

UNEXPECTED ADVANTAGES OF LESS ACCURATE PERFORMANCE MEASUREMENTS. HOW SIMPLE PRESCRIPTION DATA CONTROL THE USE OF DRUGS IN A COMPLEX SETTING

Accepted for publication in Public Administration.

Antoinette de Bont and Kor Grit

iBMG/ Erasmus University Rotterdam, The Netherlands

debont@bmg.eur.nl

Abstract

In this paper we argue that performance measurement can better be done by general, less accurate measurements than by complex – and possible more accurate - ones. The conclusions of this study are drawn from a case study of the Dutch Foundation for effective use of medication. While most studies about performance measurements focus on the management of public service organizations, this case study - informed by the literature from Science and Technology Studies – focuses on the active role of the measurements themselves. Indicators, we show, do not have to be as complex as the practices they represent, as long as they are part of a chain of intermediary data that allow traveling from the general or simple indicators to detailed data in day-to-day practices and vice versa. Furthermore, general indicators enable stakeholders to take distance from each other. Rather than the involvement of stakeholders, it is this reflexive distancing that explains the degree of compliance to performance measurement and thereby the prospect for effective co-governance.

UNEXPECTED ADVANTAGES OF LESS ACCURATE PERFORMANCE MEASUREMENTS. HOW SIMPLE PRESCRIPTION DATA CONTROL THE USE OF DRUGS IN A COMPLEX SETTING

INTRODUCTION

Performance measurement in the public sector is a contested area. The goal of measurement is unclear as indicators are linked to many functions, such as creating transparency, learning, appraisal and sanctioning (Radin 2006). Performance indicators give a poor representation of the complexity, variety and dynamics of public service production and can create more perverse than beneficial effects. Because it is hard to define performance of professional work, performance measurement systems tend to grow into complex systems (Power 1997; De Bruijn, 2002; Jos 2004; Dubnick, 2005; Radin 2006; Sanderson 2006). Improved representation is not automatically the best answer to these problems. There is more at stake than realizing an approximation of reality, such as competing values and professional accountability. Paradoxically, the limited representation of the complexities of public services seemed to be a strong point of the particular measurement (Klein and Carter 1988; Radin 2006; De Bruijn 2002). Many students of performance measurement have already contributed to the challenge of unraveling this enigma. Here are some answers:

- The majority of performance indicators function as tin-openers rather than dials (Klein and Carter 1988). “By opening up a ‘can of worms’ they do not give answers but prompt interrogation and inquiry, and by themselves provide an incomplete and inaccurate picture” (Carter 1994: 213). This creates the possibility of clarifying that a movement in a particular direction, such as towards lower unit costs or reduced length of patient stay, is desirable without setting a specific target or precise standard. Performance measurement provides no answers, but inspires actors to ask the right questions (Van Peursem et al 1995).
- Every single performance indicator should be interpreted as a poor representation of performance. If we acknowledge that a public performance can never be identified by one indicator, we need to allow more variety in performance measurement (Lindgren 2001; Radin 2006; De Bruijn 2002; Kanter and Summers 1987, Noordegraaf and Abma 2007). The poor quality of a single indicator is a drive to tolerate variety in the system of performance measurement or to develop a repertoire of performance measures. Variety and redundancy have a number of advantages, such as moderating incentives for perverting behavior and dealing with, instead of destroying it, the ambiguous picture of performance measurement (Bevan and Hood 2006; De Bruijn 2002). When public performance is multiple, the measurement of this performance should also be multiple.
- Performance measurement is a game, which implies that every performance indicator has a

limited life. Performance measurement will cease to be effective after some time, because professionals learn to play with it or the beneficial effects of performance measurement are realized or worked out (De Bruijn 2002). This dynamics of performance measurement maximizes, rather than minimizes, the perverse effects. Challenging the involved actors requires a lively and dynamic system of performance measurement. Bevan and Hood (2004) advised a game-proof design by introducing more uncertainty into the specifications of targets by making them transparent in process and in retrospect but not in real time.

- Performance measurement is basically an issue of trust and interaction (De Bruijn 2002; Noordegraaf and Abma 2007). Interaction can improve the confidence stakeholders have in each other and in the system of performance measurement. Measurements should be reasonable, i.e. connected to the goals of the various actors involved (Radin 2006; De Bruijn 2002). Both managers and professionals must agree upon the reasons, evidence and rationales that are relevant. De Bruijn suggests to set up rules how to use performance measurements that make stakeholders use the measurements in a predictable way and mitigate its impact (De Bruijn 2002)

What these answers have in common is that good performance measurement seems to find a balance between recognizing the complexity of public services without becoming as complex as the practices they represent. We found various contributory factors in literature that can explain why the limited representation is a strong point of the particular measurement, such as performance indicators can be used as tin-openers and give reason for more interaction, necessitate the development of multiple indicators and stimulate a lively practice of measurement. In this article we also want to contribute to the design of performance indicators that takes into account the complexity without becoming complex themselves. Our contribution to the debate would be to add another relevant alternative definition of ‘validity’, which acknowledges more radically that accuracy is not the main issue. In this paper we open the black box of a performance indicator. A performance indicator, we show, is much more differentiated and an active entity. Reduction is a step by step process starting at complexity and ending in one indicator. An *active* concept of validity and representation, developed in the field of Science and Technology Studies, can show in a new way how measurements can keep its simplicity and still be effective in complex settings.

The aim of our study is to understand how measurements can keep its simplicity and still be effective. We focus on a particular case, i.e. the use of prescription data by insurers to control the use of medicines in general practice. In the Netherlands insurers ask general practitioners to analyze prescription data in order to enhance the cost-effective and clinical effective use of medication. The prescription data in this particular case are rather general. It is just one measurement, i.e. the number of prescriptions of a certain drug per general practitioner. Also no corrections are made for

the particular population of patients of a general practitioner – or any other bias. The validity of prescription data as indicator for the cost-effective and clinical effective use of medication was a problem at first, but after the singularization of the sources from which the data was retrieved to one source, the accuracy of the data was hardly discussed any more. Note, the measurement itself was not changed; only the number of sources from which the data was retrieved was limited to one, namely the national database of pharmacists. Also no additional measures were developed. Neither have stakeholders been involved in the development of the measurement. How could simple or general measurements, with all its limitations, work?

Our analysis focuses upon the construction and use of prescription data. Thereby we are especially interested in the role of measurements as intermediaries between the complexities of clinical work and more simple performance goals. Given the complexity of pharmaceutical care, the question whether physicians prescribe drugs rationally (both clinically effective and cost-effective) cannot easily be answered by an inspection of their work. At the very least, an inspection of their work would produce seemingly inconclusive evidence. The question can only be settled through a formal inquiry which accurately but selectively represents clinical practice in such a way that specific relationships become visible while others, of course, remains hidden (see Latour 1987). The relationships specific to the medical practice re-presented by the numbers would otherwise remain hidden in the innumerable relationships that make up medical work. In this respect, measurements are inherently simple. Yet, measurements can coordinate activities and events at various locations and times without the need to bring all these elements together in one total measurement. In other words, measurements can link general and aggregated data with individual decisions and actions in day-to-day practice (Latour 1987).

The plan of the paper is as follows. First, we explain our theoretical perspective of measurements. Next, we analyze practices of prescription measurement in Dutch primary care. Thereby we focus upon the networks of measurements linking those exercising political power with the persons, processes and problems that they seek to govern (Rose 1999). In the last section, we present our conclusions where we reflect upon the contribution of our study to the performance measurement debate.

THE ACTIVE ROLE OF MEASUREMENT IN PUBLIC MANAGEMENT

Our approach is informed by the literature from science and technology studies (STS) and related fields of study (Latour 1987; Rose 1999; Callon 2001; Mackenzie 2003; Poster 2004; Mackenzie 2007). As Rose (1999) has explicated, measurements achieve political power. Measurements, for example, legitimize the power authorities or institutions claim. Numbers also transform individuals into percentages and thereby invisibly align private values with public goals. Last but not least,

measurements construct the public domain – “the domain of politics is numerically made” (Rose 1999: 198). How can we explain this active role of measurement in politics and public management?

Performance measurement is one aspect of a network consisting of many relationships. Measuring performance takes only some of these relationships into account in the calculations; others (the majority) are excluded (Callon 2003). Measuring performance means disentangling and dissociating the actors involved from their daily complexity (Callon 2003). The main quality of a measurement, according to Latour, is not to be objective or representative, but to be *presentable* – that is, *to present* absent objects or people to those that one needs to convince (Latour 1995). The development of a registry requires things or events to be inscribed in standardized forms and the inscriptions to be transported to a central locale, where they can be compiled, aggregated, compared and calculated. Yet, essential to measurements, however, is that they be *traceable*, that they allow one *to travel* from the complex field down towards individual graphs or measurements and from the measurements back to the entire field. Therefore, general or ‘simple’ measurements can control complex performances but only if they allow one to travel from detailed data to a more general indicator, and, more importantly, vice versa.

In the selection of what to measure and how to measure it implicit political judgments are made. In this paper we study the development of measurement as a layering of values: what is counted becomes normalized and what is not counted remains invisible (for a similar argument see Radin, 2006). This normativity of administration and other formal systems has been debated extensively in STS (Radder, 1998, Thevenot and Boltanski, 2006). A distinctive insight of STS studies is that measurement systems constitute or create the world they appear to represent; numbers do not merely inscribe a pre-existing reality in a database (MacKenzie, 2006; Rose, 1999). Performance measurements do not simply count elements in a pre-existing world, but help to create a world for which the measurement is a truer reflection (MacKenzie 2003). An illustration of this active role of measurement is given by MacKenzie (2003; 2007), who describes how the Black-Scholes-Merton theory has become the dominant theory for calculating the expected price of options in financial markets although the fit between the model and the prices of options on the market was not always close. According to Mackenzie, the success of the Black-Scholes-Merton theory – in terms of its use – cannot be explained either by the validity of the model or by the accuracy of the data, but by the model’s data sheets that help traders solve a variety of operational and organizational problems. The sheets were beautifully designed for their intended role of distributing relevant information and prices based on the Black-Scholes-Merton model. They included “what options traders using the Black-Scholes-Merton model needed to know, but no more than they needed to know – there is virtually no redundant information on a sheet – hence their easy

portability” (MacKenzie 2007: 62). As the sheets were used by almost every trader, because they helped them deal with several day-to-day problems in a relatively easy way, the financial markets began behaving more and more according to the assumptions of the Black-Scholes-Merton model. In other words, the fit between the model and the prices of options was not the starting point for the use of the model, but the end result.

We show in this paper that there is a constitutive relationship between specific forms of measurement and modern practices of self-governance. This idea is inspired by Foucault’s work on ‘governmentality’, where to govern human beings is to presuppose the freedom of the governed. Foucault explains how ‘techniques of the self’ respect the autonomy of the citizen yet mould them in such a way that their behavior is suitable for specific views about collective and individual welfare (Foucault 1988). Similarly, Rose (1999) warns us that we should not think of measurement systems purely as technologies of domination, for they can also be technologies of autonomization and responsabilization. Rose argues that numbers and measurement systems have a role in subjectification – “they turn the individual into a calculating self endowed with a range of ways of thinking about, calculating about, predicting and judging their own activities and those of others” (Rose 1999: 214). For instance, performance measurement systems create calculating selves, like workers trying to maximize their achievements in norms that could be calculated and embedded in the organization with the aid of the same measurement systems. People are enmeshed in these calculation networks both as objects and active participants. We analyse how a measurement system respects the autonomy of doctors and at the same time stimulates a practice of self-governance that fits in with the aims of political authorities.

Once a network is in place, new conduits of power can be exercised over distant events and processes: the things or events that one measures can be moved (to be re-presented) to those one wants to convince. Only at the registry – the centre of calculation – are the inscriptions aggregated, compared, and compiled. As we will describe pharmacists and the insurers gained a prominent position since they aggregate, compare and compile the data, and bring the prescription data to the doctors they want to convince (see Latour 1987). However, measurements should allow one to travel back from the measurements (the centre) towards the field (the periphery). Hence, neither insurers nor pharmacists can exercise power over the events and processes taking place in the clinic as the measurements allow only the doctors to travel from the general data down to the day-to-day decisions about prescriptions, and vice versa.

METHODS

Our study is designed as a case study with a particular attention to the context, rather than a certain

policy plan or intervention. The primary unit of analysis is a situation (see Clarke 2005). The situation we analyze is the way local pharmaceutical consultation groups (in Dutch *farmacotherapeutisch overleg*, abbreviated as FTO groups) use prescription data to enhance the clinical and cost effectiveness of the use of drugs on behalf of Dutch Insurers, which are increasingly responsible for the rational use of drugs in primary care since the introduction of regulated competition in the Netherlands. In this study we combine analysis of dominant discourses and organizational settings with detailed study of the tools general practitioners and pharmacists use to retrieve, analyze and present prescription data. Our analysis of this case focuses on the infrastructure developed to produce and compare prescription data. In particular, we deconstruct the series of intermediaries (tables, datasets, and forms) that make prescription data into a performance measurement. This series of intermediaries, we show, explain how general and aggregated data are linked with individual decisions in day-to-day practice - the chain of data allows the traveling from indicators to detailed data in day-to-day practice and vice versa. We would like to stress that non human elements, like databases and forms, might condition the interactions, just like human actors do. Non human agents can channel the flow of regulation more effectively than the representative actors can speak for them (see also the former section). These material infrastructures as we will show, explain the transformation of the relationship between state-actors and public service workers, or between regulators and professionals.

Our approach was developed in the context of two studies about the role of insurers in Dutch Health care that we conducted on behalf of the Ministry of Health and the Dutch Health Care Insurance Board. In 2004, for the Dutch Ministry of Health, we studied the new role of pharmacists as pharmaceutical care providers supposed to contribute to a more rational use of medicines (Niezen 2005). We conducted in-depth interviews with 11 key respondents, including the general manager of the database for pharmaceutical statistics, an IT expert from the Royal Dutch Association for the Advancement of Pharmacy (*Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie*, KNMP) who was responsible for the development of web-based tools to analyze prescription data, and representatives from the Royal Dutch Medical Association (*Koninklijke Nederlandsche Maatschappij tot bevordering der Geneeskunst*, KNMG) who discussed the role of pharmacists with us. Additionally, we carried out an extensive document study. In this study, just like the study we present in the next paragraph, the ministry of health had a strong say in how the problem of study was framed.

The second evaluation study was conducted in 2006. On behalf of the Dutch Healthcare Insurance Board (*College voor Zorgverzekeringen*, CVZ), we evaluated the possibilities for insurance companies to enhance the rational use of drugs by involving local pharmacotherapy

consultation groups of physicians and pharmacists in their policy measures (De Bont and Grit 2007). We conducted interviews with physicians and pharmacists (n=10) who were involved in projects sponsored by insurers to improve the pharmaceutical consultation meetings. Besides the interviews we conducted an extensive document study of recognized 'best practice' projects conducted by the consultation groups. As such, these projects were well described in professional journals. We also formally evaluated three of the projects.

These studies resulted in a self-organized network in which we could collaborate with policy makers, physicians, pharmacists and insurers towards understanding the use of prescription data. In this informal and self organized network we shared information through story telling searching for useful concepts. This network was of crucial importance for the richness of the data (Wrede 2005). Thanks to this network our study draws on the socially situated and distributed expertise about policy and practice (Jasanoff 2007).

Our analysis of the data is theoretical informed by the framework we introduced in the former section section. Main focus points were the development of databases and performance indicators, the use of data infrastructures, the functioning of FTO groups, debates about prescription behavior, the role of the different actors and their relationships. For this paper we also re-analyzed the interview transcripts and documents of the former two studies.

MEASURING AND CONTROLLING PRESCRIPTIONS IN PRIMARY CARE IN THE DUTCH CONTEXT

The Dutch Health Insurance Act came into effect on 1 January 2006. The new act integrates social (sickness benefit funds) and private health insurance for curative care. It requires all Dutch health insurers to operate under private law and obliges all Dutch citizens to buy individual health insurance coverage (Van der Ven and Schut 2008). The new system is aimed at increasing efficiency and responsiveness through enhancement of health insurers' competences (selective contracting) and the introduction of more competition among health insurers. Consumers can chose among insurers, and insurers can selectively contract healthcare providers. Although the introduction of this act was seen as an important marker for the start of the healthcare market in the Netherlands, market-oriented healthcare reforms had already been underway for quite some time (Helderman *et al.* 2005). In 1988 advocates of a new market-oriented program began seriously challenging the existing policy programs. Top-down rationing policies were subjected to growing criticism because of the lack of incentives for efficiency improvements and innovations to the whole healthcare system. However, the necessary technical and institutional preconditions for regulated competition

were not yet in place (Helderman *et al.* 2005; Van der Ven and Schut 2008).

With the gradual evolution of the Dutch healthcare system into regulated competition, insurers have become gradually an important sponsor of the healthcare system. Insurers gained more freedom to negotiate on price, service, and quality of care. The insurers' new role began in the domain of pharmaceutical care (VWS 2000). For instance, insurers must specify the contractual rights for the pharmaceuticals covered by their policies. Yet they are allowed to contract providers selectively and motivate consumers to use preferred providers or preferred drugs, for example, by providing financial incentives. Some large insurers are experimenting with some form of bonuses for prescribing generics. Others have begun setting up primary healthcare centers and pharmacies (Van der Ven and Schut 2008).

Dutch general practitioners (GPs) and pharmacists are organized in local pharmaceutical consultation groups (FTO groups). The FTOs started off in a city or region as small groups of GPs with one or two local pharmacists who informed each other about (new) pharmaceuticals (Paes and Cornips 1987; Jansen 2004). In the 1990s, the Ministry of Health began supporting the development of a more formal system of FTO groups, to stimulate a more rational use of drugs. The Minister of Health introduced a price-reference system but realized that the desired efficiency improvement would not be obtained by top-down measurement. The ministry then expected the FTOs to fill a crucial role in developing a more effective system of self-governance. Besides gathering and reviewing information about pharmaceuticals, these consultation groups would have to discuss the prescription behavior of individual GPs and pharmacists. Nowadays, there is a nation-wide web of 800 pharmaceutical consultation groups, including 90% of all GPs and pharmacists (DGV 2007). The work of these FTOs is closely monitored. Every year, the frequency of meetings, the availability of prescription data at meetings, the minutes and reports, and local prescription guidelines are measured. If an FTO meets at least four times a year, with 75% of its members present, and minutes are made of each meeting, most insurance companies reward the doctors in the group with an annual bonus of Euro 1500. Since 1999, the insurance companies have been encouraging FTOs to analyze prescription data. They ask the FTOs to set targets for drugs, to measure how often they prescribed a particular drug and deliver information on the numbers of patients receiving the drug prescribed. For instance, most insurers expect that for 80% of all new patients, GPs will prescribe the preferred drugs chosen by their consultation group. As this example shows political judgments are implicit in the measurements. Additionally, the measurements do not merely inscribe elements in a pre-existing world, but, as we will show later, helped to create a world that behave more and more according to the numbers.

Pharmacists play a major but far from transparent role in preparing FTO meetings and setting the

stage for target-setting in order to enhance the rational use of drugs. In fact, the role of pharmacists in the Dutch healthcare system is heavily debated. Since 2001 the Ministry of Health has taken steps to expand the role of pharmacists (Pronk *et al.* 2002; Muijters *et al.* 2004; IGZ 2003) from merely dispensing drugs to providing pharmaceutical care as well, including monitoring the rational use of drugs. The Royal Dutch Medical Association (KNMG), however, is against expanding the role of pharmacists. GPs want pharmacists to stick to dispensing a prescribed drug and not interfere with the prescription since the GPs reason that if a pharmacist changed a prescription this would affect the confidential relationship between doctor and patient (Cohen 2001; NIVEL 2004; NIVEL 2005; TNO 2005). The debate about the role of pharmacists has boiled down to the question whether a GP should add the patient's diagnosis to a prescription. Pharmacists claim they need the diagnosis to check and select the best drug. According to GPs, pharmacists do not need to know the diagnosis because in their view the pharmacist's only task is to dispense the prescribed drug.

Nonetheless, it is pharmacists, not GPs, who deliver the data for the FTO groups. The data infrastructure has opened up the discussion of GPs' professional behavior – a discussion which, incidentally, pharmacists became welcome to join in. We claim that pharmacists can take on this role because the Royal Dutch Association for the Advancement of Pharmacy (KNMP) has created the infrastructure for pharmacists to retrieve and deliver the data in the preferred format. The KNMP first developed a national database with data from the community of pharmacists. A database for pharmaceutical statistics was, the KNMP believed, of strategic importance for the profession. Since 1990, the KNMP has been registering dispensation data through the Foundation for Pharmaceutical Statistics (*Stichting Farmaceutische Kengetallen*, SFK). The data includes the drug or medical aid supplied on prescription, the dispensing pharmacy, the health insurance company that does or does not reimburse the remedy, the prescribing doctor and the patient for whom the prescription was issued (www.sfk.nl). Since 2003, the KNMP has also developed software (SFK web tools) for individual pharmacists to retrieve and analyze the data stored in the SFK database. The KNMP developed the tools in collaboration with DGV, the Dutch Institute for Rational Use of Medicine (Metz *et al.*, 2005). The tools include all the data and graphs a FTO group needs for a consultation meeting. The SFK uses these tools to automatically retrieve and analyze data for the pharmacist. Note: in contrast to GPs who receive a bonus of Euro 1500 basically for attending FTO meetings, pharmacists are not paid for the work of delivering the prescription data.

It is possible to obtain similar data from GPs' medical records; however, this is a cumbersome retrieval method. First, the GPs have to find the data in their own records, perhaps using the prescription codes DGV has instructed them to use. DGV, however, assumes that GPs code their data with standard codes whereas the GPs might have good reason, as Winthereik *et al.* (2007) have shown, to use other codes or even to record their data without codes. Second, the GPs

have to collate their separate data sets in one file in order to enable comparisons of data from the whole FTO group, a challenging task considering that GPs work in different places and with different systems. Last but not least, the data have to be analyzed and the statistics presented in hand-drawn graphs so that the group can actually discuss the findings.

Before the advent of SFK web tools, data were retrieved from various sources. GPs retrieved their own data from their electronic records and the pharmacists used the SFK. As GPs in the various regions or cities worked with different systems and used different codes, it was not unusual for them to make their own datasets and, as a result several different GP datasets would be presented for discussion by the FTO groups – in addition to the data from the pharmacists. All these different datasets on the table led to discussions of validity since, for example, a GP's position in the rankings could differ from one dataset to another. In due course, the SFK data was used more often as it was much easier to retrieve than any other available dataset, which makes their measurements more *presentable* than others. Moreover, the pharmacists were keen to deliver it to demonstrate their proficiency in their role as pharmaceutical care providers. With the arrival of the SFK web tools, all data could be registered centrally and at last the measurements could meet the presentation requirements. Following the arrival of web-based tools, the SFK data became the only dataset used for prescription management. Having just one dataset stopped the discussions over data validity. In this respect, the web-based tools played an essential role in the creation of validity, not because the SFK data were any more accurate than the data from GPs' records, but because the SFK tools construct a world in which these data fit.

The prescription data GPs need to deliver to the insurers are relatively general. Basically they show only the number of patients per doctor who received a certain drug. This measurement of prescriptions may count patients but it excludes all other factors that can explain why a certain drug was prescribed, such as population characteristics (age, socioeconomic status, environment and lifestyle) and the role of other doctors. Measurement by one simple indicator ignores the complex nature of the performance of physicians and pharmacists. For example, a drug is an important link in the delicate network between GPs, specialists and patients. Changing a prescription might affect the GPs confidential relationship with both the specialist and patient. Therefore GPs are inclined to prolong prescriptions issued by specialists, even when they have to overrule their own guidelines to do so (NIVEL 2004; NIVEL 2005; TNO 2005). A drug is also a trump card in negotiations between patients and physicians. Dutch GPs often advise patients to take an aspirin as a painkiller. Patients, however, do not always regard aspirin as a serious outcome of a consult with the doctor because they already use it as self medication. In these cases, the GP might decide to prescribe something else which in the eyes of the patient is a 'real' drug. Most of these relations are excluded. Hence, the patients who get certain drugs prescribed are turned into percentages, and with them the physicians

who prescribe the drugs.

However, the FTO groups succeeded in aligning simple indicator measurements with complex clinical practice, with the help of DGV and the SFK. DGV (2005) developed material that lets the GP trace an individual patient record and reconstruct the various rationales that made the GP decide to prescribe a drug. Figure 1 shows the tool that GPs use to track back from their total number of prescriptions down to the level of individual patients. For example, with this tool the GP can select the drug-prescription records that deviate from local guidelines, such as patients who use a drug longer than a certain time, or patients who are not prescribed the preferred drugs. Since these data are *traceable*, the GPs can recollect their decisions, and as they allow the GP to go back to the details of an individual patient, the simple measurements are precise enough. These simple indicators help GPs to critically reflect on their own and each other's practices and learn from one another, without destroying their individual autonomy. Even if GP behavior may not be determined by these performance indicators alone, these simple measurements do highlight inter-doctor variation, and the differences trigger discussion about the behavior of individual professionals. FTO groups are thus challenged to start a practice of self-governance. As Rose explicated numbers not only transform individuals in percentages but -thereby - also align private values with public goals, such as the rational use of drugs.

[Insert figure 1 somewhere here]

There is a constitutive relationship between general forms of measurement and systems of self-monitoring and self-governance which in turn fits the goals of political authorities. Our analysis of the construction and use of general measurement explains how compliance with the goals of insurers is possible even though professionals distrust interventions by insurers. A crucial explanatory factor is that while the simple indicators enable the professionals to retrieve more detailed data, the data permit the insurers to break with complex practices. In this respect, simple measurements have an extra advantage; they set actors *at a distance*. Rather than technologies of domination, general prescription data act as technologies of autonomization. In our case only GPs are allowed full access to their patient records, thus no one else can take over their position (or decisions). The insurer has no access to the GPs' individual patient data. In this respect, multi-dimensional measurement might actually hinder compliance with performance measurement because it would show more patient characteristics, such as diagnosis, age, sex and co-morbidity and shorten the distance between physicians and insurers. The notion of distance also explains why health insurance companies do not use their own data to control prescriptions. Insurers have always kept their own prescription data for the drugs they reimburse. In 1988 the Healthcare Insurance

Board started up a database called GIP (*Geneesmiddelen Informatie Project*, medicine information project) to analyze these data. GIP contains expenditure information on (extramural) drugs and prescription-related data on drugs from nine health insurance organizations that have almost half of the Dutch population as their clients (7.5 million clients) (www.GIP.nl). In other words, instead of using their own resources, the insurers ask GPs to deliver the required data they already have for themselves. The insurers' behavior constructs a distance between themselves and the professionals that enables co-governance. Insurers, GPs and pharmacists maintain a distance between each other to be able to control performance together.

CONCLUSION

The aim of this article is to contribute to the design of simple indicators that can take the complexity of professional work into account. As was explained in the introduction the effectiveness of performance indicators depends on its simplicity. Its simplicity not only makes performance measurement into a strong communication tool, it also prevents strategic behavior (De Bruijn 2007).

We started our analysis by describing the political context of measurements as the role they fulfill relies on relationships provided by political authority (Singh 2002, Tuohy 2003; Fox 2006;). As we showed in this paper, the roles of two central actors –i.e. the insurers and the pharmacists– were contested. As the Dutch Health Care Act came into effect in 2006, insurers were supposed to become the sponsors of the new health care market. Yet, insurers were hesitant to take a role in the health care governance of the quality of care. Pharmacists were contested for merely dispensing drugs, but were not allowed either to advance their roles as pharmaceutical care givers. What we show in this paper is how simple performance measurements created a new kind of political authority for both insurers and pharmacists that emanates from databases - which is not the same as legitimating and supporting the political authority of insurers, which is still contested as in the past. The political authority of measurement is not characterized by rule-following, since there are no strict rules for the GP to follow. Measurements create ranking among GPs, which generates discussions of behavior. Meanwhile, owing to their limitations, the simple, general indicators see to it that potential evaluators (e.g. insurers) are held at a sufficient distance. The prescription data function as what Foucault calls 'techniques of the self' that respect the individual autonomy of GPs yet mould their behavior in such a way that it fits with the collective aim of a rational use of drugs. The degree of compliance to performance measurement and thereby the prospect for effective co-governance can be explained by reflexive distancing, that is 'general' indicators enable stakeholders to distance themselves from each other.

Although the description of the political context is proven relevant, we are hesitant to include

formally institutionalized politics as an explanation of how performance measurements work. Such an explanation would separate politics from administration, while our ambition is to link the technical, practical and operational practices surrounding the construction of data to broader and more political questions about governance. As Gullick (1936) wrote, “The real work of government is not to be found behind the Greek columns on public buildings, but rather on the land, among the people” (cited in Rose, 1999). In our case study, how governance works is found in information technologies and systems, such as financial instruments, databases and calculations.

To contribute to the work of Carter, De Bruijn, Radin and others, we introduced (active) concepts culled from science and technology studies to the analysis of the working and the dynamics of performance measurement. By analyzing the construction and use of general prescription data by pharmacists and GPs in the Netherlands, we were able to describe the advantages of a simple and general measurement to performance management in comparison to a more extensive set of performance indicators. First, simple performance measurement acts *as a coordination mechanism* between general (policy) goals and individual decisions or behavior in day-to-day practices. The ‘poor’ quality of general prescription data was an invitation for GPs *to go back* to their patient records to get additional data. Once the statistical analyses were done, the files were not closed. On the contrary, sum totals were neither the outcome nor the end result of the ongoing process of making use of prescription data. Along with the prescription data, GPs were offered several tools to trace individual records and recall specific decisions the GP has made in the clinic. These tools allowed GPs to travel from graphic visualizations of sum totals down to the complex totality of their work, and back up to the simplicity of the indicator. As the tools allow the GPs to mediate between the simple indicators and the complexity of their work, simple indicators did not need to be replaced or enhanced by more complex sets of indicators. Consequently, the perverse effects of performance measurement kept minimized.

Secondly, simple indicators put insurers at *the ‘right’ distance*. The involvement of stakeholders in the process of developing measurements is often seen a key condition for measuring performance, as it is one way of preventing measurements from becoming too distant from local practices. Yet, our findings indicate that it is this very distance between stakeholders that explains the degree of compliance that is achieved – not their proximity. As we explained in the theoretical section, with central registration new conduits of power can be exercised over distant events and processes because the input can be aggregated, compared, and compiled so that all the things or events that one measures can be re-presented to others. Yet the networks of measurement connecting those exercising political power with the processes, problems and persons they seek to govern, also provide the latter with power. As we showed in this article, both the prescription data and the tools that allow GPs – only GPs - to go back to the medical records reconstruct the

relationships between the insurers on the one hand and the GPs on the other hand. Essential to measurements, is that they allow one to travel from the complex totality of the field down to graphic visualizations of individual measurements, and back up to the entire field and in our case, only GPs can trace the measurements back and forth.

Our research specifies two hypotheses for how simple performance indicators can function as successful tin-openers. First, effective performance measurement do not imply more indicators but to move backwards to where the performance started. i.e. back to the field. Second, auditing authorities should better not be involved in the opening of the ‘can of worms’. They should reflect upon the right distance. We are very cautious to generalize these claims to other settings or countries, since performance measurement is always embedded in a specific setting. Yet we hoped to offer a framework for further research that is needed to test these hypotheses.

LITERATURE

Bevan, G. and C. Hood. 2006. ‘What’s Measured is what Matters: Targets and Gaming in the English Public Health Care System’, *Public Administration*, 84, 3, 517-538.

Bont, A. de and K. Grit. 2007. *Sturingsinstrumenten voor zorgverzekeraars. Een bestuurskundige analyse van de mogelijkheden om het farmacotherapeutisch overleg te optimaliseren*. Rotterdam: iBMG.

Callon, M. 2001. ‘Writing and (Re)writing Devices as Tools for Managing Complexity’, in: Law, J. and A. Mol (Eds.), *Complexities in Science, Technology and Medicine*. Durham: Duke University Press.

Callon, M. 2003. ‘Actor-Network Theory – the Market Test’, published at: <http://www.comp.lancs.ac.uk/sociology/papers/Callon-Market-Test.pdf>.

Carter, N. 1994. ‘Performance Indicators; ‘backseat driving’ or ‘hands off’ control?’, in: McKevitt D. and A. Lawton (Eds), *Public Sector Management. Theory, Critique and Practice*. London: SAGE Publications.

Clarke, A. (2005). *Situational Analysis. Grounded Theory After the Postmodern turn*. Thousands Oaks: Sage.

Cohen, D., M. McCubbin, J. Colli and G. Perodeau. 2001. ‘Medications as Social Phenomena’. *Health*, 5,4, 441-469.

De Bruijn, D. 2002. *Managing performance in the public sector*. London: Routledge.

DGV. 2005. *FTO module Maagklachten, afbouwen van onterecht chronisch gebruik van maagzuurremmers. Werk materiaal voor een FTO bijeenkomst*. Utrecht: DGV.

- DGV. 2007. *FTO Peiling 2007. Farmacotherapieoverleg in 2007*. Utrecht: DGV.
- Dubnick, M., 2005, Accountability and the promise of performance: in search for the mechanisms, *Public Performance and Management Review*, 28, 3, 376-417.
- Foucault, M. 1988. *Politics, philosophy, culture: interviews and other writings of Michel Foucault, 1977-1984*. London: Routledge.
- Fox, N., K. Ward, A. O'Rourke. 2006. 'A Sociology of Technology Governance for the Information Age: The Case of Pharmaceuticals, Consumer Advertising and the Internet', *Sociology*, 40, 315-334.
- Gullick 1936, 'Introduction', in L. Harrison and E. Lane, *After Repeal: A Study of Liquor Control Legislation*, New York: Harper, pp. i-xxxv.
- Helderman, J. and F. Schut, 2005, 'Market-Oriented Health Care Reforms and Policy Learning in the Netherlands'. *Journal of Health Politics, Policy & Law*, 30, 1-2, 189 –221.
- IGZ. 2003. *Voorwaarden voor verantwoorde farmaceutische zorgverlening door openbare apothekers onderzocht*. Den Haag: IGZ.
- Jansen, J., Van Dijk, L., Wagner, C. and D. de Bakker (Eds.). 2004. *Professionalisering van het FTO in Noord-Holland-Noord*. Utrecht: NIVEL.
- Jasanoff, Sheila, 2007, *Designs on Nature: Science and Democracy in Europe and the United States*, Princeton University Press.
- Jos, P., M. Tomkins, 2004, 'The accountability paradox in an age of reinvention', *Administration & Society*, 26, 3, 255-81.
- Kanter, R.M. and D.V. Summers. 1994. 'Doing Well While Doing Good: dilemmas of performance measurement in nonprofit organizations and the need for a multiple-constituency approach', in: McKevitt D. and A. Lawton (Eds), *Public Sector Management. Theory, Critique and Practice*. London: SAGE Publications.
- Klein, R. and Carter, N., 1988, 'Performance measurement: a review of concepts and issues', in D. Beeton (Ed.), 1988, *Performance Measurement: Getting the concepts right*, (London, Public Finance Foundation).
- Latour, B. 1987. *Science in action: how to follow scientists and engineers through society* Cambridge: Harvard University Press.
- Latour, B. 1995. 'The "Pedofil" of Boa Vista—A Photo-Philosophical Montage'. *Common Knowledge* 4,1, 145-187.
- Lindgren, L. 2001. 'The Non-profit Sector Meets the Performance-management Movement: A Programme-theory Approach'. *Evaluation*, 7,3, 285-303.
- MacKenzie, D. 2003. 'An Equation and its Worlds: Bricolage, Exemplars, Disunity and Performativity in Financial Economics'. *Social Studies of Science*, 33, 12, 831 - 868.

- MacKenzie, D. 2007. 'Is Economics Performative? Option Theory and the Construction of Derivatives Markets', in D. Mackenzie et al, *Do Economists Make Markets? On the Performativity of Economics*. Princeton: Princeton University Press. Page 54-86.
- Metz, J, P. Smet, J. Tinke (2005). *Het ontwikkelen, implementeren van praktijkprogramma's rond FTO en FPZ*. Den Haag: KNMP.
- Muijrsers, P., J. Knottnerus, J. Sijbrandij, R. Janknegt and R. Krol. 2004. 'Pharmacists in primary care. Determinants of the care-providing function of Dutch community pharmacists in primary care'. *Pharmacy World Science*, 26, 256-262.
- Niezen, M. , A. de Bont en H. Stoevelaar. 2005. *Evaluatie MJA deeltraject. Het waarborgen en verbeteren van de kwaliteit van de farmaceutische zorg*. Rotterdam: IBMG.
- NIVEL 2004. *Professionalisering van het FTO in Noord-Holland Noord*. Utrecht: NIVEL.
- NIVEL 2005. *Doelmatig voorschrijven van antidepressiva in de huisartspraktijk*. Utrecht: NIVEL.
- Noordegraaf, M. and T. Abma. 2007. 'Management by Measurement? Public Management Practices Amidst Ambiguity', *Public Administration*, 81, 4, 853-871.
- Paes, A. and P. Cornips. 1987. 'Ontstaan en verloop van farmocotherapie-overleg', *Pharmaceutisch weekblad*, 122,10, 255-230..
- Poster, M. 2004. 'The Information Empire.' *Comparative Literature Studies*, 41,3, 317-333.
- Power, M. 1997. *The audit society: rituals of verification*. Oxford, Oxford University Press.
- Pronk, M., L. Blom, R. Jonkers, E. Rogers, A. Bakker en K. De Blaey. 2002. 'Patient oriented activities in Dutch community pharmacy: diffusion of innovations', *Pharmacy World Science*, 24, 154-161.
- Radder, H. (1998). 'The Politics of STS', *Social Studies of Science*, 28, 2, 325-328
- Radin, B. (2006). *Challenging the performance movement. Accountability, complexity and democratic values*. Washington: Georgetown university Press.
- Rose, N. 1999. *Powers of Freedom. Reframing Political Thought*. Cambridge: Cambridge University Press.
- Sanderson, I. 2006. 'Complexity, "practical rationality" and evidence-based policy making'. *Policy and politics: studies of local government and its services*, 34, 1, 115-132.
- Singh, J. P. 2002. 'Introduction: Information technologies and the changing scope of global power and governance', in Rosenau, J and J. Singh (Eds.). *Information technologies and global politics*. Albany: State University of New York Press.
- Thevenot, L. and Boltanski, 2006, *On justification. Economies of worth*, Princeton, Princeton University Press.
- Tuohy, C. 2003. 'Agency, Contract, and Governance: Shifting Shapes of Accountability in the Health Care Arena', *Journal of health politics, policy and law*, 28, 2-3, 195-216.

- TNO. 2005. *Doelmatig voorschrijven van nieuwe geneesmiddelen; off label voorschrijven van cox-2 remmers als casus*. Leiden: TNO.
- Van Peurse, K.A., M.J. Pratt and S.R. Lawrence. 1995. Health management performance. A review of measures and indicators. *Accounting, Auditing and Accountability Journal*, 5, 34-70.
- Ven, W., van der and Schut, F. 2008, 'Universal Mandatory Health Insurance In The Netherlands: A Model For The United States?'. *Health Affairs*. 27, 3, 771-781.
- VWS. 2000. *Kabinetstandpunt geneesmiddelenbeleid*. GMT/ G2064781. Den Haag: VWS.
- Winthereik, B., I. van der Ploeg and. M. Berg. 2007. 'The Electronic Patient Record as a Meaningful Audit Tool: Accountability and Autonomy'. *Science Technology Human Values*. 32, 6-25.
- Wrede, S., Benoit, C., Lynn Bourgeault, I., E.R. van Teijlingen, J. Sandall and R.G. De Vries. 2006. Decentred comparative research: Context sensitive analysis of maternal health care. *Social Science & Medicine*, vol. 63, issue 11, pages 2986-2997.

Patient's name, sex, date of birth	Dispensing pharmacist [name]	Drug used [name]	Indication according to protocol to reduce the use of the drug	GP's assessment: Is it possible to reduce the patient's use of the drug?	Action to be taken	Consult date	Consult result	Result after three months
				0 Yes 0 Perhaps 0 No	0 Yes 0 Perhaps 0 No			
				0 Yes 0 Perhaps 0 No	0 Yes 0 Perhaps 0 No			
				0 Yes 0 Perhaps 0 No	0 Yes 0 Perhaps 0 No			

Figure 1: A table that Dutch GPs use to trace their prescription behavior down to the level of individual patients. Source: DGV, the Institute for Rational Use of Medicines.