. Murray. 1987. *Cognition as Intuitive Statistics*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Gigerenzer, G. et al. 1989. *The Empire of Chance. How Probability Changed Science and Everyday Life*. Cambridge, MA: Cambridge University Press.
Gill, P. W. et al. 1973. Observer Variation in Clinical Diagnosis - A Computer-Aided Assessment of its Magnitude and Importance in 552 Patients with Abdominal Pain. *Methods of Information in Medicine* 12: 108-13.
Ginsberg, A. S. 1972. The Diagnostic Process Viewed as a Decision Problem. In *Computer Diagnosis and Diagnostic Methods*, edited by J. A. Jacquez, 5-9.
Goldwater, L. J. 1959. Clinical Judgment and Common Sense for the Cardiac Patient. *Journal of the American Medical Association* 169: 598.
Goldwater, L. J., L. H. Bronstein, and B. Kresky. 1952. Study of One Hundred Seventy-Five 'Cardiacs' Without Heart Disease. *Journal of the American Medical Association* 148: 89-92.
Gooding, D. 1992. Putting Agency Back into Experiment. In *Science as Practice and Culture*, edited by A. Pickering, 65-112.
Goody, J. 1977. *The Domestication of the Savage Mind*. New York: Cambridge University Press.
Gordon, D. R. 1988. Clinical Science and Expertise: Changing Boundaries Between Art and Science in Medicine. In *Biomedicine Examined*, edited by M. Lock, and D. R. Gordon, 257-295.
Gorry, G. A., J. P. Kassirer, A. Essig, and W. B. Schwartz. 1973. Decision Analysis as the Basis for Computer-Aided Management of Acute Renal Failure. *American Journal of Medicine* 55: 473-84.
Graves III, W., and J. M. Nyce. 1992. Normative Models and Situated Practice in Medicine. Towards More Adequate System Design and Development. *Information and Decision Technologies* 19: 143-149.
Greenbaum, J., and M. Kyng, ed. 1991. *Design at Work: Cooperative Design for Computer Systems*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
Greenberg, R. N. 1978. An Argument for Research in the Medical School Curriculum. *Journal of the American Medical Association* 239: 1162-3.
Greenfield, S., F. E. Bragg, D. L. McCraith, and J. Blackburn. 1974. Upper-respiratory Tract Complaint Protocol for Physician-Extenders. *Archives of Internal Medicine* 133: 294-9.
Grimm, R. H., K. Shimoni, W. R. Harlan, and E. H. Estes. 1975. Evaluation of Patient-Care Protocol Use by Various Providers. *New England Journal of Medicine* 292: 507-11.
Groen, G. J., and V. L. Patel. 1985. Medical Problem-Solving: Some Questionable Assumptions. *Medical Education* 19: 95-100.
Grol, R. 1989. De verspreiding van NHG standaarden onder huisartsen. *Huisarts en Wetenschap* 32: p.494-7.
Gustafson, D. H., R. L. Ludke, P. J. Glackman, F. C. Larson, and J. H. Greist. 1972. Wisconsin Computer Aided Medical Diagnosis Project - Progress Report. In *Computer Diagnosis and Diagnostic Methods*, edited by J. A. Jacquez, 255-278.
Haggerty, R. 1973. Effectiveness of Medical Care. *New England Journal of Medicine* 289: 372-3.

Haraway, D. J. 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature.* New York: Routledge.
Haraway, D. J. 1994. A Game of Cat's Cradle. Science Studies, Feminist Theory, Cultural Studies. *Configurations* 1: 59-71.
Harper, R. H. R., and J. A. Hughes. 1993. 'What a F-ing System! Send 'em all to the same place an then expect us to stop 'em hitting'. Making Technology Work in Air Traffic Control. In *Technology in Working Order,* edited by G. Button, 127-144.
Hart, A., and J. Wyatt. 1989. Connectionist Models in Medicine: an Investigation of their Potential. In *Proceedings AIME 89. Second European Conference on Artificial Intelligence in Medicine,* edited by J. Hunter, J. Cookson, and J. Wyatt, 115-124. Berlin: Springer-Verlag.
Hartland, J. 1993a. *The Machinery of Medicine. An Analysis of Algorithmic Approaches to Medical Knowledge and Practice.* PhD Thesis, University of Bath.
Hartland, J. 1993b. The Use of 'Intelligent' Machines for Electrocardiograph Interpretation. In *Technology in Working Order,* edited by G. Button, 55-80.
Haugeland, J. 1985. *Artificial Intelligence: the Very Idea.* Cambridge: MIT Press.
Hayes-Roth, F., D. A. Waterman, and D. B. Lenat. 1983. *Building Expert Systems.* London: Addison-Wesley.
Heathfield, H. A., and J. Wyatt. 1993. Philosophies for the Design and Development of Clinical Decision-Support Systems. *Methods of Information in Medicine* 32: 1-8.
Heims, S. J. 1980. *John von Neumann and Norbert Wiener. From Mathematics to the Technologies of Life and Death.* Cambridge: MIT Press.
Helman, C. G. 1988. Psyche, Soma and Society: The Social Construction of Psychosomatic Disorders. In *Biomedicine Examined,* edited by M. Lock, and D. R. Gordon, 95-122.
Heritage, J. 1984. *Garfinkel and Ethnomethodology.* Cambridge: Polity Press.
Hershey, J. C., and J. Baron. 1987. Clinical Reasoning and Cognitive Processes. *Medical Decision Making* 7: 203-11.
Hill, M. N., and C. S. Weisman. 1991. Physicians' Perceptions of Consensus Reports. *International Journal of Technology Assessment in Health Care* 7: 30-41.
Hirschauer, S. 1991. The Manufacture of Bodies in Surgery. *Social Studies of Science* 21: 217-319.
Hirschauer, S. forthcoming. Shifting Contradictions. Doing Sex and Doing Gender in Medical Disciplines. In *Differences in Medicine,* edited by M. Berg, and A. Mol.
Holmes-Rovner, M . 1992. Methods for Medical Decision Making. Presidential Address. *Medical Decision Making* 12: 159-62.
Horrocks, J. C., W. A. F. McAdam, G. Devroede, A. A. Gunn, and N. Zoltie. 1976. Some Practical Problems in Transferring Computer-Aided Diagnostic Systems from One Geographical Area to Another. In *Decision Making and Medical Care,* edited by F. T. de Dombal, and F. Gremy, 159-163.
Horrocks, J. C., A. P. McCann, J. R. Staniland, D. J. Leaper, and F. T. de Dombal. 1972. Computer-aided Diagnosis: Description of an Adaptable System, and Operational Experience with 2,034 Cases. *British Medical Journal* 2: 5-9.
Horstman, K. forthcoming. Chemical Analysis of Urine for Life Insurance: The Construction of Reliability. *Science, Technology and Human Values.*
Howard, J. M. 1961. Cooperative Clinical Research Programs between Medical Institutions. *Journal of the American Medical Association* 175: 705-6.
Hubbard, W. N. 1964. Welcome and Prologue. In *The Diagnostic Process,* edited by J. A. Jacquez, 3-5.
Hughes, D. 1988. When Nurse Knows Best: Some Aspects of Nurse/Doctor Interaction in a Casualty Department. *Sociology of Health and Illness* 10: 1-22.
Hunter, K. M. 1991. *Doctor's Stories. The Narrative Structure of Medical Knowledge.* Princeton: Princeton University Press.
Hurst, J. W. 1971. Ten Reasons Why Lawrence Weed is Right. *New England Journal of Medicine* 284: 51-2.
Ikonen, J. K. et al. 1983. Presentation and Diagnosis of Acute Abdominal Pain in Finland: A Computer Aided Study. *Annales Chirurgiae et Gynaecologiae* 72: 332-6.
Ingelfinger, F. J. 1973. Algorithms, Anyone? *New England Journal of Medicine* 288: 847-8.
Ingelfinger, F. J. 1975. Decision in Medicine. *New England Journal of Medicine* 293: 254-5.
Jacoby, I. 1988. Evidence and Consensus. *Journal of the American Medical Association* 259: 3039.

Jacoby, I., and S. G. Pauker. 1986. Technology Assessment in Health Care: Group Process and Decision Theory. *Israel Journal of Medical Sciences* 22: 183-90.
Jacquez, J. A. 1964. *The Diagnostic Process*. Ann Arbor, Michigan: Mallory Lithographic Company.
Jacquez, J. A. 1972. *Computer Diagnosis and Diagnostic Methods*. Springfield, IL: Charles C. Thomas.
Jencks, S. F. 1992. Accuracy in Recorded Diagnoses. *Journal of the American Medical Association* 267: 2238-39.
Jordan, K., and M. Lynch. 1992. The Sociology of a Genetic Engineering Technique: Ritual and Rationality in the Performance of the 'Plasmid Prep'. In *The Right Tools for the Job*, edited by A. E. Clarke, and J. H. Fujimura, 77-114.
Jordan, K., and M. Lynch. 1993. The Mainstreaming of a Molecular Biological Tool. A Case Study of a New Technique. In *Technology in Working Order*, edited by G. Button, 162-78.
Judd, W. H. 1960. Physicians in a Changing World. *New England Journal of Medicine* 263: 894-9.
Kahan, J. P., D. E. Kanouse, and J. D. Winkler. 1988. Stylistic Variations in National Institutes of Health Consensus Statements, 1979-1983. *International Journal of Technology Assessment in Health Care* 4: 289-304.
Kahneman, D., P. Slovic, and A. Tversky. 1982. *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
Kanouse, D. E., and I. Jacoby. 1988. When Does Information Change Practitioners' Behavior? *International Journal of Technology Assessment in Health Care* 4: 27-33.
Kanouse, D. E. et al. 1989. *Changing Medical Practice through Technology Assessment. An Evaluation of the NIH Consensus Development Program*. Ann Arbor, MI: Health Administration Press.
Kaplan, B. 1987. The Medical Computing 'Lag': Perceptions of Barriers to the Application of Computers to Medicine. *International Journal of Technology Assessment in Health Care* 3: 123.
Kaplan, B. 1995. The Computer Prescription: Medical Computing, Public Policy, and Views of History. *Science, Technology and Human Values* 20: 5-38.
Kassirer, J. P., and R. I. Kopelman. 1988. Intuitive and Inspirational, or Inductive and Incremental? *Hospital Practice* 23: 21-4,26-7.
Kassirer, J. P., and R. I. Kopelman. 1990. Diagnosis and Decisions by Algorithms. *Hospital Practice* 25: 23-4,27,31.
Kassirer, J. P., A. J. Moskowitz, J. Lau, and S. G. Pauker. 1987. Decision Analysis: A Progress Report. *Annals of Internal Medicine* 106: 275-91.
Kassirer, J. P., and S. G. Pauker. 1978. Should Diagnostic Testing be Regulated? *New England Journal of Medicine* 299: 947-9.
Kling, R. 1991. Computerization and Social Transformations. *Science, Technology, and Human Values* 16: 342-367.
Kling, R., and S. Iacono. 1984. The Control of Information Systems Developments After Implementation. *Communications of the ACM* 27: 1218-26.
Kling, R., and W. Scacchi. 1982. The Web of Computing: Computer Technology as Social Organization. *Advances in Computers* 21: 1-90.
Knaus, W. A. 1986. Rationing, Justice, and the American Physician. *Journal of the American Medical Association* 255: 1176-7.
Knaus, W. A., D. P. Wagner, and J. Lynn. 1991. Short-term Mortality Predictions for Critically Ill Hospitalized Adults: Science and Ethics. *Science* 254: 389-394.
Knorr-Cetina, K., and K. Amann. 1990. Image Dissection in Natural Scientific Inquiry. *Science, Technology and Human Values* 15: 259-283.
Knorr-Cetina, K. D. 1981. *The Manufacture of Knowledge*. Oxford: Pergamon Press.
Knottnerus, J. 1987. Bouwstenen voor een rationele medische besluitvorming. *Medisch Contact* 42: 501-4.
Komaroff, A. L. 1982. Algorithms and the 'Art' of Medicine. *American Journal of Public Health* 72: 10-2.
Komaroff, A. L. et al. 1974. Protocols for Physician Assistants. Management of Diabetes and Hypertension. *New England Journal of Medicine* 290: 307-12.
Komaroff, A. L., B. Reiffen, and H. Sherman. 1973. Problem-Oriented Protocols for Physician-Extenders. In *Applying the Problem-Oriented System*, edited by H. K. Walker, J. W. Hurst, and M. F. Woody, 186-96. New York: Medcom Press.

Koontz, A. R. 1959. The Humanities in Medicine. *Journal of the American Medical Association* 170: 72-3.
Korpman, R. A., and T. L. Lincoln. 1988. The Computer-stored Medical Record. For Whom? *Journal of the American Medical Association* 259: 3454-6.
Kosecoff, J. et al. 1987. Effects of the National Institutes of Health Consensus Development Program on Physician Practice. *Journal of the American Medical Association* 258: 2708-13.
Kulikowski, C. A. 1977. Problems in the Design of Knowledge Bases for Medical Consultation. Paper presented at the *IEEE Proceedings of the First Annual Symposium on Computer Applications in Medical Care*,
Langlotz, C. P. 1989. The Feasibility of Axiomatically-based Expert Systems. *Computer Methods and Programs in Biomedicine* 30: 85-95.
Latour, B. 1986. Visualisation and Cognition: Thinking with Eyes and Hands. *Knowledge and Society* 6: 1-40.
Latour, B. 1987. *Science in Action*. Milton Keynes: Open University Press.
Latour, B. 1988. *The Pasteurization of France*. Cambridge: Harvard University Press.
Latour, B. 1993. *We Have Never Been Modern*. Cambridge: Harvard University Press.
Latour, B., and S. Woolgar. 1986. *Laboratory Life. The Construction of Scientific Facts*. Princeton, N.J.: Princeton University Press.
Lau, J. M., J. P. Kassirer, and S. G. Pauker. 1983. Decision Maker 3.0. Improved Decision Analysis by Personal Computer. *Medical Decision Making* 3: 39-43.
Lave, J. 1988. *Cognition in Practice*. Cambridge: Cambridge University Press.
Law, J. 1987. Technology and Heterogeneous Engineering: The Case of Portuguese Expansion. In *The Social Construction of Technological Systems*, edited by W. E. Bijker, T. P. Hughes, and T. J. Pinch, 111-134.
Law, J., ed. 1991. *A Sociology of Monsters: Essays on Power, Technology and Domination*. London: Routledge.
Law, J. 1993. *Organizing Modernity*. Oxford: Blackwell.
Lawrence, C. 1985. Moderns and Ancients: The 'New Cardiology' in Britain 1880-1930. *Medical History* Supplement No. 5: 1-33.
Leaper, D. J., J. C. Horrocks, J. R. Staniland, and F. T. de Dombal. 1972. Computer-Assisted Diagnosis of Abdominal Pain using "Estimates" Provided by Clinicians. *British Medical Journal* 4: 350-4.
Ledley, R. S. 1965. *Use of Computers in Biology and Medicine*. New York: McGraw-Hill.
Ledley, R. S., and L. B. Lusted. 1959. Reasoning Foundations of Medical Diagnosis. Symbolic Logic, Probability, and Value Theory Aid Our Understanding of How Physicians Reason. *Science* 130: 9-21.
Lipscombe, B. 1989. Expert Systems and Computer-controlled Decision Making in Medicine. *AI and Society* 3: 184-97.
Lipscombe, B. 1991. *Minds, Machines & Medicine: An Epistemological Study of Computer Diagnosis*. PhD Thesis, University of Bath, UK.
Lister, J. 1957. Time to Think - About Socialized Medicine. *New England Journal of Medicine* 256: 221-2.
Littenberg, B., and H. C. Sox Jr. 1988. Evaluating Individualized Medical Decision Analysis. *American Journal of Medicine* 84: 289-90.
Lock, M., and D. R. Gordon, ed. 1988. *Biomedicine Examined*. Dordrecht: Kluwer Academic Publishers.
LoGerfo, J. P. 1977. Variation in Surgical Rates: Fact vs. Fantasy. *New England Journal of Medicine* 297: 387-9.
Lomas, J. 1990. Promoting Clinical Policy Change: Using the Art to Promote the Science in Medicine. In *The Challenges of Medical Practice Variations*, edited by T. F. Andersen, and G. Mooney, 402-11.
Lomas, J. 1993. Making Clinical Policy Explicit. Legislative Policy Making and Lessons for Developing Practice Guidelines. *International Journal of Technology Assessment in Health Care* 9: 11-25.
Lomas, J. et al. 1988. The Role of Evidence in the Consensus Process. Results From a Canadian Consensus Exercise. *Journal of the American Medical Association* 259: 3001-5.
Lomas, J. et al. 1989. Do Practice Guidelines Guide Practice? The Effect of a Consensus Statement on the Practice of Physicians. *New England Journal of Medicine* 321: 1306-11.
Lomas, J. A. et al. 1991. Opinion Leaders vs Audit and Feedback to Implement Practice Guidelines. *Journal of the American Medical Association* 265: 2202-7.

Löwy, I. 1993. Decision-making in a Clinical Trial: The Introduction of Interleukin-2 Therapy into the Cancer Clinics. Paris: INSERM. Unpublished manuscript.
Löwy, I. 1994. Nothing More to Do: Palliative Care versus Experimental Therapy in Advanced Cancer. *Science in Context* forthcoming.
Luff, P., and C. Heath. 1993. System Use and Social Organisation. Observations on Human-computer Interaction in an Architectural Practice. In *Technology in Working Order*, edited by G. Button, 184-210.
Lundberg, G. D. 1981. Acting on Significant Laboratory Results. *Journal of the American Medical Association* 245: 1762-3.
Lundsgaarde, H. P. 1987. Evaluating Medical Expert Systems. *Social Science and Medicine* 24: 805-819.
Lusted, L. B. 1968. *Introduction to Medical Decision Making*. Springfield, IL: Thomas.
Lusted, L. B. 1971. Decision-making Studies in Patient Management. *New England Journal of Medicine* 284: 416-24.
Lusted, L. B. 1975. In the Process of Solution. *New England Journal of Medicine* 293: 255-6.
Lynch, M. 1984. 'Turning up Signs' in Neurobehavioral Diagnosis. *Symbolic Interaction* 7: 67-86.
Lynch, M. 1985. *Art and Artifact in Laboratory Science. A Study of Shop Work and Shop Talk in a Research Laboratory*. London: Routledge and Kegan Paul.
Lynch, M. 1991. Laboratory Space and the Technological Complex: An Investigation of Topical Contextures. *Science in Context* 4: 51-78.
Lynch, M. 1993. *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. New York: Cambridge University Press.
Lynch, M., E. Livingston, and H. Garfinkel. 1983. Temporal Order in Laboratory Work. In *Science Observed*, edited by K. D. Knorr-Cetina, and M. Mulkay, 205-238. London: Sage.
MacMahon, B. 1955. Statistical Methods in Medicine. *New England Journal of Medicine* 253: 646-52,688-93.
Margolis, C. Z. 1983. Uses of Clinical Algorithms. *Journal of the American Medical Association* 249: 627-32.
Marks, H. M. 1988. Notes from the Underground: The Social Organization of Therapeutic Research. In *Grand Rounds. 100 Years of Internal Medicine*, edited by R. C. Maulitz, and D. E. Long, 297-336. Philadelphia: University of Pennsylvania Press.
Marks, H. M. forthcoming. *A Rational Therapeutics*.
Markussen, R. 1994. Dilemmas in Cooperative Design. In *PDC'94: Proceedings of the Participatory Design Conference*, edited by R. Trigg, S. I. Anderson and E. A. Dijkstra-Erickson. Chapel Hill, NC.
Måseide, P. 1983. Analytical Aspects of Clinical Reasoning: A Discussion of Models for Medical Problem Solving. In *The Social Organization of Doctor-Patient Communication*, edited by S. Fisher, and A. D. Todd, 241-265. Washington, D.C.: Center for Applied Linguistics.
Maxmen, J. S. 1987. Long Term Trends in Health Care: The Post Physician Era Reconsidered. In *Indicators and Trends in Health and Health Care*, edited by D. Schwefel, 109-15. Berlin: Springer.
May, W. E. 1985. Consensus or Coercion. *Journal of the American Medical Association* 254: 1077.
Mayne, J. G., W. Weksel, and P. N. Sholtz. 1968. Toward Automating the Medical History. *Mayo Clinic Proceedings* 43: 1-25.
McDermott, W. 1971. Medicine in Modern Society. In *Cecil-Loeb Textbook of Medicine*, edited by P. B. Beeson, and W. McDermott, 1-3. Philadelphia: W B Saunders.
McDonald, C. J. 1976. Protocol-Based Computer Reminders, The Quality of Care and the Non-Perfectability of Man. *New England Journal of Medicine* 295: 1351-5.
McGehee Harvey, A. et al. 1984. *The Principles and Practice of Medicine*. New York: Appleton.
McNeil, N., S. G. Pauker, H. C. Sox, and A. Tversky. 1982. On the Elicitation of Preferences for Alternative Therapies. *New England Journal of Medicine* 306: 1259-62.
Meneely, G. R., O. Paul, H. F. Dorn, and T. R. Harrison. 1960. Cardiopulmonary Semantics. *Journal of the American Medical Association* 174: 1628-9.
Mesman, J. 1993. The Digitalization of Medical Practice. Paper presented at the *Annual Meeting of the Social Studies of Science*, West-Lafayette, IN, USA, October 1994.
Miller, P. L. 1986. The Evaluation of AI Systems in Medicine. *Computer Methods and Programs in Biomedicine* 22: 5-11.

Miller, R. A., and F. E. Masarie Jr. 1990. The Demise of the 'Greek Oracle' Model for Medical Diagnostic Systems. *Methods of Information in Medicine* 29: 1-2.
Miller, R. A., M. A. McNeil, S. M. Challinor, F. E. Masarie, and J. D. Myers. 1986. The INTERNIST-1/QUICK MEDICAL REFERENCE Project - Status Report. *Western Journal of Medicine* 145: 816-22.
Miller, R. A., H. E. Pople, and J. D. Myers. 1982. INTERNIST-I, An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine. *New England Journal of Medicine* 307: 468-476.
Mirowski, P. 1992. When Games Grow Deadly Serious: The Military Influence on the Evolution of Game Theory. In *Toward a History of Game Theory*, edited by E. R. Weintraub, 227-55. Durham: Duke University Press.
Mol, A. 1993. Decisions No One Decides About. Anemia in Practice. Paper presented at *Ethics in the Clinic. An international conference on normative and sociological aspects of clinical decision making*, Maastricht, March 1993.
Mol, A. forthcoming. Missing Links, Making Links. The Performance of Some Atheroscleroses. In *Differences in Medicine*, edited by M. Berg, and A. Mol.
Mol, A., and M. Berg. 1994. The Principles and Practices of Medicine: The Co-existence of Various Anemias. *Culture, Medicine and Psychiatry* 18: 247-65.
Mol, A., and J. Law. 1994. Regions, Networks and Fluids: Anaemia and Social Topology. *Social Studies of Science* 24: 641-71.
Morgan Jr., W. M. 1988. Clinical Approach to the Patient. In *Cecil Textbook of Medicine*, edited by J. B. Wijngaarden, and L. H. Smith, 70-4.
Moser, R. H. 1974. An Anti-intellectual Movement in Medicine? *Journal of the American Medical Association* 227: 432-4.
Moskowitz, A. J., B. J. Kuipers, and J. P. Kassirer. 1988. Dealing with Uncertainty, Risks, and Tradeoffs in Clinical Decisions. A Cognitive Science Approach. *Annals of Internal Medicine* 108: 435-49.
Mulkay, M. 1988. Don Quixote's Double: a Self-exemplifying Text. In *Knowledge and Reflexivity*, edited by S. Woolgar, 81-100.
Neal, M. P. 1951. Diagnostic Drifts, Deceptions and Common Misses. *Journal of the American Medical Association* 146: 537-41.
Neellon, F. A. 1974. POR Relations. *New England Journal of Medicine* 290: 854-5.
Newell, A., and H. A. Simon. 1972. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
Nyce, J. M., and W. Graves III. 1990. The Construction of Knowledge in Neurology: Implications for Hypermedia System Development. *Artificial Intelligence in Medicine* 2: 315-22.
O'Connell, J. 1993. Metrology: The Creation of Universality by the Circulation of Particulars. *Social Studies of Science* 23: 129-73.
Oliver, M. F. 1985. Consensus or Nonsensus Conferences on Coronary Heart Disease. *Lancet* May 11: 1087-9.
Oughterson, A. W. 1955. Surgery, Science and Society. *New England Journal of Medicine* 252: 463-7.
Overall, J. E. 1972. Empirical Approaches to Classification. In *Computer Diagnosis and Diagnostic Methods*, edited by J. A. Jacquez, 5-9.
Pasveer, B. 1992. *Shadows of Knowledge. Making a Representing Practice in Medicine: X-Ray Pictures and Pulmonary Tuberculosis, 1895-1930.* PhD Thesis, Universiteit van Amsterdam.
Patil, R. S., P. Szolovits, and W. B. Schwartz. 1982. Modeling Knowledge of the Patient in Acid-Base and Electrolyte Disorders. In *Artificial Intelligence in Medicine*, edited by P. Szolovits, 191-226.
Patterson, J. T. 1987. *The Dread Disease. Cancer and Modern American Culture*. Cambridge: Harvard University Press.
Pauker, S. G., G. A. Gorry, J. P. Kassirer, and W. B. Schwartz. 1976. Towards the Simulation of Clinical Cognition. Taking a Present Illness by Computer. *American Journal of Medicine* 60: 981-96.
Pearson, S. D. et al. 1992. The Clinical Algorithm Nosology: A Method for Comparing Algorithmic Guidelines. *Medical Decision Making* 12: 123-31.
Pellegrino, E. D. 1964. The Diagnostic Process - I. In *The Diagnostic Process*, edited by J. A. Jacquez, 143-49.
Perrolle, J. A. 1991. Expert Enhancement and Replacement in Computerized Mental Labor. *Science, Technology and Human Values* 16: 195-207.

Pickering, A., ed. 1992. *Science as Practice and Culture*. Chicago: University of Chicago Press.
Plante, D. A., J. P. Kassirer, D. A. Zarin, and S. G. Pauker. 1986. Clinical Decision Consultation Service. *American Journal of Medicine* 80: 1169-76.
Politser, P. 1981. Decision Analysis and Clinical Judgment. A Re-Evaluation. *Medical Decision Making* 1: 361-89.
Pople, H. E. 1982. Heuristic Methods for Imposing Structure on Ill-Structured Problems: The Structuring of Medical Diagnostics. In *Artificial Intelligence in Medicine*, edited by P. Szolovits, 1357-60.
Potthoff, P., M. Rothemund, D. Schwefel, R. Engelbrecht, and W. v. Eimeren. 1988. Expert Systems in Medicine: Possible Future Effects. *International Journal of Technology Assessment in Health Care* 4: 121-33.
Pylyshyn, Z. W. 1980. Computation and Cognition: Issues in the Foundations of Cognitive Science. *The Behavioral and Brain Sciences* 3: 111-169.
Raiffa, H. 1968. *Decision Analysis. Introductory Lectures on Choices under Uncertainty*. Reading, MA: Addison-Wesley.
Rees, C. 1981. Records and Hospital Routine. In *Medical Work: Realities and Routines*, edited by P. Atkinson, and C. Heath, 55-70. Farnborough: Gower.
Reggia, J. A., and S. Tuhrim, ed. 1985. *Computer-Assisted Medical Decision Making*. New York: Springer-Verlag.
Reiser, S. J. 1978. *Medicine and the Reign of Technology*. Cambridge: Cambridge University Press.
Rennie, D. 1981. Consensus Statements. *New England Journal of Medicine* 304: 665-6.
Reverby, S. 1981. Stealing the Golden Eggs: Ernest Amory Codman and the Science and Management of Medicine. *Bulletin of the History of Medicine* 55: 156-71.
Richards, E. 1991. *Vitamin C and Cancer: Medicine or Politics*? New York: St. Martin's Press.
Riesenberg, D. 1989. Economics is Everybody's Business. *Journal of the American Medical Association* 262: 2897.
Rizzo, J. A. 1993. Physician Uncertainty and the Art of Persuasion. *Social Science and Medicine* 37: 1451-9.
Robinson, M. 1991a. Computer Supported Co-operative Work: Cases and Concepts. Paper presented at *GROUPWARE '91*, Amsterdam, 29-30 October.
Robinson, M. 1991b. Double-Level Languages and Co-operative Working. *AI and Society* 5: 34-60.
Robinson, M. 1994. Intimacy & Abstraction & Maps & Terrains. Paper presented at conference *Locating Design, Development and Use*, Øksnoen, May 13-18.
Rosenberg, C. E. 1987. *The Care of Strangers. The Rise of America's Hospital System*. New York: Basic Books.
Rothman, D. J. 1991. *Strangers at the Bedside: A History of How Law and Bioethics Transformed Medical Decision Making*. New York: Basic Books.
Ruby, A., and J. Morganroth. 1970. The Need for a Medical Ideology. *Journal of the American Medical Association* 212: 2096-7.
Rutstein, D. 1962. Another Look at the Patient. *New England Journal of Medicine* 267: 338-42.
Rutstein, D. D. 1961. Maintaining Contact with Medical Knowledge. *New England Journal of Medicine* 265: 321-4.
Saetnan, A. R. 1991. Rigid Politics and Technological Flexibility: The Anatomy of a Failed Hospital Innovation. *Science, Technology and Human Values* 16: 419-447.
Savage, L. J. 1972a. Diagnosis and the Bayesian Viewpoint. In *Computer Diagnosis and Diagnostic Methods*, edited by J. A. Jacquez, 131-8.
Savage, L. J. 1972b. *The Foundations of Statistics*. New York: Dover.
Schaffner, K. F., ed. 1985. *Logic of Discovery and Diagnosis in Medicine*. Berkeley: University of California Press.
Schoemaker, P. J. H. 1982. The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations. *Journal of Economic Literature* 20: 529-63.
Schwartz, E. B. 1970. Medicine and the Computer: The Promise and Problems of Change. *New England Journal of Medicine* 283: 1257-64.
Schwartz, S. I. 1987. Tempering the Technological Diagnosis of Appendicitis. *New England Journal of Medicine* 317: 703-4.
Schwartz, W. B. 1979. Decision Analysis: A Look at the Chief Complaints. *New England Journal of Medicine* 300: 556-9.

Schwartz, W. B., G. A. Gorry, J. P. Kassirer, and A. Essig. 1973. Decision Analysis and Clinical Judgment. *American Journal of Medicine* 55: 459-72.
Schwartz, W. B., R. S. Patil, and P. Szolovits. 1987. Artificial Intelligence in Medicine. Where Do We Stand? *New England Journal of Medicine* 316: 685-8.
Searle, J. 1984. *Minds, Brains and Science*. Cambridge, MA: Harvard University Press.
Shapin, S., and S. Schaffer. 1985. *Leviathan and the Air-pump. Hobbes, Boyle, and the Experimental Life*. Princeton: Princeton University Press.
Shortliffe, E. H. 1987. Computer Programs to Support Clinical Decision Making. *Journal of the American Medical Association* 258: 61-6.
Shortliffe, E. H. 1991. Medical Informatics and Clinical Decision Making: The Science and the Pragmatics. *Medical Decision Making* 11: 2-14(Suppl).
Shortliffe, E. H., S. G. Axline, B. G. Buchanan, T. C. Merigan, and S. N. Cohen. 1973. An Artificial Intelligence Program to Advise Physicians Regarding Antimicrobial Therapy. *Computers and Biomedical Research* 6: 544-60.
Shortliffe, E. H., B. G. Buchanan, and E. A. Feigenbaum. 1979. Knowledge Engineering for Medical Decision Making: A Review of Computer-Based Clinical Decision Aids. *Proceedings of the IEEE* 67: 1207-24.
Shortliffe, E. H. et al. 1984. An Expert System for Oncology Protocol Management. In *Rule-Based Expert Systems*, edited by B. G. Buchanan, and E. H. Shortliffe, 653-665.
Silverman, D. 1987. *Communication in Medical Practice*. London: Sage.
Silverstein, M. D. 1988. Prediction Instruments and Clinical Judgment in Critical Care. *Journal of the American Medical Association* 260: 1758-9.
Simon, H. A. 1979. From Substantive to Procedural Rationality. In *Philosophy and Economic Theory*, edited by F. Hahn, and M. Hollis, 159-62. Oxford: Oxford University Press.
Simon, H. A. 1981. *The Sciences of the Artificial*. Cambridge: MIT Press.
Simon, H. A. 1986. Alternative Visions of Rationality. In *Judgment and Decision Making*, edited by H. R. Arkes, and K. R. Hammond, 97-113. Cambridge: Cambridge University Press.
Singleton, V. forthcoming. Stabilizing Instabilities: The Laboratory in the UK Cervical Screening Program. In *Differences in Medicine*, edited by M. Berg, and A. Mol.
Slack, W. V. 1972. Patient Power: a Patient-Oriented Value System. In *Computer Diagnosis and Diagnostic Methods*, edited by J. A. Jacquez, 5-9.
Smith, L. H. 1988. Medicine as an Art. In *Cecil Textbook of Medicine*, edited by J. B. Wyngaarden, and L. H. Smith, 1-5.
Society for Medical Decision Making Committee on Standardization of Clinical Algorithms. 1992. Proposal for Clinical Algorithm Standards. *Medical Decision Making* 12: 149-54.
Sokal, R. R. 1964. Numerical Taxonomy and Disease Classification. In *The Diagnostic Process.*, edited by J. A. Jacquez, 143-49.
Southgate, M. T. 1975. Demythologizing Medicine; Or, How to Stop Worrying and Love the Computer. *Journal of the American Medical Association* 232: 515-6.
Sox Jr., H. C., C. H. Sox, and R. K. Tompkins. 1973. The Training of Physician's Assistants. The Use of a Clinical Algorithm System for Patient Care, Audit of Performance and Education. *New England Journal of Medicine* 288: 818-24.
Star, S. L. 1989a. Layered Space, Formal Representations and Long-distance Control: The Politics of Information. *Fundamenta Scientiae* 10: 125-154.
Star, S. L. 1989b. The Structure of Ill-Structured Solutions: Boundary Objects and Heterogeneous Distributed Problem Solving. In *Distributed Artificial Intelligence Vol. 2*, edited by M. Huhns, and L. Gasser, 37-54. London: Pitman.
Star, S. L. 1991. Power, Technologies and the Phenomenology of Conventions: On Being Allergic to Onions. In *A Sociology of Monsters*, edited by J. Law, 26-56.
Star, S. L. 1992. Craft vs. Commodity, Mess vs. Transcendence: How the Right Tool Became the Wrong One in the Case of Taxidermy and Natural History. In *The Right Tools for the Job*, edited by A. E. Clarke, and J. H. Fujimura, 257-286.
Star, S. L. 1994. Misplaced Concretism and Concrete Situations: Feminism, Method and Information Technology. Unpublished manuscript, University of Illinois, Champaign, IL, USA.
Star, S. L., and J. R. Griesemer. 1989. Institutional Ecology, 'Translations', and Boundary Objects: Amateurs and Professionals in Berkely's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science* 19: 387-420.
Starr, P. 1982. *The Social Transformation of American Medicine*. New York: Basic Books.

Stevens, R. 1989. *In Sickness and in Wealth: American Hospitals in the Twentieth Century*. New York: Basic Books.
Stoop, A., and M. Berg. 1994. Spiegels voor de arts. Beslissingsondersteunende technieken en medische praktijken. *Gezondheid: Theorie in Praktijk* 2: 265-78.
Strasser, P. H., J. C. Levy, G. A. Lamb, and J. Rosekrans. 1979. Controlled Clinical Trial of Pediatric Telephone Protocols. *Pediatrics* 64: 553-7.
Strauss, A., S. Fagerhaugh, B. Suczek, and C. Wieder. 1985. *Social Organization of Medical Work*. Chicago: University of Chicago Press.
Suchman, L. 1987. *Plans and Situated Actions. The Problem of Human-machine Communication*. Cambridge: Cambridge University Press.
Suchman, L. 1993a. Centers of Coordination. A Case and Some Themes. Paper presented at *NATO Advanced Research Workshop on Discourse, Tools and Reasoning*, Lucca, Italy, November 2-7.
Suchman, L. 1993b. Technologies of Accountability. Of Lizards and Aeroplanes. In *Technology in Working Order*, edited by G. Button, 113-26.
Suchman, L. 1994. Working relations of technology production and use. *Computer Supported Cooperative Work* 2: 21-40.
Suchman, L., and R. H. Trigg. 1993. Artificial Intelligence as Craftwork. In *Understanding Practice. Perspectives on Activity and Context*, edited by S. Chaiklin, and J. Lave, 144-78. Cambridge: Cambridge University Press.
Sultan, C., M. Imbert, and P. G. 1988. Decision-making System (DMS) Applied to Hematology. Diagnosis of 180 Cases of Anemia Secondary to a Variety of Hematologic Disorders. *Hematologic Pathology* 2: 221-8.
Sutherland, J. W. 1986. Assessing the Artificial Intelligence Contribution to Decision Technology. *IEEE Transactions on Systems, Man and Cybernetics* SMC-16: 3-20.
Szolovits, P., ed. 1982. *Artificial Intelligence and Medicine*. Boulder, Colorado: Westview Press.
Szolovits, P., and W. J. Long. 1982. The Development of Clinical Expertise in the Computer. In *Artificial Intelligence and Medicine*, edited by P. Szolovits, 79-117.
Szolovits, P., R. S. Patil, and W. B. Schwartz. 1988. Artificial Intelligence in Medical Diagnosis. *Annals of Internal Medicine* 108: 80-87.
Szolovits, P., and S. G. Pauker. 1978. Categorical and Probabilistic Reasoning in Medical Diagnosis. *Artificial Intelligence* 11: 115-44.
Talbott, J. H. 1961. The Art of Medicine. *Journal of the American Medical Association* 175: 898-9.
Talmon, J. L., R. A. J. Schijven, P. J. E. H. M. Kitslaar, and R. Penders. 1987. An Expert System for Diagnosis and Therapy Planning in Patients with Peripheral Vascular Disease. In *Proceedings European Conference on Artificial Intelligence in Medicine*, edited by J. Fox, M. Fieschi, and R. Engelbrecht, 37-44. Berlin: Springer-Verlag.
Tielens, V. 1989. Standaarden: het gezicht van de huisarts. *Huisarts & Wetenschap* 32: 3.
Timmermans, S. 1993. Calling the Patient: Interpretative Information Work and the Closure of a Resuscitative Trajectory. Unpublished manuscript, University of Illinois, Campaign, IL.
Timmermans, S., and M. Berg. forthcoming. Standardization in Action: Achieving Universalism and Localization in Medical Protocols.
Timmermans, S., G. C. Bowker, and S. L. Star. forthcoming. The Architecture of Difference: Visibility, Controllability, and Comparability in Building a Nursing Intervention Classification. In *Differences in Medicine*, edited by M. Berg, and A. Mol.
Tompkins, R. K., W. D. Kniffin Jr., H. C. Soc Jr., C. H. Sox, and A. D. Kaplan. 1973. Use of a Clinical Algorithm System in a Physician's Assistant Program. In *Applying the Problem-Oriented System*, edited by H. K. Walker, J. W. Hurst, and M. F. Woody, 5-9. New York: Medcom Press.
Turkle, S. 1984. *The Second Self: The Human Spirit in a Computer Culture*. New York: Simon & Schuster.
Tversky, A., and D. Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* 185: 1124-31.
Vaisrub, S. 1971. Etched in Twilight. *Journal of the American Medical Association* 215: 1318.
van Bemmel, J. H., and J. L. Willems, ed. 1989. *Handboek Medische Informatica*. Utrecht: Bohn, Scheltema en Holkema.
van Dijk, P. H., and J. Bomhof. 1991. Consensusontwikkeling: een basisbehoefte? *Medisch Contact* 46: 345-8.

Van Way III, C. W., J. R. Murphy, E. L. Dunn, and S. C. Elerding. 1982. A Feasibility Study of Computer Aided Diagnosis in Appendicitis. *Surgery, Gynecology & Obstetrics* 155: 685-8.
Vandenbroucke, J. P. 1988. Kwantitatieve nieuwlichterij in de geneeskunde: een poging tot ordening. *Nederlands Tijdschrift voor de Geneeskunde* 132: 337-40.
Vang, J. 1988. The Consensus Development Conference and the European Experience. *International Journal of Technology Assessment in Health Care* 3: 65-76.
Vogel, M. J., and C. E. Rosenberg, ed. 1979. *The Therapeutic Revolution: Essays in the Social History of American Medicine.* University of Pennsylvania Press.
Wachtel, T., A. W. Moulton, J. Pezzullo, and M. Hamolsky. 1986. Inpatient Management Protocols to Reduce Health Care Costs. *Medical Decision Making* 6: 101-9.
Wagner, I. 1994. Hard Times. The Politics of Women's Work in Computerized Environments. Paper presented at the *5th IFIP Working Conference on Women, Work and Computerization*, Manchester, July 2-5.
Warndorff, D. K., C. J. M. Pouls, and J. A. Knottnerus. 1988. Medische besliskunde in nederland: een literatuuronderzoek. *Medisch Contact* 43: 1174-78.
Warner, H. R., A. F. Toronto, and L. G. Veasy. 1964. Experience with Baye's Theorem for Computer Diagnosis of Congenital Heart Disease. *Annals of the New York Academia of Science* 115: 558-67.
Warner, J. H. 1985. Science in Medicine. *Osiris* 1 (2nd series): 37-58.
Warner, J. H. 1986. *The Therapeutic Perspective. Medical Practice, Knowledge and Identity in America, 1820-1885.* Cambridge: Harvard University Press.
Weaver, R. R. 1991. *Computers and Medical Knowledge. The Diffusion of Decision Support Technology.* Boulder: Westview Press.
Webster, J. 1991. Advanced Manufacturing Technologies: Work Organisation and Social Relations Crystallised. In *A Sociology of Monsters*, edited by J. Law, 192-222.
Weed, L. L. 1968. Medical Records that Guide and Teach. *New England Journal of Medicine* 278: 593-600,652-7.
Weed, L. L. 1971. *Medical Records, Medical Education, and Patient Care.* Chicago: Year Book Medical Publishers.
Weinstein, M. C. et al. 1980. *Clinical Decision Analysis.* Philadelphia: Saunders.
Weizenbaum, J. 1976. *Computer Power and Human Reason. From Judgment to Calculation.* San Francisco: W. H. Freeman and Company.
Wennberg, J. 1984. Dealing with Medical Practice Variations: a Proposal for Action. *Health Affairs* 3: 6-32.
Wennberg, J. E. 1991. Unwanted Variations in the Rules of Practice. *Journal of the American Medical Association* 265: 1306-7.
Whalen, J. 1993. Accounting for 'Standard' Task Performance in the Execution of 9-1-1 Operations. Paper presented at the *Annual Meetings of the American Sociological Association*, Miami, August 1993.
White, I. 1988. W(h)ither Expert Systems? - A View from the Outside. *AI and Society* 2: 161-71.
Wiersma, T. 1992. Overwegingen bij de NHG-standaard cholesterol. *Huisarts & Wetenschap* 35: 97-100.
Willems, D. 1993. Namen voor een ziekte. *Gezondheid: Theorie in Praktijk* 1: 144-155.
Williamson, J. W. 1973. Evaluating the Quality of Medical Care. *New England Journal of Medicine* 288: 1352-3.
Wilson, D. H., P. D. Wilson, R. G. Walmsley, J. C. Horrocks, and F. T. de Dombal. 1977. Diagnosis of Acute Abdominal Pain in the Accident and Emergency Department. *British Journal of Surgery* 64: 250-4.
Winickoff, R. N., K. McCue, and F. Perlman. 1977. Upper Respiratory Infection. In *A Problem-Oriented Textbook with Protocols*, edited by A. L. Komaroff, and R. N. Winickoff, 11-40. Boston: Little, Brown and Company.
Winograd, T., and F. Flores. 1986. *Understanding Computers and Cognition. A New Foundation for Design.* Norwood, N.J.: Ablex.
Wittgenstein, L. 1958. *Philosophical Investigations.* Oxford: Basil Blackwell.
Wood, D. 1992. *The Power of Maps.* New York: Guilford Press.
Wooffit, R., and N. Fraser. 1993. We're Off to Ring the Wizard, the Wonderful Wizard of Oz. In *Technology in Working Order*, edited by G. Button, 211-230.
Woolgar, S. 1985. Why Not a Sociology of Machines? The Case of Sociology and Artificial Intelligence. *Sociology* 19: 557-572.

Woolgar, S. 1987. Reconstructing Man and Machine: A Note on Sociological Critiques of Cognitivism. In *The Social Construction of Technological Systems*, edited by W. E. Bijker, T. P. Hughes, and T. J. Pinch, 311-328.
Woolgar, S. 1988. *Science: the Very Idea*. Chicester: Ellis Horwood.
Woolgar, S., ed. 1988. *Knowledge and Reflexivity: New Frontiers*. London: Sage.
Woolgar, S. 1989. Representation, Cognition and Self: What Hope for An Integration of Psychology and Sociology? In *The Cognitive Turn*, edited by S. Fuller, 201-224. Dordrecht: Kluwer.
Woolgar, S. 1991. Configuring the User: The Case of Usability Trials. In *A Sociology of Monsters*, edited by J. Law, 57-99.
Wouters, P. 1992. Een nieuw teken aan de wand: de uitvinding van de Science Citation Index 1948-1964. *Kennis en Methode* 16: 313-36.
Wulff, H. R. 1981. *Rational Diagnosis and Treatment*. Oxford: Blackwell.
Wyatt, J. 1989. Lessons Learned from the Field Trial of ACORN, an Expert System to Advise on Chest Pain. In *Proceedings Medical Informatics 1989, Singapore*, edited by P. Manning, O. Zhineng, and B. Barber, North Holland: Elsevier.
Wyatt, J. 1991a. *A Method for Developing Medical Decision Aids Applied to ACORN, a Chest Pain Advisor*. University of Oxford.
Wyatt, J. 1991b. Use and Sources of Medical Knowledge. *Lancet* 338: 1368-73.
Wyatt, J., and P. Emerson. 1990. A Pragmatic Approach to Knowledge Engineering with Examples of Use in a Difficult Domain. In *Expert Systems: Human Issues*, edited by D. Berry, and A. Hart, 65-78. Cambridge, MA: MIT Press.
Wyatt, J., and D. Spiegelhalter. 1991. Evaluating Medical Expert Systems: What To Test, and How? In *Knowledge Based Systems in Medicine: Methods, Applications and Evaluation*. Proceedings of the Workshop *System Engineering in Medicine*, Maastricht, March 16-18, 1989, 578-81.
Wyngaarden, J. B. 1979. Biological Science and Medical Practice. In *Cecil Textbook of Medicine*, edited by P. B. Beeson, W. M. Dermott, and J. B. Wijngaarden, 6-8.
Wyngaarden, J. B., and L. H. Smith, ed. *Cecil Textbook of Medicine*. Philadelphia: W. B. Saunders.
Wynne, A. 1988. Accounting for Accounts of the Diagnosis of Multiple Sclerosis. In *Knowledge and Reflexivity*, edited by S. Woolgar, 101-122.
Wynne, B. 1988. Unruly Technology: Practical Rules, Impractical Discourses and Public Understanding. *Social Studies of Science* 18: 147-67.
Young, A. 1981a. The Creation of Medical Knowledge: Some Problems in Interpretation. *Social Science and Medicine* 15B: 379-86.
Young, A. 1981b. When Rational Men Fall Sick: An Inquiry into Some Assumptions Made by Medical Anthropologists. *Culture, Medicine and Psychiatry* 5: 317-335.
Yu, V. L. et al. 1979. Antimicrobial Selection by a Computer. *Journal of the American Medical Association* 242: 1279-82.
Zerubavel, E. 1979. *Patterns of Time in Hospital Life*. Chicago: University of Chicago Press.
Zimmerman, D. H. 1970. The Practicalities of Rule Use. In *Understanding Everyday Life*, edited by J. D. Douglas, 221-38. Chicago: Aldine Publishing Co.
Zuboff, S. 1988. *In the Age of the Smart Machine. The Future of Work and Power*. New York: Basic Books.
Zubrod, C. 1984. Origins and Development of Chemotherapy Research at the National Cancer Institute. *Cancer Treatment Reports* 68: 9-19.
Zwaard, A., S. Zijstra, and R. Grol. 1989. Kwaliteits- en deskundigheidsbevordering rond NHG standaarden. *Huisarts & Wetenschap* 32: 501-4.

Curriculum Vitae

Marc Berg was born March 19, 1964, in Leiden, the Netherlands. After finishing high school in 1982 he spent a year in Greencastle, IL, USA (DePauw University), attending a liberal arts program. From 1983 - 1990 he studied medicine, and from 1985 - 1990 he studied health sciences at the University of Limburg, Maastricht. In 1990 he worked four months as a research assistent for Dr. A. Mol at the Department of Philosophy of the Faculty of General Sciences. In 1991, still at the University of Limburg, he became an AIO (Assistent in Opleiding; a PhD student) at the Department of Health Ethics and Philosophy of the Faculty of Health Sciences. In 1994 he became a temporary 'universitair docent' (a lecturer) at the latter department, where he is now a post-doc researcher.