BEHIND THE SCENES OF PERFORMANCE

Performance, practice and management in medical research[1]

Ad A.M. Prins
Department of Science Dynamics Amsterdam
Nieuwe Achtergracht 166
1018 WV Amsterdam
The Netherlands

Summary

In this paper the research performance and management of medical research groups will be evaluated taking cognizance of cognitive and local aspects of the specialties involved. A methodology for the interpretation of performance indicators is offered, using dimensions of the degree of uncertainty over information, organizational dependency in research and patient care, and the task conceptions of actors. It concludes: 1. that formal indicators can be interpreted in terms of cognitive and local aspects of research groups; 2. that task conceptions concerning cognitive and organizational aspects of patient care and research are crucial for the interpretation of performance; and 3. that these task conceptions can be conceived of as collective styles, which form a subject for research management and research policy. Therefore, cognitive and local aspects should be studied to make a valid and policy-relevant contribution.

1. Introduction

In most current studies of research performance, the sounds of "laboratory life" seem far away. Performance is usually measured in terms of indicators such as the number and impact of papers, and additionally by peer review. However useful these studies may be to research policy, their acceptance in the relevant fields is sometimes less than one might expect. Doubts about their validity are sometimes expressed both by the scientists, who are perhaps tempted to point to local circumstances, and by sociologists of science, who sometimes question the cognitive

[1] Acknowledgements: This paper is based on a policy study for a university hospital in The Netherlands, financed by the medical faculty of this university. Thanks are due to many (anonymous) physicians who were willing to tell us in detail about their work and their working conditions. Thanks to the persistent work of Albert Boekhorst, the necessary data could be gathered and processed in the limited period of time that is allowed for policy studies. The author would also like to thank Stuart Blume, for his stimulating comments on earlier versions of this paper, and Gene Moore for editing the paper.

significance of the indicators. Moreover, studies based on indicators leave room for questions about the way in which these indicators are to be interpreted or evaluated for policy measures. Besides this "gap" between indicators and their interpretation, the results of performance studies are particularly problematic in the applied sciences, since utility is in general not taken into account. This paper raises the question of whether performance indicators can be understood in terms of the cognitive nature of tasks and the local circumstances of organization and management in research groups in medical science.

2. The methodology of performance studies in the applied sciences

Despite recognized difficulties,[8] performance studies are sometimes carried out in fields of applied research. One example is the evaluation of health research in The Netherlands. In fact, this evaluation consisted of a number of studies concerning the costs as well as the productivity and quality of health research.[12] As the Dutch study indicates by its combined evaluation of both costs and performance, the increased interest in performance studies is largely due to the wish to optimize the allocation of declining research funds and to a concern for the quality of scientific research.

These kinds of assessments, often denominated as "external",[21] are of a nature different from that of more traditional "internal" assessments such as peer review. While peer review is traditionally accepted among scientists, the legitimation of external assessments has to be established, which is in most cases done through the use of peer review or a combination of peer review and formal techniques. Policy makers seem to assume that formal techniques provide non-experts with their own source of insight into the quality of research, and that the formal status of the methods adds to the acceptability of the results. Not surprisingly, the development of external methods for evaluation and assessments has until now been concentrated on the use of formal techniques, such as citation data and other bibliometric indicators, and several technology indicators. That bibliometric indicators provide at best a partial measure of the quality of research performance, is generally accepted.[9] Therefore, several methodologies have been developed, usually relying on the combined interpretation of these indicators and the evaluations of peers.[8] Even so, the evaluative use of these indicators is considered problematic by policy researchers. As Martin and Irvine state, bibliometric indicators cannot measure quality since these indicators are influenced by a vast array of "other factors". Martin and Irvine consequently propose a methodology which tries to control for the influences of other factors, by comparing groups of "like with like" and through their use of "convergent indicators". Others, like Moed, propose to use these indicators in a "pre-evaluative" manner and to rely partially on

the opinions of peers and the scientists involved in these studies in order to avoid rash interpretation.[10]

A further problem in performance studies is their application to applied sciences. Applied sciences, including technological and medical sciences, have distinctive social contexts in which research results are evaluated and to which researchers direct their activities.[2] Evaluation not only by peers but by users and the public also, is evident and potentially directs performance. It is therefore unlikely that performance studies in the applied sciences will produce relevant and valid results if their methodology is borrowed from performance studies in the basic sciences.

The perspective of this paper attempts to overcome these difficulties in relation to research in clinical medicine. Its central concerns are: first, the interpretation of performance indicators in terms of cognitive and local aspects of research; and secondly the policy and managerial issues that follow. Three dimensions are used: 1. the cognitive nature of the uncertainties encountered in the organization of research; 2. the organizational characteristics of patient care; and 3. the task conceptions of research and its applied context, as they become visible in local styles of research.

How can we conceptualize the cognitive nature of a research area? Whitley offers a comparative theory of the organizational characteristics of different disciplines including social as well as cognitive characteristics, which helps us to understand the nature of the relevant differences.[22]

According to Whitley, the main characteristic of science is the search for reputation. Researchers can only maintain and enhance their position in their field by gaining reputation on the basis of accepted knowledge claims. As a consequence, researchers are socially and cognitively dependent on other researchers, and on previously accepted results and methods of research. The dependencies can be found in different dimensions: in shared means for reducing task uncertainty (such as the use of methods and instruments to obtain information as unambiguous as possible), in functional dependency (concerning the use of the research results of colleagues), and in strategical dependency (concerning research agendas, institutional issues, and the public issues to be dealt with).

3. Dependencies and uncertainties in applied sciences: The case of medicine

[2] The applied sciences are in this respect an interesting field for science studies in general: the context of utility is positively identifiable, in contrast to a general and unspecific "social context".

In the exploration of medicine we have to adopt a more comprehensive view of organizational characteristics. Researchers have mixed roles in patient care and research, and have to satisfy the demands of patient care. Furthermore, various forms of research are differently related to the goals of patient care.[4] Therefore, we have to elaborate a framework that fits research and patient care as well.

Clinicians will not only seek recognition from colleagues in their scientific field, but will also strive to maintain professional standards in their patient care. The nature of the collegial control exerted over these two aspects of their activities might be quite different.[5] Professional control over work is based on claims that physicians have an "exclusive capacity to evaluate and solve life-threatening emergencies".[3] In contrast, control in research is based on the evaluation of novelties and knowledge claims. While reputational and professional control have some degree of autonomy in common, the autonomy of professional control is typically seen as localized in departments and hospitals.[16] This local control is most apparent in task conceptions influencing effectively the arrangements of work in patient care.[14] Given the different nature of the two forms of control and the mixed role of physicians/researchers, some ambiguity in the effectiveness of these forms of control is to be expected. Yet clinicians will maintain a coherent *style* in which they seek to overcome the uncertainties stemming from this ambiguity and which reflects their own sense of tasks in research and patient care.

Another aspect of professional control concerns the way in which colleagues are involved in controlling and managing their work. Specialized professionals maintain functional relations in the accomplishment of their tasks. For physicians this is particularly true in the use of various techniques and the expertise necessary for diagnosis and treatment. The techniques and expertise that are available to perform diagnosis and treatment lead to a division of tasks that can (in most cases) be performed and managed individually. Such task division is complemented by processes involving the exchanges of information and the coordination necessary to perform each task according to the goals and standards of the department. Uncertainty of information sets limitations to the exchange of information and the degree to which this can be done in an unambiguous way, thus affecting the possibilities of task division and control.

There may be various patterns of task division and exchange of information.[19] Firstly, when actors use the same kind of information to perform similar tasks leading to the same kind of product, we may speak of *pooled interdependence*. Secondly, in organizations with sequential or

[3] Freidson, op. cit., p. 119.

serial interdependence, actors work on the same product using different kind of information, without exchange of information. A third type is the *reciprocal organization*, in which the product is made by a reciprocal exchange of information. A fourth type can be added: the cluster form, or *professionalized organization*. Different tasks are performed by actors who do not exchange information for production purposes, but will do so only to maintain an *a posteriori* line of control.[5]

Local professional task conceptions, as well as these patterns of organization, are issues in the management of groups. Managerial activities such as the maintainance of certain task divisions will be determined both by local task conceptions, since all professionals within a group will to some extent have control over their own task,[4] and by the availability and use of certain instruments and standards of methods and specializations in each field. The influence of the (collective) task conception on managerial activities may therefore differ, according to specific task divisions. In departments with professional task divisions, the control of individual professionals over their own work will be substantial, leaving few possibilities for managerial control. In highly interdependent departments with a reciprocal tasks division, the integration of tasks requires a stable collective task conception within the group, in which management plays a stabilizing and sanctioning role. In such departments, management plays a hierarchical role in the maintainance of the task division and the dominant collective task conception.
In summary, we shall analyse the performance of clinical groups on three dimensions. The first of these involves *the organizational and cognitive characteristics of research*, in terms of (a) the uncertainties in research and (b) the (organizational) interdependencies contained in the actions to reduce these uncertainties. Secondly we shall compare this with *the organizational setting of researchers in their role of physicians*. In addition to these structural aspects of the accomplishment of research, which can be analyzed without concern for the motivations of actors, there is strong suggestion that individual and collective conceptions about the accomplishment of various tasks should also be a part of our analysis. Our third dimension in the analysis concerns conceptions about the kind of decisions to be taken in research and patient care. We shall analyse these conceptions and the consonant actions to be taken in research and patient care, as *individual and collective styles*.

[4] In organizational literature, the issues of authority and control in relation to organizational characteristics is scarcely dealt with when professional organizations are at stake. See e.g.: W.R. Scott, (1981), who mentions only "collegiate power" (p.282). For our purpose it is sufficient here, to note some constraints upon collective opinions about tasks and policy stemming from organizational characteristics, and the managerial aspects of the development of both organizations and coherent collective opinions. The professionalized character of control in groups suggests also that the power exerted by the "dominant coalition" in such groups is rather of an endorsed than of an authoritive kind.

The management issue will be analysed here only briefly, we shall focus on two aspects of management: namely, the relation of the collective styles to the organizational characteristics of the groups, and the management of styles.

4. Behind the scenes of performance

In recent Dutch studies on the performance of medical research, the issues of the cognitive nature of various disciplines and the peculiar influence of patient care are scarcely addressed. The output study of the Dutch Advisory Council for Science Policy (RAWB), is based on peer review and bibliometric indicators.[11] The RAWB study sought to evaluate clinical research groups from various disciplinary backgrounds. Each research group was classified by its specialty. For each specialty a restricted list of journals was composed, the impact score of which was compared with the citation rates of each group. This enabled the evaluators to identify groups "above average". Furthermore, both national and international peers were asked to mention outstanding or unique research groups.[12] The study made no distinction between clinical and other kinds of research, which provoked some criticism of the validity of the evaluation on the part of clinical groups.

The evaluation of the Royal Dutch Academy of the Sciences (KNAW), undertaken in 1985, and finally published in 1988, was based on the same formal indicators, used in the RAWB report.[7] Their own peer review of the groups nevertheless also indicated more groups performing above average.

These evaluations led to fierce public debate after policy measures were announced including the closing down of one academic hospital. Some commentators suggested that the tasks of clinical staff in education were insufficiently dealt with; others that the performance of clinical groups in research was negatively influenced by the very nature of their clinical tasks.[2],[20]

The study presented here takes a different approach to the evaluations as its starting point: How can performance in clinical research - as so assessed - be explained? Four groups in a university hospital were studied. Two groups were assessed positively (Cardiology and Neurology), while the two other groups seemed be performing only moderately (Diagnostic Radiology and Obstetrics/Gynaecology).

5. The organizational setting of patient care

For our investigation we will first explore the organizational setting of researchers in their roles of physicians, before we address the issues linked with the uncertainties and interdependencies

in research.
In the four groups we studied, we found patterns in the exchange of information and task divisions that display some of the characteristics mentioned in the typology above.

Cardiology.

In Cardiology there are several working units, each with different, specialized tasks concerning diagnosis and treatment. These units, including the polyclinical unit, the diagnosis unit, the care unit, intensive care and the catheterization-room, exhibit features of serial interdependence: information is passed on from the diagnosis unit to the care unit and from the care unit to intensive care, and from intensive care to the catheterization room, and vice versa. The unit of diagnostics and treatment is central in this exchange of information. The diagnosis and treatment of patients depends on the combination of the work of these various cardiological units, which is effected by the reciprocal exchange of the information produced by each of the units.

Neurology.

The Neurological Department shows a rather different organizational structure. In addition to a general clinic, which combines treatment and diagnosis for a variety of diseases, it also includes several specialized units. We can distinguish for instance the clinic for muscular diseases, the unit for children and the recently founded unit for dementia. Also, there is a neurobiochemical laboratory, an EEG/EMG lab and an experimental unit. The special units require a high degree of specialized knowledge, which only can be passed on to colleagues trained in this specialty. The exchange of information between the units for treatment and the various labs is limited and specialized. The labs provide the treatment units only with information on the basis of which the neurologist must decide what the diagnosis is. Such diagnosis is reported to be difficult given the variation in neurological diseases and the degree of specificity with which each can be diagnosed. The exchange of information between the treatment units occurs from the general clinic to the units, while there is hardly any exchange among the specialized units. Most neurologists here specialise in one of the special departments. They also perform tasks in the general clinic, where they pool their tasks. Aspects of the work within the Neurological and Cardiological Departments will be discussed later in more detail.

Obstetrics and Gynaecology.

Although Obstetrics and Gynaecology belong to the same department, we must distinguish both subdepartments. In Obstetrics, the division of tasks is limited by the time-consuming character of the obstetrical work. The initiation of deliveries is not stimulated but naturally expected. Therefore, there is a continuous flow of deliveries. The division of tasks of the obstetricians is therefore characterized by shift work: day and night obstetricians are involved in the medical work. Technical influences on the division of work, as in Cardiology, have some impact in Obstetrics: some of the obstetricians are specialised in diagnostic echoscopy, within a separate unit. (Another specialized unit is the small subdepartment of Neonatology.)

Gynaecology, the other main activity of the Obstetrics/Gynaecology department, has a very different task structure. For instance; in gynaecological oncology the task structure is mainly determined by the surgery that is necessary for treatment. After careful diagnosis, oncological treatment consists in most cases of a combination of chemotherapy and surgery after careful diagnosis. Diagnostical difficulties includes the examination to determine the kind of cancer, and the stage to which the tumour has developed. Oncological surgery requires years of training, and the practising of the routines and skills involved in the surgical protocol.

The gynaecologists of this department manage these uncertainties by pooling tasks and by assisting each other in the complex and time-consuming operations. There is a small pool of surgeons (a team) who practice special methods of operation. A wider circle of physicians is involved in more commonly known surgical matters and in diagnosis. One physician is specialized in *in vitro* fertilization.

Diagnostic Radiology.

From the point of view of patient care, a Radiodiagnostical Department can be considered as a service department: radiologists do not have a bedside practice, but lend their expertise and instruments to (other) physicians for treatment and diagnostical activities.

Tasks, and the uncertainties that go with them, are therefore generated by the contexts of use, brought in by these other physicians. In order to perform their tasks well, radiologists have to work in close cooperation with other specialists. Nevertheless, they have also to perform according to the standards of Diagnostic Radiology. The demands of other physicians have therefore to be "translated" and matched with radiological standards. In cases of complex instrumentation (such as the CT scanner) one radiologist is specialised to work solely with this instrument. Most tasks are pooled, in several pools, including surgery, neurology and internal medicine. The servicing character of the tasks does not imply an exchange of information within

each pool, as in the Obstetrics department. In order to maintain radiological standards, a daily meeting is held to discuss "difficult" röntgen-pictures, which can in itself be considered as a "professional line of control".

Although there are obvious differences between the departments, most of them have a combination of pooling and professional characteristics in common, such as in Neurology, Gynaecology and Diagnostic Radiology. Cardiology and Obstetrics have a different type of organization. Table 1. summarizes these organizational characteristics.

~~~~~~~ABOUT HERE TABLE 1.~~~~~~~

6. The entwinement of research and care: the nature of uncertainty in research.

How is clinical research done, and what is its relation to patient care? Is for instance the patient load an explanation for lower output rates, so as often claimed?

~~~~~~~~~~~ABOUT HERE TABLE 2.~~~~~~~~

As Table 2. indicates, there is indeed a negative "partial coherence" between the average numbers of papers per staff member and the indicators for the load of tasks in patient care. If the task load in routines is relatively high, the research output seems to be low, as Diagnostic Radiology shows. Also, if the task load in numbers of patients is higher, the average research output seems to be lower. The Obstetrics/Gynaecology group shows this coincidence also in the reverse direction: the lowest number of patients coincides with the highest average research output.
Still, such coherence does not provide insight into how research is related to patient care. In recent literature it is sometimes assumed that typical clinical research methods such as cohort studies, trials and case studies are connected in various ways with patient care.[4] These studies

address clinical questions such as the diagnosis of a disease, or assessing the benefits and risks of certain treatments. Each method seems to have a specific relation with patient care. All of these studies are based on information from patients; however, the number of patients may vary from one (case studies) to hundreds or even thousands (cohorts, trials). Furthermore, there is considerable variation in the design of the research: in contrast to case and cohort studies, various forms of trials have protocols in which decision rules are formalized about the choice of patients and treatment before the gathering of data has begun. This literature explains little about the way methods are connected with patient care, or about the implications of this connection for research. For our purpose it is necessary to probe this relationship.

Most research papers of the four groups provide evidence that research is based on information that is also used in patient care. Thus, the selection of patients for research is mainly based on the same diagnostic process that underlies patient care. Let us consider two examples, one from the Cardiology Department, and another from the Department of Neurology.

The cardiological study aims at a method for selecting patients with a particular movement of the walls of the heart after an infarction. The selection of patients is based on identifiable video pictures of the echocardiograms with which the movements are made visible. Furthermore...

> *"(t)he diagnosis of acute myocardial infarction was based on a typical history with evolving electrocardiographic changes or a diagnostic rise and fall in the activities of cardiac isoenzyme (MB), of creatine kinase or both. A value of more than twice the upper limit of normal (4U/L) was considered to be abnormal"*

The selection of patients for purposes of research is in this case entirely based on records of patient history, lab information, and the video registration of echocardiograms, techniques that belong -in this hospital- to the standard procedures concerning heart infarctions. No extra information is collected; the selection is possible because of the registration of information and the certainty with which norms can be maintained in patient care as well as in research. The selection of patients that meet the criteria for the research sample is made *post hoc* (for example in terms of the upper limit of enzymes like MB and CK).

The selection of patients for the neurological study is rather different. The purpose of the study is the effect of a certain medicine on the clinical course of a type of brain haemorrhage. Here, as in the case of the cardiological article, the diagnostic techniques are used both for research and patient-care. However, the outcomes of the diagnostic procedures are reported to be less certain

than in the cardiological case. Because CT results are inconclusive, angiography (an unpleasant invasive technique) has to be used. We should also note that the use of diagnostic techniques is influenced by the methodology of the research: the decision to perform angiography seems to derive from the purpose of the study:

"A CT scan was obtained on admission, weekly, and after clinical deterioration. Four vessel angiography was performed if the patient's clinical condition permitted .... If the CT scan was negative and therefore doubt existed about whether an aneurysm had ruptured, angiography was carried out before randomization."

In this particular case, diagnostic problems also influence the inflow of patients to the hospital. The kind of haemorrhage studied is rare. The researchers have therefore informed local practitioners about their project, to secure the inflow of these patients to their hospital. The phenomena are nonetheless not very clear to general practitioners, which causes an abundant inflow of patients who have only a severe headache.
Diagnostical and taxonomical uncertainties seem to be characteristic of neurology: as a neuromuscular specialist in this group stated regarding a published study: "As long as we do not have biochemical evidence, there is no better classification. If you would ask me whether I am certain in my diagnosis that it is not another disease, I am not".

Similar uncertainties appear in other studies of the four groups. It appears from our examples that they are based on difficulties in obtaining precise information via diagnostic procedures. In this way the applied character of clinical medical research influences the opportunities for research. The number of patients, necessary to obtain a research sample that meets statistical requirements can also be a difficulty. In the Neurology group cooperation seemed to be necessary with other hospitals, since some of the diseases were so rare that the influx rate for one hospital would be too low. A second consequence of various uncertainties is the degree to which routinization can take place, both in research and in patient care. Apparently, the possibilities for research vary, not only per research project but also, and more structurally, per group.

7. The interdependencies in clinical research

Clinical research is in most instances based on information gained in patient care. Therefore, we might expect information flows and task divisions in patient care to be as crucial in clinical research. This indeed can be shown, even when we restrict the analysis to the coauthorships that

occur within each of the four groups.[5]

Coauthorships within the four groups reveal in two cases patterns of cooperation parallel to the organization of patient care.[6]

The distinct units for diagnosis and treatment in *Cardiology* (Fig 1.) reappear as patterns of coauthorships: colleagues working in the same unit tend to publish more frequently with each other than with their colleagues from other units. Nevertheless, the various units are connected with each other, as we expected on the basis of the serial and reciprocal character of patient care in this group.

~~~~~ABOUT HERE FIGURE 1~~~~~

Again, in *Neurology* (Fig. 2), the pattern of coauthorships in this group represents a peculiar characteristic of patient care organization. Several units, such as the general treatment unit, and two specialized units are visible, as are the various laboratories. Recently established specialized units such as the one for dementia are not (yet) visible. Coauthorships exist only within the specialized units and among researchers of these units and the laboratories. Coauthorships between these units and the general clinic do not take place.

~~~~~ABOUT HERE FIGURE 2~~~~~

The percentage of coauthored publications contained in the graphs of Diagnostic Radiology and Obstetrics/Gynaecology suggests that cooperation in these two groups takes place principally within the groups rather than both inside and outside the groups. Again, the patterns are less readily interpretable in terms of the structuring of patient care. In *Obstetrics/Gynaecology* (Fig. 3) there is a vague delineation of the Gynaecological Department. The specialized surgical team

[5] In all groups cooperation with other units occur in patient care. This cooperation and its significance for research is left out for the brevity of our analysis.

[6] In figures 1 to 4, coauthorships within the four groups are depicted. The figure of Cardiology (Fig. 1.) is based on coauthorships over three years; the other figures are based on coauthorships over a period of 8 years. We used a combination of two techniques: network-analysis and multi-dimensional scaling. In terms of network analysis, each point in the figures represents an author, and each line the incidence of one or more coauthorships in the past years between the authors connected with such a line. N-cliques, in: [17]. In multi-dimensional scaling the distribution of the authors in these pictures is a function of the number of coauthorships they share with neighbours. The more coauthorships researchers have in common, the more closely they are depicted. The position of each author is thus expressed in terms of relative distance from coauthors. The dimensions of the relative distances among authors is reduced to two dimensions via the algorithm of Minissa. The dimensions in the graphs have only a "geographical" meaning. [13].

apparently does not publish much as a team, nor does the subdepartment of neonatology appear to be a coherent research group. Furthermore, the unit for ultrasound echoscopy in Obstetrics does not seem to be a source of information for research for this group, However, the figure of coauthorships does reveal a characteristic comparable to that of the Cardiology group: the dominant position of the head of the department, with whom every researcher in this group seems to work.

~~~~~ABOUT HERE FIGURE 3~~~~~

Coauthorships of *Diagnostic Radiology* (Fig. 4) occur in only a few cases: more than half of the members of this department have not shared any coauthorship in the past eight years. Furthermore, while in other departments full professors/heads of departments tend to publish with several colleagues at least, in Diagnostic Radiology this does not occur.

~~~~~ABOUT HERE FIGURE 4~~~~~

Thus while Cardiology and Neurology show an organizational similarity between research and patient care, Diagnostic Radiology and Obstetrics/Gynaecology do not. This finding can hardly be explained by the characteristics of the specialties: e.g. although Diagnostic Radiology is a service department, it does not follow that cooperation within this group would be impossible. Although organizational similarity in some groups indicates some degree of effectiveness of the organization of research, and thus might serve as an explanation for success, it is certainly not a sufficient explanation. We do not yet understand why such effectiveness cannot be found in the other groups. We will therefore turn to our third dimension of analysis: the conception of tasks.

8. The conceptions of tasks and uncertainties in patient care and research.

Thus far we have dealt with the relation of patient care and research as if information were

passed on from patient care to research, and as if physicians switched from their role in patient care to their role as researchers without any problems. This is, however, an oversimplification. First of all, we will introduce two different perspectives related to patient care and research, after which we will discuss the preferences for each perspective in the four groups.

In interviews held with members of all departments, respondents frequently opposed the needs of individual patients to research on large numbers of patients during diagnostical activities:[7]

"One can view patients in two different ways. One can regard a patient as a human being in illness, or regard him as a member of a population with a disease." (a neurologist)

Diagnostics for research or for strictly individual patient care typically follow different "logics" as they imply a different basis for decisions and activities. Physicians may legitimately claim to know enough for treatment in the case of an individual patient without defining precisely all the aspects of this single case in terms of its relevant population. Thus, decisions to exclude uncertain or ambiguous information in diagnostics, can be different from those required in research; in the case of research, all aspects have to be dealt with via tests to prove that a given patient is a veritable member of the population under investigation. So, in large-scale studies, a population has to be composed via uniform and repeated diagnostical procedures of numbers of patients who match the demands of research. This will lead to quite different activities compared with those for patient care or for case studies:

"Take for instance amenorrhoea, a disease which prohibits menstruation in a patient. Now, one can diagnose and treat this "conservatively" (with hormones), or one can perform all kinds of tests in order to know precisely what is going on. In that case patients have to be hospitalized, and take lumbar punctures, cisterna punctures, kidney biopsies etcetera. Then you are able to publish for instance the microscopic image and have nice publications with a lot of parameters. The vast majority however is not relevant for treatment. Now imagine what the consequences are for patients! I think it leads to excesses in medicine." (a gynaecologist)

---

[7] We selected 26 phycisians actively performing research and publishing several articles in the last four years: the respondent group included the professor/head of the department of each group. In order to minimize biases of representation in opinions on patient care and research due to specific tasks in patient care or degree of involvement in research, two criteria of selection for the sample of other researchers were used. The first was an equal distribution of researchers over the various units of patient care. The second was an equal distribution of respondents in central and peripheral positions in the patterns of coauthorships in the figures above: researchers in the centre of a network might reveal other aspects and also other opinions on research and patient care than researchers in the periphery of that network [6].

The uncertainties in the diagnostics of amenorrhoea are according to this gynaecologist to be interpreted *first* in terms of patient care, and certainly *not as important* as the uncertainties of some research project.

There is considerable variation within the four groups with respect to the way diagnostics and its uncertainties are to be handled,[8] yet there is no reason to suppose that these variations are characteristic of the cognitive nature of the work of each group. Although we found dominant conceptions in each group on matters of tasks and interpretations of the meaning of diagnostical uncertainty, in each group we found also some deviant opinions.

In Cardiology the difference between individual diagnostics and diagnostics for research (large numbers of patients) is in most cases barely manifest. As a cardiologist in echocardiography states:

"A lot of patients in this clinic, well, that's research and patient care at the same time. Because we try to put each patient in some kind of research project. That's how it goes".

In other groups, especially in Diagnostic Radiology and in Obstetrics/Gynaecology, some of the respondents were more critical regarding the significance of research protocols for patient care.[9]

In the interviews, the respondents from Diagnostic Radiology and Obstetrics/Gynaecology substantiated this criticism for their own field. For the obstetricians, pregnant women are not primarily ill: they consider pregnancy not as abnormal or a disease. Secondly, members of this department favoured home delivery, which in the Netherlands is supported by an elaborate organization of midwives. In this way -as these respondents stated- hospitalization is minimized. Respondents are reluctant to rely unreflexively on tests which can only be interpreted in terms of statistical norms: "...normal values ..... become absolute criteria, which is absolutely ridiculous: there are vast numbers of people with abnormal values in their bodies without any

[8] Such different points of view can be understood in terms of an ongoing debate in the philosophy of medicine. See for instance for the naturalistic point of view [1], and for the normative point of view [3]. In this respect, the local form of task conceptions is not so much contained in these or similar points of view, but rather in the specific use of such points of view in a local context. This becomes clear from a comparison of the different task conceptions of Radiodiagnosticians and Obstetricians/Gynaecologists: Both are critical ("normative") though certainly not equal.

[9] Critical or normative opinions may concern for instance: 1. the reduction of complaints of patients, symptoms and other evidences of diseases to its biological content; 2. the hospitalization or iatrogenesis of complaints and symptoms (physicians are too readily prepared to intervene); 3. the abstraction of the individual history of a patient into a typical case of a (sick) population; and 4. the dichotomizing of the outcome of all sorts of diagnostical tests into "normal" vs. "abnormal" [18].

complaints", as one obstetrician stated. Some of the respondents characterised this point of view as markedly different from that of other Obstetrical groups in The Netherlands.
In Diagnostic Radiology the reluctance to subordinate diagnostical activities to research is related to the interests of the individual patient. In this department it is an explicit policy issue that the development of procedures in taking röntgen pictures, as well as the reading of the pictures, be done by radiologists themselves (and not by other physicians), in order to minimize the dose of radiation for each individual patient.[10] This brings in some potential conflict with physicians who want a particular group of patients screened in a similar manner: cooperation with other groups is therewith strained.

9. Collective styles in the various organizational settings

Respondents show a variety of preferences in the different ways diagnostical uncertainty can be defined. In general, when asked to range activities like patient care, education and research according to their importance, each respondent identified patient care as the top priority, the importance of research varying from "as important as patient care" to "second best". These preferences concern also the decisions to be made in uncertain situations for example the use of invasive forms of diagnosis for research purposes, and they have consequences for both patient care and research.
Such preferences, or styles can be identified not only at the level of individuals but also collectively, at the level of the group. One can speak of a dominant style in research, which is the dominant style shared by the majority of the group.

There is, of course, no imperative reason why a certain style should predominate in a certain group. But there are constraints stemming from the organizational demands of care and of research. Conversely, styles influence the development of the organization. Although each professional will have his or her individual conception of patient care and research, such conceptions are in most cases shared by colleagues in the group. Such conceptions do not stand on their own, but relate to the development of the organization of patient care, and to the sanctions used among colleagues. Ideas about how to develop research and patient care are shared, elaborated and put into practice by numbers of colleagues. The necessity of this will be clear in the instance of large-scale studies which involve a number of cooperating researchers. The maintenance of shared ideas will also lead to the sanctioning of deviant ideas and

[10] This is explicitly mentioned in the yearbooks of this department.

behaviour. The ways these sanctions are carried out are not the same for every group.

In Cardiology the serial/reciprocal character of patient care, and the predominant preference for large-scale studies implies that none of the physicians should "cop out" with the excuse of a different opinion on research. Both the organizational structure of patient care and the distribution of research tasks are maintained in a hierarchical manner, with almost "military precision". Senior members see themselves as "officers with soldiers (junior physicians)", while to younger researchers "toys (are) given" in order to accomplish narrowly defined research tasks. Such a hierarchical character is absent in Neurology: its pooled character does not require strict agreements, and the element of professional organization in its patient care would offer a way out if demands in styles became too strict. The collective style has collegiate features: the main event is the informal research meeting for all members, in which research initiatives are taken and publications and progress in projects are evaluated.
In Obstetrics, the same kind of organizational structure (pooling) involves a similar "collegiate" character of collective style. The predominant antipathy to hospitalization, which existed until recently, supported research projects based on laboratory studies of (for instance) placentas instead of information directly gained from hospitalized pregnant women. One of the respondents gave an example of the collective pressure in favour of this kind of research: a former member wanted to gain information from pregnant women. He thought it best to keep secret the tests he wanted to perform, in order to avoid disapproval in the group.

10. The influence of collective style on research

The collective styles, revealed in the interviews with the groups correspond with the kind of research carried out in each group. This holds both for journals and for the types of studies.
Table 3. shows a considerable variation in the journals addressed by the groups. Cardiology and Neurology direct themselves to international journals with a rather high status. Diagnostic Radiology also addresses international journals, but seems to address mostly journals with a lower status. The Obstetrics and Gynaecology groups seem to be oriented to Dutch journals. These findings correspond with respondent perceptions of what the publication policy of each group was. While respondents in the Cardiology and Neurology groups stated it was important for them to publish in the most important and prestigious journals, this was not the case for Diagnostic Radiology and Obstetrics/Gynaecology. The Obstetrics/Gynaecology group did orient themselves to a specific audience: the preference for home deliverance coincides with publications for (Dutch) midwives and G.P.'s in Dutch scientific papers and educational text-

books (Monographs).

~~~~~~~ABOUT HERE TABLE 3.~~~~~~~~

Also, the types of studies vary among the groups. We can distinguish several types of studies, which not only use different not research designs but also have different types of results:

1. The *large-scale clinical study*, based on large numbers of patients, and carried out by several authors. In this category we count cohort studies, and the various forms of trial studies. As we have seen in the examples of the Cardiological and Neurological studies, these studies require cooperation form various specialized parts of the groups. Furthermore, they require a uniform procedure of testing, diagnosis and treatment of all relevant patients. The methods of large-scale studies allow for experiments with drugs, treatment or medical devices.
2. *Case studies*, which are also clinical studies, are the product of observations of usually one or two authors. These studies are not viewed as research as some respondents claimed.[4] Observational studies such as case studies concern the detailed description of the course of a disease in one or several patients. The elaborated testing procedures for research can be restricted to those patients.
3. *Laboratory studies*, performed mainly on the basis of information gathered in a laboratory. In these studies the laboratory staff played an important role in most instances, though physicians also participate in laboratory work.
4. *Follow-up studies*, performed in cooperation with (or sometimes performed solely by) social scientists.

The distribution of these various types of studies within the groups is far from equal: respondents in Diagnostic Radiology, for instance, published far more case studies than other groups, while in Obstetrics educational publications and lab studies outnumbered the multiauthored clinical studies (Table 4).

~~~~~~~~~~ABOUT HERE TABLE 4.~~~~~~~~

These various types of studies are linked with the collective styles of the groups: In Diagnostic Radiology the conduct of case studies coincides with the dominant view that the patient has to come first. Large-scale studies are left to the initiative taken by physicians of other groups.[11] The predominant attitude against hospitalization in Obstetrics, which existed up until recently, supported publications for educational purposes (in Table 4. counted as Others). Note also the unique occurrence of follow-up studies here, performed to find out what the long-term social and physical consequences of certain kinds of treatment will be.

The preference for certain types of studies, an element of the dominant style in a group, is related to the organizational similarity between patient care and research noted above. It now becomes possible to see why such similarity occurs in Cardiology and Neurology, but not in Obstetrics and Diagnostic Radiology. Large-scale studies require the cooperation of colleagues, and the collection of a variety of information used in patient care. Hence functional dependencies in patient care will become visible in coauthor patterns. In case studies and in textbook publications, such cooperation is not necessary, and cannot be detected in our graphs. Moreover, prestigious journals are more likely to publish large scale studies than case studies.[4] The *types* of research carried out by Cardiology and Neurology thus explain their success in the RAWB study. The organizational similarities in our network graphs are apparently not only to be understood as "organizational" dependencies as such, but also as dependencies due to certain methods.[12]

Our conclusion must therefore be that collective styles influence not only our ideas about the most preferable types of studies, but are crucial in the explanation of the performance of medical clinical research groups. Furthermore, outcomes of indicator studies do not seem to be indifferent with respect to the types of research undertaken in the groups. Though other types of research may make important contributions in the long run, large scale studies may 'perform' better in analyses based on indicators in the short term.

---

[11] The case studies in neurology are also published by those who are interested in large scale studies. One of them stated: "Some of the phenomena are quite seldom or very hard to detect. I wouldn't say it is real research, it is interesting though". The higher degree of task uncertainties in neurology influences thus the kind of research: though the preferences are like those in the cardiology group, we will find case studies in neurology, while in cardiology case studies are absent.

[12] It should be noted here that methodologies change over time through the development of new methods or the refinement of existing ones, and perhaps also due to users or the public interest. Hence the structural equivalence in cooperation observed here cannot be used as an indicator of proper performance without an understanding of the methods that lead to such forms of cooperation.

11. The management of styles

As one might expect, some of the aspects of collective styles of research and patient care in these groups relate to issues of management. On the basis of two issues, personnel management and time management, it can be demonstrated how these management issues are entwined with collective style, and how these styles become a tradition due to a prolonged reinforcement of the collective style through management and sanctions. As we have seen, the collective style is enhanced and maintained by managerial decisions: collective styles are a subject for management.

In the views of all respondents, the possibilities of performing research of whatever sort were closely linked to the issue of time. Typical research activities, such as the analysis of findings and the writing of papers, had to be done after the accomplishment of tasks in patient care. In two departments, Cardiology and Neurology, arrangements exist to allow their members to perform research. In Neurology, the specific character of the task division, e.g. the pooling of manpower, allows each member three months off duty from his or her tasks in the general clinic in a period of one year. Colleagues would stand in, all staff members agreeing to work a little harder for the compensation of the loss of manpower. This arrangement is illustrative of the collegiate character of the collective style in this group.

In Cardiology, the serial character of patient care cannot allow for such a substitutional arrangement. Here a substitution is made through a temporary replacement by trainees in the same unit, whose performance is supervised by the exempted staff member. This kind of arrangement implies extra financial burdens to the group (in order to appoint an extra trainee) which is compensated by external grants for the research project that is being carried out.

In Diagnostic Radiology and Obstetrics/Gynaecology) such arrangements do not exist. The respondents of these groups who were actively involved in research complained about the lack of time and of any arrangements. In Diagnostic Radiology, patient care was considered so time-consuming that no time arrangements in favour of research could be made, even for those researchers who were active in research. They complained about the unwillingness - as they saw it- of the group management to make time for research. In Obstetrics, the head of department was until recently the only member of the group permanently involved in research (hence in each research project), while other members could only find some time incidentally. Alternatively, they did their research while they were working on their dissertations; that is, when most of them were working in the laboratory. After that, they seldom found time to do research. This explains the character of the coauthor network (Figure 3).

In Cardiology and in Obstetrics one could speak of a conscious and rather emphatic tradition in

research and patient care. In both groups the former heads of the departments were strong proponents of the predominant styles in their groups.[13] They typically appointed staff members under conditions which were consonant with their opinions about research and patient care. In Cardiology only those trainees were -and still are- appointed, who had demonstrated research involvement and who had written papers before they graduated. In Obstetrics the consistency of the group was thought to be enhanced by the appointment of a general practitioner (which is very unusual) in order to strengthen relations with the community organizations providing for home delivery.

Such personnel policy can also be used in order to found, to elaborate or even to break with a tradition. In Neurology research was until a few years ago the realm of but one man -the head of the department- with a preference for the in-depth case-studies that can be regularly found in neurology. The recent appointment of a new head of department led to the departure of those members (12) who were not interested in research, and the research was now directed to large-scale studies.

In Obstetrics, the appointment of new staff members has led to research projects with a more clinical character than before. This development can be valued as a veritable change in the existing tradition, since they were not trained in this department. Thus they were not influenced so much by the existing style as former trainees from this department, who had been trained by the former head.

12. Conclusions

The evaluation of performance in clinical medical science, at least in the Netherlands, has until now been based on the assumption that clinical groups can be evaluated in the same manner as basic research groups. Moreover, research groups within different clinical fields are typically evaluated with the same set of criteria. Although attempts have been made to account for differences among various specialties by comparing "like with like", no attention has been given to the cognitive nature of their work and its relation to the local circumstances of performance. However, problems arose with respect to the use of indicators as a criterion for evaluation in an applied context and also with the legitimate use of indicator studies for policy purposes. Partially this is due to the interpretative "gap" between the nature of performance indicators and the criteria for policy evaluation which have to consider both the cognitive aspects and the various

[13] Respondent views could be cross-checked by an analysis of the inaugural lectures of former and recent heads of departments.

goals of medical research.

In this paper a methodology is proposed for studying the cognitive and local characteristics of research in medical groups to explain their performance. First of all, the *cognitive and organizational characteristics of research* are studied in terms of (a) the uncertainties in research and (b) the organizational interdependencies designed to reduce these uncertainties. A second dimension concerned the *organizational setting of researchers in their medical professional roles*. Our third dimension embodied the *local professional task conception*, which we analysed as the individual and collective style. This dimension contained cognitive aspects (task uncertainties) as well as local conceptions of the way in which tasks in patient care and research are to be assessed and combined.

The four groups vary considerably in terms of the uncertainties to be encountered in research. These uncertainties stem largely from ambiguities in diagnostics and therapeutics in patient care. Such uncertainties occur not only in single research projects but affect in a more structural way the degree and the pattern of task divisions both in patient care and in research. We found that task divisions and the exchange of information characteristic of patient care prove important for the ways in which cooperation in research takes place.

Research performance cannot be explained *directly* by cognitive aspects such as uncertainty of information and by organizational patterns. For example, although markedly different in its pattern of task division and in its overall degree of uncertainty in information in comparison with Cardiology, the Neurology group was assessed as "above average". What proved to be most crucial was the perception of tasks by the clinician/researcher. The clinical context can be perceived as more important than research. This perception of tasks is not a matter only of personal opinion; there is an interaction with tasks and organizational structures, with the ways in which research and patient care are carried out. These perceptions are closely linked with the cognitive features of the specialities and relate to the kind of decisions to be taken in diagnostics and therapeutics.

Local styles articulate these perceptions of the ways in which patient care and research are to be performed both individually and collectively. As for research, cooperation in most current methods of clinical research is inevitable, yet it can be avoided or even blocked as a consequence of individual or collective style. The styles of the various groups explain the differences in research output; e.g. the preference for large-scale studies in the Cardiology and Neurology group is based on a dominant collective style. This kind of research is facilitated by time management and by the selection and management of personnel. The other two groups have different styles, leading to quite different forms of research output. Again, the way these collective styles or traditions are implemented is influenced to a great extent by the cognitive

nature of patient care and research, and by the organizational demands that go with it.

It appears that the possibilities for management and policy measures vary greatly, even though outcomes of indicator studies might be "the same": the possibilities for providing adequate time arrangements for research are not equal for each group. Each type of organization demands a specific arrangement. Furthermore, both the existing traditions and the compositions of the groups in terms of personnel will set constraints or even constitute drawbacks to attempts to management or research policy.

Lastly, indicator studies based on citation and impact rates seem to have a certain bias with respect to the types of studies undertaken: large scale studies seem to contribute disproportionately more to a better performance than follow-up studies or case studies. Policy measures based on indicator studies may encourage groups to perform a type of study that will lead to higher performance levels in the short term, for example large scale studies. However, other types of studies may in the long run also lead to important contributions. Thus, policy measures may unintentionally lead to a conflict between short term strategies and long term goals. Although performance studies may urge groups to do *more* research, they do not provide criteria for the *kind* of research that will lead to better results, or what the most preferable strategy for improvement would be.

This indicates that cognitive and local factors can and should be part of a valid and policy relevant analysis of performance in medical research, and conceivably in applied research in general. Although formal indicators are attractive for policy makers in assessing performance in the applied sciences, the neglect of cognitive influences and related informal aspects such as local professional styles will in the end lead to flaws in the analysis, and perhaps also to flaws in research policy.