Problem structuring in Health Technology Assessment

An argumentative approach to increase its usefulness

Problem structuring in Health Technology Assessment

An argumentative approach to increase its usefulness

Een wetenschappelijk proeve op het gebied van de Medische Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor aan de Radboud Universiteit Nijmegen
op gezag van de Rector Magnificus prof. mr. S. C. J. J. Kortmann,
volgens besluit van het College van Decanen in het openbaar te verdedigen
op dinsdag 4 november 2008 om 15.30 uur precies

door

Margriet Moret-Hartman

geboren 27 April 1976
te Est en Opijnen

Promotores:
Prof. dr. G. J. van der Wilt
Prof. dr. J. Grin, *Universiteit van Amsterdam*

Co-promotor:
Dr. R.P.B. Reuzel

Manuscriptcommissie:
Prof. dr. E. van Leeuwen (voorzitter)
Prof. dr. R. Hoppe, *Universiteit Twente, Enschede*
Prof. dr. T. J. de Cock Buning, *Vrije Universiteit, Amsterdam*

ISBN 978-90-9023326-0

Gedrukt door: Ponsen & Looijen BV, Wageningen

Table of Contents

1

Introduction: Improving the usability of Health Technology Assessment

In an interview for a Dutch newspaper, Chalmers stressed the need for more research that provides answers to clinically relevant questions (NRC, 28th October, 2006). Indeed, there are many examples of research projects whose findings appear to have little relevance to clinical practice. Although a cost-effectiveness study demonstrated that ventilation tubes have limited value in the treatment of otitis media for most children without complaints (Hartman et al., 2001), these results hardly affected the clinical practices of otolaryngologists (van der Wilt et al., 2004). Based on two studies demonstrating a decrease in the number of diagnostic tests, perioperative complications, and treatment costs (Langmeijer et al., 1996; Rutten et al., 1995), the Health Council advised all hospitals to implement an outpatient anaesthesiological assessment (Gezondheidsraad, 1997). Despite these findings and recommendations, anaesthesiologists in one hospital considered the implementation of such outpatient assessments to be of little benefit. They felt that their existing pre-operative anaesthesiological practices were sufficient, and saw no need to change them. Possible savings in the number of diagnostic tests and costs were limited (Hartman & van der Wilt, 1999).

These examples illustrate a problem that has been frequently discussed in Health Technology Assessment (HTA). Although much evidence has become available, this information is often not applied to health policy or clinical practice. This thesis addresses the problem of the limited impact of HTA research on health policy and clinical practice and investigates the ways by which we can improve the usability of research for health policy and clinical practices.

Health Technology Assessment and use

Recent years, there is a rapid growth in the number of HTA studies being performed (Banta, 2003; Draborg et al., 2005). HTA is a type of policy research that aims to provide information concerning medical technologies in order to support health care decision making. Ideally, such assessments not only evaluate the effectiveness of the intervention but also take into account its potential economic, social, cultural, legal, ethical, and/or organisational consequences (Raad voor Gezondheidsonderzoek, 1998).

HTA can be considered a specialised field of Technology Assessment. The Office of Technology Assessment (OTA) was established in the 1970s to assess rapidly developing technologies. The reports from the OTA included comprehensive policy

research on the social consequences of the short and long term use of technological innovations and presented an example of a new analytical tool to direct policy making (Banta, 2003; Lehoux & Blume, 2000). Although HTA can be considered a specialised form of Technology Assessment, its assessments are generally less extensive. In practice, these studies emphasise the efficiency of interventions, while corresponding ethical and organisational implications are only superficially addressed (Banta, 2003; Berg et al., 2004; Lehoux & Blume, 2000; Reuzel & van der Wilt, 2000).

In the Netherlands, HTAs are usually cost-effectiveness analysis, undertaken by universities or university hospitals. Furthermore, extensive reviews on the state of affairs concerning medical subjects are provided by the Health Council. Until 2000, the national HTA programme, 'Developmental Medicine' (ontwikkelings-geneeskunde) was under the auspices of the Netherlands Sickness Funds Council (the current Health Care Insurance Board) (Raad voor Gezondheidsonderzoek, 2004; Berg et al., 2004). The Dutch Ministry of Health established this fund in order to enable the evaluation of the costs and effects of new, innovative clinical interventions. Presently, a great deal of research on the efficacy of interventions is subsidised by the 'Health Care Efficiency Research Programme' from the Netherlands Organisation for Health Research and Development (ZON-MW). Research may be commissioned directly by the Ministry of Health, the Health Care Insurance Board, or clinical practitioners. The information provided by the HTA is used by both policy makers (at the Ministry of Health or organisations such as the Health Care Insurance Board) and professionals in the field (physicians and hospital administrators). Professionals may apply the HTA results directly to their area of speciality or they may be influenced indirectly through subsequent policy measures.

Although, in many cases, HTAs appear to affect health care decisions, their actual contribution has been frequently debated (Drummond et al., 1997; Drummond & Weatherly 2000; Duthie et al., 1999; Graf van der Schulenburg, 2000; Oliver et al., 2004; Williams & Bryan, 2007). Policy decisions concerning the introduction of (new) technologies in health care are not always based on the results of HTA studies (van den Heuvel et al., 1997; Berg et al., 2004), cost-considerations, are rarely incorporated in clinical guidelines (Berg et al., 2004; Niessen et al., 2007), and behavioural changes as the result of a HTA occur infrequently (Raad voor Gezondheidsonderzoek, 1998).

Currently, two solutions that seek to increase the applicability of research are being instituted, namely the standardisation of research methods and the active implementation of research findings in clinical practice. Standardising the research methodology is expected to increase the credibility of the results and thereby their utility. Actively promoting the use of evidence and developing guidelines for summarising the evidence seek to overcome the passive dissemination of new evidence, thereby decreasing the gap between the available information and its use in clinical practice (Bero et al., 1998; Grol et al., 1998; Haines & Donald, 1998; Haynes & Haines, 1998;, Grol & Grimshaw, 2003; Grol & Wensing, 2004; Gagnon et al., 2006).

Limited usability of HTA

The abovementioned approaches may improve the utilisation of HTAs to some extent. However, there may be an alternative solution to increase the degree to which research-based findings are used in decision making. This solution is based on the 'argumentative approach' developed by the policy sciences. A possible explanation for the limited impact of HTAs is that they do not sufficiently answer the questions deemed important by their potential users, namely policy makers and health care professionals. These users are likely to see the problems different than the problems identified by the researchers. Interventions that are judged as a solution from one perspective might be irrelevant from another perspective.

Theoretical framework: the argumentative approach in policy analysis

Also within policy sciences, the relevance of policy and the need for adequate implementation of that policy has been discussed. The development and implementation of a policy is shaped by the interactions between all actors involved (stakeholders). Stakeholders are actors directly involved in the decision making process and actors who are, or might be, affected by any action taken by an organisation or group. (Derthick, 1972; Mazmanian & Sabatier, 1983; Pressman & Wildavsky, 1973). As Elmore (1985) has argued, this implies that policy design should iterate between forward mapping, translating the social problem into a policy problem and then generating policies to deal with the problem, and backward mapping, analysing ex-ante the extent to which implementers and target groups will

likely respond to these policies in ways that help resolve the social problem prior to implementing the policy.

The theory of argumentative policy analysis provides a possible explanation for the failed implementation of policy measures. The basic idea of the argumentative approach is that actors' behaviours can be explained by different views on a problem and the argumentation behind these views (Fischer & Forrester, 1993; Yanow, 1996; Grin & van de Graaf, 1996a; Hoppe & Peterse, 1998). Central to the argumentative approach (Fischer & Forester, 1993; Fischer, 1999; Fischer, 2003), is the notion that action is driven by processes of problem setting (Schön, 1983) in which actors iterate between forward mapping and backward mapping in order to define problems and solutions that correspond with each other and the actors' normative and empirical backgrounds. The way in which a problem is defined depends on the assumptions the actors have about the situation and their beliefs regarding what is considered good practice (normative values). Based on these insights, authors from the argumentative approach emphasised the importance of considering these differences in problem definitions and the underlying arguments during policy development.

In reaction to the problems that arise during the implementation of policy programs, much attention has been paid to the evaluations of implemented interventions. Evaluation is usually restricted to assessing whether a specific programme or intervention has, or has not, fulfilled the programme objectives. The problem is that the outcome of an evaluation depends heavily on the criteria used to evaluate it (Fischer, 1999). Disputes between actors on the meaning of a specific intervention frequently can be explained by their disagreements with respect to which criteria actors consider relevant for assessing the success or failure of an intervention.

The evaluation of cochlear implants provides a concrete example of the consequences of divergent assumptions. A cochlear implant (CI) is a small electronic device that can help a person who is profoundly deaf perceive sound. The implant consists of an external portion that is placed behind the ear and a second portion that is surgically implanted under the skin. The stated objective behind implementing CI was to allow deaf children to hear. To measure the effectiveness of the implants, most evaluations included the perception and production of speech as an outcome measure. Based on the results of an evaluation performed in the Netherlands (Severens et al., 1997), the Ministry of Health was advised to include this technology in the health package. However, the ministry could not reach a decision when

confronted with protests from the deaf community. For many deaf people, the ability to perceive sound through technology was an insufficient reason to adopt the new technology. Apparently, the evaluation did not provide sufficient evidence that could be used by all the stakeholders to make decisions regarding cochlear implants.

According to Fischer (1999), the main problem is that such evaluations fail to take into account the underlying assumptions and normative values which influence the evaluation of specific interventions. Frequently, the assessment is limited to simply evaluating whether the objectives of the programme are met. It is a mistake, however, to assume that a single analyst can objectively identify which specific objectives should be met. This requires an analysis about the scope and purpose of public policy. When is a programme successful? When is an intervention worthwhile?

Framework for a full evaluation

Fischer (1999) presented a framework for including empirical and normative concerns in evaluation and termed this the 'logic of practical deliberation'. A full evaluation should not only include an assessment of the situational context of the intervention but also an assessment of the more general assumptions and normative values. An evaluation of the situational context, also called a first order evaluation, includes both the measurement of outcomes and the identification of relevant criteria for determining success. An evaluation of the more general assumptions includes an assessment of the underlying assumptions and normative values of the stakeholders. This is termed a second order evaluation. A summary of these four levels is presented in Figure 1.1 and will be explained in more detail below.

The first level of evaluation measures whether an intervention meets the defined outcome measures (verification of program outcomes). For example, the evaluation of CI consisted of measuring the perception and production of speech in children who had received a CI. Most evaluations take place at this level. If there is no consensus on the relevance of an intervention, Fischer proposes that the evaluation should move to the next level, namely validation. Evaluation at this level deals with identifying relevant criteria. During this phase, one should assess whether outcome measures are relevant and valid in the given situation. Which outcome measures are relevant depends on how the problem is defined and the purpose of the improvement.

Returning to the problem of CI, some years later, Reuzel (2004) performed an interactive analysis on the use of CI for deaf children. In this analysis, he aimed to identify criteria for success and acceptability. He organised an interactive discussion among stakeholders in order to discuss the problems and relevant criteria for success.

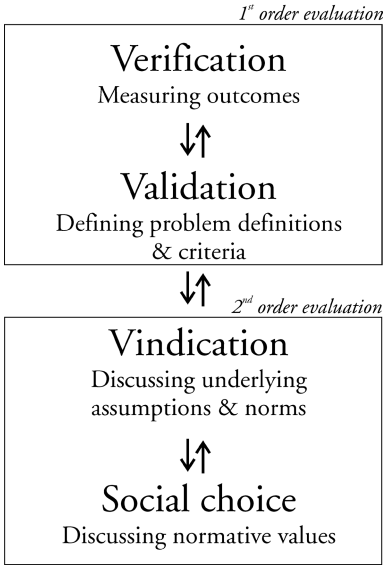

Figure 1.1 Four levels of a full evaluation (Fischer 1999)

In the case of CI, possible stakeholders included physicians, parents, counsellors, and deaf children. The interactive analysis indicated that some stakeholders considered the effect of CI on the social and emotional well-being of children implanted with the device more relevant than the potential perception of sound. The problems were defined differently by the advocates and opponents of CI. As a result, the choice with respect to outcome measures was not obvious. Evidently, each actor's preference with respect to the outcome measure was dependent on that actor's definition of the problem and the various actors had divergent problem definitions. It is important to note that a problem is frequently defined as the gap between the actual situation and the desired situation. Both the actual and desired situation might be viewed differently by different stakeholders.

If there is dissension regarding the nature of the problem, the evaluation should shift to the next level, which involves an assessment of the actors' underlying

theoretical and normative assumptions. This is termed vindication. People can have assumptions on the mechanisms behind intervention (theoretical assumptions) and on how these interventions are effective in meeting the goals of the society (normative assumptions). People working in different disciplines will frame the situation or a problem differently (Rein & Schön, 1993, p. 147). Whether someone considers a state of affairs to be evidence supporting a hypothesis depends on his or her background beliefs and assumptions (Longino 1990, p. 43). These considerations are no longer limited to the interventions. What we see then is that the evaluation moves towards the broader social context in which the evaluation takes place. This is then a second order evaluation.

With respect to the problem of CI, differences in the perception of the problem between the opponents and advocates of its use were likely the result of the differing assumptions concerning deafness itself. Deafness can be seen as a pathological disorder which prevents people from fully participating in society. In this context, spoken language is considered necessary for individuals to adequately function in society. Children with hearing deficits should thus receive a CI so they can learn how to communicate orally so that they can communicate and function within the larger community. However, others regard deafness as the distinct linguistic and cultural aspect of a group of people. They considered the acquisition of sign language necessary for the individual's social and emotional development. Sign language is considered by these actors to be someone's natural language. Furthermore, they would contend that the introduction of CI could have consequences for the future of the deaf community (for example, the financing of certain facilities).

If it is not possible for the concerned parties to reach agreement on which assumptions are legitimate, the evaluation must move to the fourth and final level. At this level, the normative values that are held by the participants (social choice) are examined, since it is believed that these normative values strongly affect which theoretical assumptions are considered relevant and how the problem is defined. The relevance of the underlying assumptions is not determined through empirical evidence but rather by an examination of the values and beliefs held by the participants (Longino 1990, p.57).

Considering the case of CI, at least two different normative positions could be identified among the participants. The first is the medical perspective which seeks to, firstly, eradicate what they see as a disability and, secondly, integrate deaf people into

hearing society. However, many people within the deaf community, value maintaining the deaf community as a distinct community with its own language and culture. These people hold the second normative perspective. In the interactive evaluation, the existence of these conflicting values were presented (Reuzel, 2004) and criteria that were considered relevant for both normative values were identified. In conclusion, in many cases the criteria used to evaluate the outcome of an intervention are inadequate. Consequently, it may be desirable to examine the issues surrounding the intervention more comprehensively. Since the underlying beliefs that determine which criteria are relevant can differ between actors, an analyst should consider carefully how the actors in a given situation define the problem and which outcome measures are relevant. When there is no agreement between the parties on the definition of the problem, the underlying beliefs and values influencing their perception of the problem should be considered.

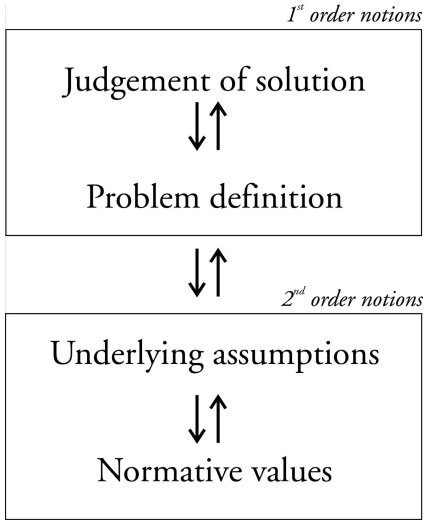

Figure 1.2 The four layers of someone's interpretive frame (Grin et al., 1997)

Interpretive frames

Grin et al. (Grin & van de Graaf, 1996b; Grin et al., 1997) developed an instrument to reconstruct someone's problem definition and the underlying argumentation. This instrument draws on the four-phase model for evaluation proposed by Fischer (1999).

The idea is to reconstruct (for example, during an interview) how interventions are judged, which criteria are considered relevant and the way in which the problem is defined the theoretical and normative assumptions that shaped them, and someone's preferences (Figure 1.2). These four layers of evaluation together constitute an actor's 'interpretive frame'. In Table 1.1, the interpretive frames of the two viewpoints that manifested in the CI case are summarised.

Careful analysis of perspectives in terms of interpretive frames helps to determine what is agreed and disagreed upon. When the interpretive frames of the actors are found to be divergent with respect to problem definitions and normative assumptions, there is a lack of congruence between interpretive frames. According to Hisschemöller (1993), the kind of research needed to solve a problem depends on the kind of problem one has to deal with. If actors agree on the problem definition and which norms are at stake, the problem is well structured (Table 1.2) and can be solved using standardised (quantitative) methods. However, if actors disagree on the nature of the problem, what information is required to evaluate an intervention and which normative values are at stake, the problem is ill-structured. If the problem is ill-structured, conventional research methods might lead to the over-simplification of the problem. As a result, the wrong problem may be solved, or the problem may be only partially solved. Addressing the wrong problem occurs frequently when there is no acknowledgement of the diverging perspectives the actors have towards the situation (Dunn, 2004; Hisschemöller, 1993). Clearly, when a problem is ill-structured, alternative methods are necessary to structure the problem.

Table 1.1. Interpretive frames concerning CI, based on Reuzel 2004

	medical perspective	member of the deaf society
judgement solution	CI is a valuable instrument	CI is only a moderately valuable aid
problem definition	Communication between deaf people and hearing people is difficult.	With CI, deaf people should face high expectations, because child with CI will be hard of hearing
theoretical assumption	Deafness is a disability. Offering spoken language is needed for rapid development of the child's orientation to spoken language. Oral language relevant for functioning in society; communication improves someone's well being.	Offering sign language needed for the child's social and emotional development. Sign language is someone's natural language.
normative value	Integrate deaf people in the community	Maintaining a deaf community with its own language and culture

Table 1.2. Types of policy problems according to Hiscchemöller (1993)

		Agreement on needed information	
		yes	no
Agreement on which norms are at stake	yes	well structured problem	moderately structured problem
	no	moderately structured problem	ill structured problem

Ignoring differences in problem definitions in HTA

Returning to the problem outlined at the beginning of this chapter regarding the limited impact of HTA, we contend that if an HTA has only a limited impact on policy and practice, it is possibly the result of the assumption that there is agreement between all stakeholders on the criteria for evaluation. HTA is usually restricted to the level of verification, assessing whether an intervention meets the predefined criteria. However, HTA might not adequately take into account the possible difficulties that result from divergent problem definitions resulting from differing assumptions and values. In the case of CI for children, the initial study was restricted to evaluating the cost-effectiveness of the procedure. This provided insufficient information to support health policy decisions made by the Ministry of Health. The evaluation in which the role of underlying assumptions and normative values were acknowledged provided additional information that was more relevant to the different actors involved. The evaluation on CI was successful because it resulted in a number of clear policy recommendations supported by all participants (Reuzel, 2004).

Critics might emphasise that the example provided by CI does not fully correspond to mainstream HTA research. In mainstream HTA, the issues seem less complex. Indeed, CI was a case in which the intervention was strongly debated from the beginning. Nevertheless, the hypothesis is that it is imperative to consider the stakeholders' perspectives, even when dealing with technologies that initially appear to be less controversial. In many cases, the actors involved may define the problem differently based on differences in education or profession. Different ways of framing a disease result in different strategies for resolution and different views on what research is considered to be relevant (van der Wilt, 1995) At the same time, the

cooperation of these actors is required to successfully implement the intervention. The direction and goals of health care policy has shifted towards policy making as a process of co-production. A functioning health care system is seen as the joint responsibility of all the relevant parties, including hospitals, health care professionals, consumers, health care insurers, and the government (Hurst, 1991; Okma, 2001).) The implementation of policy measures may only be effective if implementers and target populations consider the proposed policy measure to be meaningful (Grin & van de Graaf, 1996a). The latter implies that the proposed policy 1) must make sense in light of their perception of the problem and 2) does not violate their normative values.

In conclusion, HTA evaluates interventions in order to support policy making. Current HTA methods are adequate for solving well-structured problems. However, when stakeholders interpret findings and situations differently, problems are likely to be ill-structured. HTA currently does not adequately assess the underlying theoretical assumptions and normative values of the policy makers and stakeholders that lead to diverging views of the problems and research. If this is true, more emphasis should be put on the possible differences in problem definitions and the underlying arguments of actors involved for improving the usability of HTA. Instead of exerting the greatest effort towards the implementation of research results after the completion of a trial, the relevance of the research to both policy making and clinical practice should be stressed before initiating the research.

This kind of 'problem oriented approach' in HTA resembles, to some extent, the new view of Technology Assessment presented by Smits & Leyten (1991). They argued that more attention should be placed on the definitions of the problem as viewed by the actors involved. In accordance with Lehoux & Blume (2000), this approach might also elaborate on the need to include a socio-political aspect in HTAs. By assessing the needs and beliefs of the actors involved, an HTA can provides an opportunity to include issues that go beyond the mere effectiveness and cost of an intervention. The interactive nature of the approach may provide an opportunity to include interests and perspectives in HTA. This problem-based approach to increasing the applicability of HTA to decision making also coincides with the Health Council's directive to emphasise the optimisation of patient care rather than focussing on the implementation of evidence (Gezondheidsraad, 2000).

In this thesis, I will further explore the hypothesis that it is imperative to consider the stakeholders' perspectives, even when dealing with technologies that initially appear to be less controversial. If this is true, it is necessary to perform an extensive process of problem structuring in which the stakeholders' perspectives in HTA are taken into account. The main questions to be answered are:

1 In cases where HTA research or the implementation of subsequent policy measures failed, were the research findings or policy measures congruent with the views of all actors involved?
   a Which actors were involved (policy makers, researchers, policy's target populations), how did they define the problem, and which theoretical assumptions and normative values were held by them?
   b To what extent are research findings or subsequent policy measures congruent with the views of actors involved (fit actors' problem definitions and not conflict with their normative values)?
2 How does involving the policy's target population's influence the design of policy research? Does their participation result in research findings or subsequent policy measures that are congruent with the views of all actors involved?
3 What are the outcomes of an interactive process of problem structuring in policy research, whereby the problem definitions, underlying assumptions, and norms are analyzed and discussed? How does this impact the interventions themselves, the criteria for success, and the research questions considered relevant?

Outline of this thesis

Reconstructing interpretive frames

The methodology of reconstructing actors' interpretive frames will be used for describing and reconstructing the actors' problem definitions and underlying arguments (Grin & van de Graaf, 1996b; Grin et al, 1997). Reconstructed interpretive frames will be used to assess 1) the type of problem (well, moderately, or ill structured) and 2) whether interventions are congruent with the interpretive frames of actors involved.

In Chapter 2, the methodology of reconstructing interpretive frames is discussed in more detail. Firstly, relevant validity criteria are discussed. Secondly, the results of an analysis of the inter-observer variability of reconstructing the interpretive frames are presented. Research questions are: Is reconstructing interpretive frames a reliable method to assess problem definitions, underlying theoretical assumptions and normative values? Is reconstructing interpretive frames a reliable method to assess the cooperation of actors on predefined solutions?

Retrospective analyses of failed research or subsequent policy measures

In Chapters 3, 4, and 5, case studies are presented in which research projects failed. Failed research includes those situations in which 1) research results are considered irrelevant by policy makers or the commissioner, or 2) the implementation of subsequent policy measures failed. Failure of subsequent policy measures is considered relevant, because HTA can be regarded as a kind of policy development. The case studies consisted of research projects that were commissioned by the Department of Policy Analysis of Medicines (PAM) of the Dutch Health Insurance Board, an advisory board for the Ministry of Health. Within the Board, PAM contributes to identifying developments that may jeopardise optimal medical care, analysing the nature and size of such threats, and ensuring that additional research that may provide a basis for resolution through policy is conducted (Ziekenfondsraad, 1998).

In Chapter 3, a case study on the drug mebeverine for patients with the irritable bowel syndrome is presented. Prior to this project, measures to exclude the drug from the health package had failed. Actors involved (policy analysts, researchers, and policy's target populations) disagreed whether the drug fulfilled the criteria to qualify for reimbursement. Research was expected to provide the lacking knowledge. The results of a preliminary study, however, were considered not useful by its commissioner. No further research was commissioned and no policy measures were implemented. Questions are: can differences be found between researchers and potential users, policy makers and professionals in the field in the definition of the problem and which norms are at stake? Are research results congruent with interpretive frames of actors involved?

Chapter 4 describes a case study in which the implementation of policy measures based on research results failed. The subject was the implementation of a

national protocol for prescribing lamotrigine, an antiepileptic drug. The recommendations in this guideline were based on the available evidence from clinical trials at that time. An inquiry had shown that this guideline was hardly used in clinical practice. In this chapter, the theory of argumentative policy analysis is used to explain the proceedings of the project. Question is whether the policy measure was congruent with interpretive frames of policy maker and members of the target population.

In Chapter 5, a project in which target populations were invited to participate in policy development is described. Two workshops were arranged which aimed to advice on policy measures to improve the efficient use of the drug interferon-beta for patients with Multiple Sclerosis. Research question to be answered is, does participation of target populations in policy development result in policy measures that are considered as solutions to problems they perceive and do not conflict with their normative values?

Monitoring case studies in which lessons are incorporated

Based on the findings in the abovementioned cases, changes were made in the processes of commissioning research by the Health Care Insurance Board. Aim was to take into account stakeholders' perspectives and prevent a shift in problem definition. In chapter 6 and 7, two case studies are presented in which these lessons were incorporated.

The first case study (Chapter 6) deals with the project on new analgesics, the cox-2 selective inhibitors. In this study, researchers were asked to provide 1) information on target populations perspectives and 2) evidence on the value of specific interventions in one project. We monitored the process of policy analysis conducted by PAM staff, the European tender for proposals, and the PAM appraisal of the research proposals that were submitted. Research question is, is the new approach to commissioning research successful in terms of increasing the cooperation of researchers and the relevance of the results?

The second case (Chapter 7) deals with the efficient use of two, novel drugs for patients with pulmonary hypertension. The PAM staff asked researchers to make an inventory of problems and judgements of solutions according to the policy's target populations. The PAM staff decided to commission the project on perceived problems separately from a research project aimed at obtaining specific information needed for

policy making. Furthermore, it allowed them to commission the projects via a non-European tender. This provided flexibility in the organisation of the tendering process and allowed for meetings between the commissioner and researchers. Research questions are: Did research proposals correspond with the problems as perceived by both those who commissioned the study and the stakeholders in the field? Were the research results translated into the implementation of concrete policy measures?

Interactive process of problem structuring

Finally, a promising methodology for problem structuring is the methodology of Interactive Technology Assessment, which is based on the methodology for fourth generation evaluation developed by Guba and Lincoln (1989). In Chapter 8, a case study is presented in which an interactive methodology is used for structuring problems. The case study dealt with acute hospital care for patients who attempted suicide by intoxication. This study aims to compare the outcome of an interactive approach for problem structuring with the outcome from a conventional approach. Research questions is: What problem definition emerges from a fourth generation approach to problem structuring, as compared to the problem definition as presented by the person who first put it on the agenda?

In Chapter 9, the findings from the case studies results will be summarised and conclusions will be drawn on the need for a problem based approach in HTA.

2

Validity and reliability of qualitative data-analysis: inter-observer agreement in reconstructing interpretive frames

Many authors have discussed criteria for assessing the quality of qualitative studies. However, relatively few authors present the results of employing criteria for validity of qualitative studies. We investigated the quality of reconstructing interpretive frames, a method for analysing interview transcripts. Aim of this method is to describe a person's perspective, distinguishing between perceived problem definitions, proposed solutions, empirical background theories, and normative preferences. Based on this description one should be able to estimate this person's cooperation on implementing specific changes in his or her practice. In this article, we assessed the inter-observer reliability of this analytical method as an indicator of its rigor. Six analysts reconstructed interpretive frames on the basis of verbatim transcripts of three interviews. As to the issues identified and the question which problems should be prioritized, the analysts only moderately agreed. However, as to the estimates of the respondents' cooperation on proposed solutions, the analysts showed remarkable unanimity.

M Moret, RPB Reuzel, van der Wilt GJ, J Grin. Validity and reliability of qualitative data-analysis: reconstructing interpretative frames. Field Methods 2007; 19(1):24-39

Most articles and books on the "quality" ("validity", "credibility", or "rigor") of qualitative research deal with the question how to assess the quality of a study that has been performed. Rarely, however, the results of employing these criteria are published (Barker, 2003; Clavarino et al., 1995). Moreover, there is discussion as to which criteria should be used to assess the "quality" of a qualitative study. Commonly, the discussion centers on the concept of truth and the question whether truth is (a) universal or local, and (b) determinable. According to many authors, criteria used for quantitative research are also applicable in qualitative research, which is to say that validity and reliability are meaningful concepts in qualitative research.

Discussing validity is not only important for estimating the trustworthiness of research findings, but also for scrutinizing the aims and scope of the methods used. In this sense, discussing validity is an instrument for improving methodology.

Validity is context-bound, however (Yanow, 2000). That is, it depends on the aims of a method and the context in which this method is employed. For example, it is well known that (western) methods for assessing health related quality-of-life are not valid in many African countries (Mkoka et al., 2003.)

Furthermore, it is important to acknowledge that qualitative research and quantitative research do not exclude each other. It is more useful to view both as approaches, which in practice may involve the employment of several different methods for data collection and analysis, some qualitative, some quantitative. Each method features its own definition of reliability. If a question is quantitative in nature, it is perfectly appropriate to use quantitative approaches, even when the subject of study is a qualitative analytical method.

In this paper, we address a method for analysing qualitative data, i.e. the reconstruction of interpretive frames. This analytical method is used within the context of so-called 'fourth generation' approaches to evaluation. It allows for eliciting stakeholders' views, in order to estimate the likelihood that these stakeholders co-operate on a set of proposed solutions, or policy interventions. Our aim is (1) to explain how validity and reliability are defined in the context in which the method is employed, (2) explain why validity and reliability are important in this context, and (3) to demonstrate how reliability can be assessed.

Validity in fourth generation approaches

Unlike researchers who are quantitatively oriented, many 'qualitative researchers' would claim that they are not interested in the truth. Rather they would inquire into a respondent's version of the truth. Still, qualitative research aims at knowledge. That is, qualitative research is still defined as a scientific endeavour that is successful if in the end it produces knowledge that is broadly accepted, even if truth is considered a local concept. It is at this point, that so-called 'fourth generation' approaches to qualitative research mark a difference. According to these approaches knowledge should not be considered as an end-point of inquiry. Instead, action (e.g. policy recommendations) should. To be sure, knowledge is important as a sound basis for action. Consequently, knowledge claims should be meticulously scrutinized, but they primarily serve deliberation processes, which should culminate in action, or change. This has important consequences for the concept of validity involved.

Guba and Lincoln (1989), inventors and two advocates of the fourth generation evaluation approach, view evaluation as a procedure 'in which a combination is made of responsive focusing (using the claims, concerns and issues of stakeholders as the organizing elements) and a constructivist methodology (which aims to develop consensus among stakeholders who earlier held different or conflicting constructions).' (p. 71) Central in their methodology is the hermeneutic dialectic process. It consists of one or more rounds of open-ended interviews with stakeholders. It starts with an interview with a first respondent to determine his or her construction of the investigated phenomenon. Next, the researcher interviews a second respondent to determine his or her construction. Furthermore, the researcher confronts the second respondent with claims, concerns, and issues raised in the interview with the first respondent. The interviewer then makes a shared construction based on these two interviews. Then a third respondent is interviewed, etc. Ideally, this process proceeds until no new information is added. In the view of Guba and Lincoln, the aim is to reach consensus.

Obviously, 'traditional' criteria (internal validity, external validity, and reliability) are not useful in this approach. Reproducibility is considered irrelevant, because in qualitative research the researcher is commonly interested in practices that are strongly bound to a specific context (including time and place). Similarly, the fourth generation researcher would not aim at generalisability. On the contrary, fourth generation evaluation should produce change in order to provide solutions for

problems conceived in a specific context. Thus, in fourth generation evaluation, reliability of interviews looses in significance, for if evaluation aims at action, rather than knowledge, then reliability in the sense of being researcher independent and yielding the same results upon repeated measurements is not only futile, but even undesirable. It is therefore that Guba and Lincoln prefer to use 'dependability' instead.

However, Guba and Lincoln have derived their criteria from the aims of fourth generation evaluation as a whole. The question remains whether these criteria apply to methods, e.g. methods for analyzing interviews, used within the process. Could concepts of validity and reliability be meaningful there? One could argue that if the analyst has difficulties with interpreting an interview, several interviews should be scheduled to adjust the interpretation until analyst and respondent agree to an interpretation that covers the respondent's version of the truth. However, we would argue that in fourth generation evaluation it is unwise to exclusively rely on the self-correcting mechanism of the hermeneutic process, if only for reasons of efficiency. Constraints on time and money call for an analytic tool that makes it possible to interpret someone's ideas in a sound way. It is at this level, then, that we do believe that reliability remains important.

Reconstructing interpretive frames

Reconstructing interpretive frames is one such method for analyzing interviews. The term 'interpretive frame' is used by Grin and van de Graaf (1996B; based on a synthesis of Schön, 1983, and Fischer, 1980) to refer to a quadruple set of elements that determine a respondent's view: context-specific problem definitions, solutions, empirical and ethical background theories, and normative preferences. Grin and van de Graaf argue that the 'second order notions' of background theories and normative preferences span the space within which problems are defined, and solutions sought. This adds some precision to understanding the process initiated in fourth generation evaluation. Careful analysis of interviews in terms of interpretive frames helps 'at the level of knowledge' sorting out what is agreed and disagreed upon, and thus helps preparing subsequent interviews. But reconstructing interpretive frames is even more useful for designing widely endorsed solutions to problems encountered, and estimating the likelihood that participants agree to these solutions. The idea is that cooperation on the implementation of policy measures depends on whether

stakeholders consider these policy measures meaningful from their own interpretive frame. A measure is considered meaningful by a particular stakeholder, if it solves his or her problems and does not violate his or her background theories and normative preferences. Thus, in designing policy measures it is relevant to identify actors involved and their interpretive frames. Clearly, the method fits in Guba and Lincoln's fourth generation approach, which similarly aims at agreement over policy measures.

Until now, little has been published on the validity and reliability of this method or comparable methods (Grin et al., 1997). To assess the inter-observer reliability of the method, we aimed at answering the following research questions:

1. To what extent do different analysts agree on (a) issues identified and (b) the most important problem definitions for each respondent?
2. To what extent do different analysts agree as to whether respondents would cooperate on a set of proposed solutions?

Methods

We used the transcripts of three interviews from a fourth generation evaluation of cochlear implantation (CI) in deaf children (Reuzel, 2004). A cochlear implant, or 'bionic ear', is a device that provides a hearing sensation to profoundly deaf people. Sounds from the environment are transduced by a microphone, processed by a so-called 'speech processor', and then transferred to the acoustic nerve through electrodes. Surgery is required to implant the receiver coil and connect the electrodes. Through extensive rehabilitation, recipients can learn to interpret the auditory input they receive. Although the technology is effective in most individuals, the technology has raised considerable controversy for its development, implementation, and evaluation have been primarily based on a medical perspective on deafness as a handicap to be eradicated. Seen from this perspective, cochlear implantation helps to ensure that deaf subjects are integrated into the 'hearing society' as much as possible. However, advocates of cochlear implantation, who have been responsible for the development and evaluation of the technology, have largely neglected Deaf concerns about the sustainability of Deaf culture and the social and emotional development of deaf children. These concerns are associated with an alternative view on deafness referring 'to socio-cultural characteristics of those hearing-impaired persons who

consider themselves to belong to a special (Deaf) community' (Tellings, 1995: 21). The perspective on deafness as a handicap is thought of as a threat to this community. Furthermore, deaf children would be in danger of experiencing social and emotional pressures, due to discrimination and high expectations, the effects of which could be serious and lasting.

A fourth generation evaluation was undertaken (Reuzel, 2004) to identify the conditions under which implementation of CI might be effective and acceptable. Moreover, it was felt that the evaluation perhaps could restore the severely deteriorated mutual trust between advocates and opponents of CI. This fourth generation evaluation was, in fact, a response to the claim of many opponents that not only CI, but also the health technology assessment studies undertaken to support policy decisions on it, were dominated by a conventional medical rationality. The project involved of a series of open-ended interviews with 51 different stakeholders. Among the most important issues that came up was communication, particularly the question whether a deaf child wearing a cochlear implant should be raised and educated using oral language, Sign language, or a combination of both. It is this issue that we have emphasized in assessing the validity of reconstructing interpretive frames.

Reconstructing interpretive frames

Using the verbatim interview transcripts, six analysts employed within our department have reconstructed the interpretive frames of three respondents. One of them (Reuzel) was the interviewer. The other analysts were familiar with evaluation studies in health care, but inexperienced in reconstructing interpretive frames, and only superficially acquainted with the CI problematic. They followed a short training (three half days, by Grin) on the theory of reconstructing interpretive frames. Moreover, they were provided with a protocol with references to Miles and Huberman (1994) on the coding of texts, and Grin et al. (1997) on the method of reconstructing interpretive frames.

Next, all analysts independently analyzed the interview transcripts, coding for the issues encountered and the four elements of the interpretive frame. As to the coding of issues, no agreements had been made beforehand. After analyzing the first transcript, the results were discussed between the analysts and John Grin. After analyzing the second and third transcripts, the analysts filled out a questionnaire,

prompting them to note the three most relevant problems for each respondent and predict the cooperation of each respondent on a set of predefined solutions. Answer categories included 'yes', 'no', or 'indeterminate'.

Three solutions were based on the problem of exposition to oral language in schools for the deaf or hard of hearing usually being insufficient for optimal use of CI:

1. Children with CI should attend mainstream schools as soon as possible,
2. Ambulatory services should be increased.
3. Regional expertise centers in which schools for the deaf, schools for the hard of hearing, and mainstream schools co-operate are necessary.

Two solutions referred to the fact that communication between hearing and deaf persons is commonly problematic, particularly in the first years of life:

1. Simultaneously offering sign language and oral language is necessary to enable early communication with deaf children.
2. Parents of deaf children should learn sign language in order to better communicate with their children.

We also noted what arguments for or against cooperation the analysts ascribed to the respondents and to which elements of the respondents' interpretive frames these arguments referred. Finally, in order to further interpret inter-analyst similarities and differences, we discussed the results during two 2-hour sessions.

Analysis

We compared the reconstructions of the analysts as to the following items:

1. The issues identified in the first interview
2. The three most important problems for each respondent
3. A prediction of the cooperation of each respondent on the five proposed solutions.

A qualitative description was given of the agreement on issues identified and the lists of three most important problems.

In order to assess the degree of agreement between analysts about the respondents' cooperation on proposed solutions, reliability coefficients were calculated. A quantitative calculation is possible because the number of outcomes is limited (three options) and have an order in ranking. One reliability coefficient was calculated for each respondent. The calculation method has been described in detail

by Streiner and Norman (1991), and is based on the analysis of variance. The overall variability in answers could be attributed to three sources of variability: variability attributable to the analysts, variability attributable to the interventions, and residual variability.

We replaced the answers (yes, indeterminate, no) by numeric values. We valued a 'yes, the respondent will cooperate' as '1', 'indeterminate' as '0', and 'no' as '-1'. We assumed that 1) an 'indeterminate' is situated between a 'yes' and a 'no'; and 2) the difference in scores between 'yes' and 'indeterminate' is equal to the difference in scores between 'indeterminate' and 'no'. On these assumptions, the outcome is independent from the selected values.

The overall variability is based on deviations between the observed values and the average value of all observations (= expected value). The variability attributable to analysts is based on the deviation between the average value for each analysts and the average value of all observations. The variability attributable to interventions is based on the deviation between the average value for each intervention and the average value of all observations.

To calculate the reliability coefficient, the variability attributable to the interventions is divided by the overall variability (variability attributable to the analysts, variability attributable to the interventions, and residual variability). The reliability is expressed as a number between 0 and 1, 0 indicating 'no reliability' and 1 indicating 'perfect reliability'. If all analysts agree, the variability attributable to the analysts is low and the coefficient is high.

Results

Five analysts reconstructed the interpretive frames for the interview with respondent A (R, W, G, M, & J). Four analysts reconstructed interpretive frames on the basis of interview transcripts pertaining to respondents B and C (R, W, G, & M). Five analysts completed the questionnaire for respondents A and B (R, W, G, M, & H). Four analysts completed the questionnaire for respondent C (R, W, G, & M).

Table 2.1. Relevant issues in the interview with respondent A, identified by the 4 analysts

identified issues	R	W	J	M
language	x	x	x	x
education	x	x	x	x
counseling of the child	x	x	x	x
indication for CI	x	x	x	x
education / counseling of parents	x		x	x
funding	x			x
identity	x			
pressure of work (counselor)	x			
role (counselor)	x		x	
social well being / communication	x			
equipment	x	x		
discussion around CI	x			
use of protocol	x			
CI in general			x	x

Identified issues

The analysts identified the following issues: language, education, counseling of the child, and eligibility criteria for CI (Table 2.1). The coding differed between analysts as to the measure of detail. For example, issues that were coded as "CI in general" by two researchers were identified by the other analysts as issues concerning the role of the counselor, social well-being, and identity. Not surprisingly, the analyst who had held the interviews (R) made the most detailed list of issues. Moreover, the analysts appeared to attach different meanings to concepts such as communication.

Most important problems

As to interview A, all analysts identified the problem that 'language input is generally insufficient for deaf children' (Table 2.2). The lists for interview B and C, however, were more diverse (Tables 2.3 & 2.4). Regarding interview B, three out of five analysts mentioned the problem of gearing language input to the capabilities of individual deaf persons. The problem that oral language alone is insufficient for a healthy social emotional development was also mentioned frequently. On the basis of interview C, two out of four analysts identified the problem that the medical perspective is

dominant in information for parents. Two analysts mentioned that deafness is more than not hearing, and that the social emotional development of deaf children is at stake. All other problem definitions were mentioned by one analyst only.

Table 2.2. Three most important problem definitions for respondent A according to the five analysts

	R	W	G	M	H
language input is generally insufficient for deaf children	x	x	x	x	x
information on deafness and CI is not objective	x				
value of CI for an individual is difficult to assess	x				
conditions for integration in regular education		x			
indication: who decides who are qualified for CI		x		x	
children cannot use CI directly				x	
CI not available for everyone due to too less financial sources					x
education should be geared to one another					x

* CI =Cochlear Implants

Table 2.3. Three most important problem definitions for respondent B according to the five analysts

	R	W	G	M	H
Gearing language input with possibilities of individual children; an individual approach in education is preferable	x	x		x	
Oral language is insufficient for social emotional development		x		x	x
Insufficient nuance in information/ expectations of parents	x				
It is not clear what is the best way to communicate		x			
Unclear how to obtain the maximum with right mix van bi-lingualism?			x		
Deaf people do not respect parents choice for CI*				x	
At schools no agreement on using oral language, sign language or both					x
Regular education whereas deaf education seems more suitable					x

* CI =Cochlear Implants

Expected co-operation on proposed solutions

Analysts were asked to consider 15 (=3x5) combinations of a respondent and proposed solution. In eight combinations, all analysts agreed about the likelihood that the respondent would cooperate on the proposed solution (Table 2.5). In one situation analysts fully disagreed. In the remaining six situations, some small differences were found. Most differences were found in the predicted cooperation of respondent A. The calculated reliability ratio was 0.73. The analysts disagreed on the cooperation on

the last three solutions. All analysts referred to respondent's background theories, but interpreted these differently. Some analysts indicated that information was missing. Apparently, the respondent's opinion was nuanced.

Table 2.4. Three most important problem definitions for respondent C according to the four analysts

	R	W	G	M
Adequate information is necessary; medical perspective in information dominant	x			x
Deafness is more than not hearing; culture and identity; acceptance of deafness by hearing people and deaf	x		x	
(Uncertainty on) social emotional development of children	x	x		
Deaf have language deficiency		x		
Self image of deaf is indistinct		x		
Deaf always push oneself to he limit			x	
Consequences of CI* are indistinct				x
People think that input of sign language suffice				x

Fewer differences were found in the predicted cooperation of respondent B and C. For both respondents, reliability ratios were 0.98. In most situations the analysts agreed and mentioned similar arguments, referring to respondent's background theories. Furthermore, the answer 'indeterminate' was given less often. In three situations, some analysts answered that the issue that the proposed solution referred to was not addressed in the interview, whereas others based their judgments on background theories. Finally, all analysts agreed that co-operation on the fourth solution was to be expected of all respondents.

Discussion

In this article, we presented the results of the reconstruction of interpretive frames by six analysts. We found that the analysts only moderately agreed on issues identified and the most important problems for the respondents. Texts coded as 'CI in general' by some analysts were coded in greater detail by others, the analyst who had held the interviews achieving the greatest detail. From the discussion between the team members, it became clear that the identification of issues depends on prior knowledge from the analyst. For example, the analyst who held the interviews probably recognized issues that had been dealt with in other interviews. Another

possible explanation is that this analyst was more experienced in reconstructing interpretive frames.

Table 2.5. Estimated cooperation of the three respondents on five proposed solutions by analyst *

	respondent A					respondent B					respondent C			
	R	**W**	**G**	**M**	**H**	**R**	**W**	**G**	**M**	**H**	**R**	**W**	**G**	**M**
1. Children with CI** should participate in regular education as soon as possible.	+	+	+	+	+	-	-	-	-	-	-	?	-	-
2. The capacity of ambulatory counselors should be increased to optimize the input of oral language at school and at home.	+	?	+	+	+	?	?	+	?	+	-	?	-	?
3. Regional expertise centers in which schools for the deaf, schools for the hard of hearing, and regular schools co-operate are necessary.	?	-	?	?	?	-	-	-	-	-	?	?	?	?
4. Offering sign language besides oral language is necessary to facilitate communication with the child.	+	+	?	?	?	+	+	+	+	+	+	+	+	+
5. Parents should learn sign language so that they can better communicate with their children.	+	?	?	-	-	+	+	+	+	+	+	+	+	+

* '+' = yes, it is likely that the respondent will cooperate on the proposed solution according to the analyst, '-' = no, the respondent will not cooperate according to the analyst, and '?' indicates that the analyst could not ascertain whether the respondent would cooperate.
** CI =Cochlear Implants

Other authors found a high agreement in coding of text between analysts (Clavarino et al., 1995). Ryan (1999) also reported on a comparison between multiple coders. However, the aim of this analysis was not to assess the validity or reliability of the coding method. The author argued that comparison of multiple analysts could also be used to identify core and peripheral issues from a text. He emphasized text fragments that are coded by several analysts. An explanation for the high agreement claimed on the basis of these studies might be that agreements regarding the codes to be used had been made before analysis. In our study, we only agreed that two kinds of codes should be used: one concerning the four elements of the interpretive frame and one concerning the issues discussed in the interview. No agreements were made on the coding of these issues.

Agreement on cooperation

Although it is not always appropriate to use quantitative measures in qualitative research, we consider quantitative measures appropriate in our analysis. For instance, results from qualitative research need not always be replicable, in case of which quantification is inappropriate. However, we asked to what extent analysts agree on the likelihood that a respondent is willing to co-operate on a set of proposed solutions, which involves a quantitative question. We found that the agreement between the analysts as to the estimated cooperation of actors B and C on the proposed solutions was fairly high. Streiner and Norman consider a score of 0.75 as a minimal requirement for a useful instrument (Streiner & Norman, 1991).

Analysts moderately agreed on the co-operation of respondent A. In the calculation of reliability coefficients, we have accounted for three sources of variability: variability attributable to the analysts, variability due to differences between the proposed policy measures, and residual variability. The extent of agreement found with respect to respondent A may result from differences between the analysts or from residual variability. In our data, the low agreement resulted from a relatively high residual variability (data not shown). Inadequate interviewing may affect the reliability assessments of expected co-operation on policy measures. Poor interviewing (on the part of either the interviewer or the respondent) could leave a respondent's view unclear or obscure, as a result of which analysts interpret it differently and the residual variability increases. The limited experiences of the analysts could also provide an explanation for the relatively high residual variability, in case of which the relatively high levels of agreement found for respondents B and C indicate a learning effect. Finally, in the questionnaire, solutions were introduced by a short description of a problem to provide the analysts with some context. Sometimes an analyst remarked that the problem appeared not to be acknowledged by the respondent. This might also affect the analyst's prediction. It might be better to present solutions alone.

It is difficult to tell which explanation accounts for the extent of agreement found. However, in any case, our calculation of the extent of agreement indicates that there is a problem concerning validity and reliability. Similarly, a high degree of agreement may be interpreted as an indicator of good reliability.

Mostly, when analysts estimated a respondent's cooperation, they referred to similar arguments, which often derive from the respondent's background theories.

Sometimes, analysts would be cautious and answer that an issue was not addressed in the interview, while others would not hesitate to predict cooperation on the basis of the respondent's background theories. As we noted, the theoretical idea is that cooperation is likely if 1) the intervention is conceived of as a solution to his or her problems, and 2) the solution does not conflict with second order notions. Apparently, it was not always easy to unambiguously assess whether a solution that is not explicitly discussed will conflict with someone's second order notions. Here again, our discussions indicated that the amount of prior knowledge could explain much of the inter-analyst variation.

Limitations our study and steps for further inquiry

A limitation of our study was that we restricted our analysis to interview transcripts. In fact, reconstructing interpretive frames involves both interviewing and analysing the interviews. It is plausible to think that if all analysts had been involved in data collection, we would have found more agreement. It would be interesting to compare outcomes of the whole process of interviewing and reconstructing interpretive frames. However, this is more complicated to arrange in practice.

Furthermore, we assessed the agreement between analysts in their estimates of respondents' co-operation on a set of proposed solutions. Even when all analysts agree on the cooperation of a respondent, the question remains whether the actor will actually cooperate. A high agreement between different analysts is not enough to establish the validity of the interpretation of data. It is possible that all analysts were wrong about the interpretation. Moreover, in reality, it is possible that the respondent changes his or her perspective. Guba and Lincoln acknowledge changes in the perspectives of the respondents as one of the hallmarks of quality. In fourth generation evaluation, the results of the data-analyses do not mark the end stage of a research project, but are submitted to the respondents in a continuous process. The next step would be to monitor whether the respondents indeed co-operate on the proposed solutions when these are implemented.

Conclusion

The method of reconstructing interpretive frames from verbatim transcribed interviews seems reliable to assess the cooperation of respondents on proposed solutions. Thus, the method is useful for identifying widely endorsed policy

interventions. Nevertheless, in some situations it might be difficult to assess cooperation on issues that were not discussed in the interview. The method is less reliable when it comes to identifying relevant issues raised by the respondents and the most important problem definitions, at least with analysts not involved in data collection. When more than one analyst cooperates in one project, it might be advisable to make agreements on the coding of issues. In particular, a shared understanding of the problematic and an associate coding system might significantly improve reliability.

3

Health Technology Assessment and ill-structured problems: a case study concerning the drug mebeverine

The practical significance of Health Technology Assessment (HTA) in policy decisions or clinical practice has been challenged. Possibly, problem definitions underlying HTA do not concur sufficiently with the problem definitions held by policy makers or clinicians. We performed an in-depth case study on mebeverine, a drug prescribed to patients with irritable bowel syndrome to explore this hypothesis.

Methods: Theoretical framework was provided by the theory of argumentative policy analysis. We analysed documents and held semi-structured interviews to collect data. We reconstructed interpretive frames to analyse actors' argumentation.

Results: The funding and usage problems relating to mebeverine were ill-structured. Actors disagreed on the information needed and the norms at stake. As a result, the problem definition shifted and the resulting problem definitions failed to correspond with the problems perceived by the target populations.

Conclusions: To ensure that future studies on health care problems are useful, it is imperative that policy makers take the problem definitions of potential users into account.

M Moret-Hartman, GJ van der Wilt, J Grin. Health Technology Assessment and ill structured problems: a case study concerning the drug mebeverine. International Journal of Technology Assessment in Health Care 2007; 23(3):316-23

Health Technology Assessment (HTA) generally aims to support health care policy-making. (Banta, 2003) However, the actual contribution of HTA to the policy making process has been questioned. (van den Heuvel et al., 1997; Rosen & Gabbay, 1999) It is possible that HTA provides its users (policy makers) with insufficient insights on the considerations and life worlds of the target populations that will be affected by potential policy measures. Such insights are needed because, in health care, policy has shifted from central regulation to an approach based on target populations. The focus is currently placed on influencing the doings and dealings of patients, physicians, and other target populations (Chernichovsky, 1995; Frankish et al., 2002; Hurst, 1991; Malcolm, 1989; Okma, 2001; Schieber, 1995). Policy problems can be formulated differently by the actors involved because information that surrounds these problems can be interpreted differently (Dunn, 2004).

Literature on argumentative policy analysis (Fischer & Forester, 1993; Hoppe & Peterse, 1998; van de Graaf & Grin, 1999) contends that in order to acquire the cooperation of target populations (crucial for effective policies), the perspectives of the target population must be taken into account during the development of policy. For instance, with respect to policies aimed at promoting sustainable development, target populations co-operate if, and only if, they consider the proposed policy measures to be meaningful. A measure is considered meaningful when it corresponds with theirs problem definition and does not conflict with their background theories or preferences (Grin & van de Graaf, 1996a; van de Graaf & Grin, 1999). Consequently, policy research should take the problem definitions of its users and target populations and the meaning they attribute to potential policy measures into account.

We examined a case of policy research on mebeverine, a drug frequently prescribed to patients with irritable bowel syndrome (IBS). Patients with IBS suffer from diarrhoea, constipation or spasms of the gastrointestinal tract for which no structural or biochemical cause can be found. The drug is thought to affect smooth muscle cells in the colon and is thus expected to relax the muscles and thereby decrease spasms (Connell, 1965; Greenwood & Mandel, 1992; Poynard et al., 1994). Although the drug is frequently prescribed, there is discussion regarding its effectiveness. (Klein, 1988) The uncertainty about the effectiveness of mebeverine has also been a relevant issue in Dutch reimbursement decisions. In particular,

mebeverine seemed to be a drug that could easily be removed from the public health care package. After failed policy, new research was proposed to inform future policy measurements on this subject. Unfortunately, the commissioner considered the outcome of a preliminary study not useful for further policy measures. As a result, no further policy measures were taken.

The aim of the present study was to explore the extent to which the failures in policy and subsequent policy research may have resulted from incompatible problem definitions among various stakeholders. The research questions were:

- How did the process of policy analysis and subsequent research proceed? Who were involved and how did they define the problem?
- Did the study incorporate problems perceived by target populations (physicians and patients) with respect to policy measures?

Methods

The theoretical framework employed in this study was argumentative policy analysis. This theory contends that successful implementation of policy requires the cooperation of the target populations. Target populations are: actors who are likely to experience the consequences of the intended policy and actors whose cooperation may be necessary for successful implementation. Cooperation from these target populations is more likely if they consider the proposed measures meaningful. This means that they expect the measures to provide a solution to problems they perceive.

We used the method of reconstructing interpretive frames to analyse target populations' argumentation (Grin et al., 1997). One's interpretive frame is comprised of problem definitions and judgement of possible solutions but also the empirical and normative background theories that shape them (Grin & van de Graaf, 1996b; Schön, 1983).

We analysed documents and held semi-structured interviews to collect data. Firstly, we analysed the process of policy analysis and the subsequent research project retrospectively. Data were obtained from the file on this project in the archive of the Heath Care Insurance Board. The file contained letters, reports from the Board, internal memos on this subject, research proposals and a research report. Secondly, semi-structured interviews were held to reconstruct actors' perspectives. Interviews were held with policy makers, researchers, physicians, and patients. We prepared for

interviews by conducting a detailed analysis of the existing literature on mebeverine and irritable bowel syndrome. In the interviews, the questions focussed on perceived problems and reasons for actions or decisions. The actors involved in the project (two policy makers from the Health Care Insurance Board (HCIB) and two researchers) and a policy maker from the Ministry of Health were asked provide reasons for the specific choices that were made during the process of policy analysis. Members of target populations (two general practitioners, one gastroenterologist, and two patients, who were founders of the IBS patient association) were asked about their use of mebeverine, their reasons for using this treatment, alternative treatment strategies, ideas about IBS, and perceived problems relating to the care for IBS patients.

Interviews were transcribed verbatim and coded. Four layers of one's interpretive frame and relevant subjects were distinguished. The interview transcripts and a conceptual version of the report were sent to respondents for comment (respondent validation).

Results

In this section, we, firstly, describe the proceedings of the policy analysis. A summary of the problem definitions and possible solutions at several time intervals is also provided in Table 3.1 Secondly, we compare policy makers' perspectives with the perspectives of target populations.

Reconstruction of policy analysis proceedings

In 1996, drugs in the Dutch public health care package were screened on the basis of the need for and effectiveness in order to develop a high quality and affordable health care package. The Ministry of Health asked the Sickness Funds Council (currently called the Health Care Insurance Board) for advice on several drugs, including mebeverine. The Ministry did not consider mebeverine to be very effective. Other interventions, such as a dietary advice and reassurance, were considered more efficient. In response to the Ministry's request for advice on the reimbursement of drugs, the Sickness Funds Council developed a decision model (Ziekenfondsraad, 1995). The criteria for assessment included the efficacy, effectiveness, therapeutic value, and efficiency of the drug in question (Table 3.2).

Table 3.1.Problem definitions and proposed solutions

	problem definition	solution
Advice Sickness Funds Council (September 1995)	Effectiveness of mebeverine is disappointing	Strict administration of the protocol should lead to exclusion from the health care package; however, because of symptoms and lack of alternatives, it should be reimbursed until further notice.
Ministry of health (December 1995)	mebeverine has little therapeutic value; other interventions more efficient	mebeverine should be excluded from reimbursement
Policy document PAM* (March 1999)	Mebeverine is reimbursed, however, still questions on its efficacy or effectiveness	Research on the effectiveness of mebeverine
Letter covering the PAM policy document (March 1999)	The industry made the court pass a sentence on the efficacy of mebeverine. However, efficacy does not lead to effectiveness automatically	Question on the effectiveness of mebeverine still needs to be answered
Policy document PAM (April 1999)	Unlikely that trial results will be useful to remover mebeverine from health care package.	Aim should be to act on prescribing patterns. Research on therapeutic value of mebeverine.
Proposal researchers (June 1999)	-	Preliminary study on the feasibility of a trial on the therapeutic value of mebeverine
Final report researchers (November 1999)	Evidence on IBS treatment is lacking	A trial on the efficacy of mebeverine and fibres is feasible
Policy document PAM (January 2001)		European tender of placebo-controlled trial on effect of mebeverine
Comment advisory committee PAM (January 2001)	Previous studies on the effectiveness of mebeverine were methodologically flawed	Impossible to prove effectiveness mebeverine in trial
Report Health Care Insurance Board (February 2002)		No randomised controlled trial on the effectiveness or therapeutic value of mebeverine

* PAM = department of Policy Analysis of Medicines

With respect to mebeverine, the Council concluded: a) that the efficacy of the drug was assessed when mebeverine was initially registered b) that convincing evidence on the drug's effectiveness was lacking; and c) that its therapeutic value is limited. Despite the apparent limitations, the Council recommended that the Ministry continue to reimburse mebeverine since the Council deemed it advisable to have at least some form of medicinal treatment and, of various alternatives, mebeverine was considered to have least amount of side effects (Ziekenfondsraad, 1995). The Ministry of Health rejected the advice provided and excluded the drug from the public health care package. This decision was challenged both by the medical profession and by the industry and the matter was taken to court. The court concluded that both the Ministry of Health and the industry acknowledged the lack of clear evidence on effectiveness but differ in the opinion whether mebeverine fulfilled the criteria for reimbursement.

Table 3.2.Reimbursement criteria (Ziekenfondsraad 1995)

Criterion	Description
efficacy	Its pharmacological action results in a therapeutic effect in clinical research (therapeutic potential)
effectiveness	Its use in clinical practice results in the aimed goal of the treatment
therapeutic value	The sum of its relevant characteristics (effectiveness, toxicity, user-friendliness, etc.) qualifying for its position relative to alternative therapeutic interventions
efficiency	A medicine is effective and the balance between therapeutic value and costs is favourable in comparison to other treatments

The court decided that mebeverine should be reimbursed because the interest in its continued use and funding far outweighed the Ministry's justification for withdrawing the drug from the health care package. Following this, the Ministry sought new means to substantiate their claim that mebeverine be removed from the health care package and thus proposed a new study that could hopefully generate evidence on the drug's lack of effectiveness

In 1999, this matter was adopted by the department of Department of Policy Analysis on Medicines (PAM) at the HCIB (Ziekenfondsraad, 1998). PAM employees performed a policy analysis and wrote a short proposal for the requested trial that was presented in a meeting with the PAM advisory committee. The committee, which consisted of various experts from the field, concluded that a placebo-controlled trial on the effectiveness of mebeverine might not be able to provide the relevant information needed for policy making. Because the Ministry of Health considered dietary advice more efficient, they contended that dietary advice should be included in the study. PAM employees then proposed to commission a trial on the therapeutic value of mebeverine compared to dietary advice. However, PAM employees were uncertain about the feasibility of such a trial. In June 1999, they proposed a preliminary study on the feasibility of a clinical trial on the therapeutic value of mebeverine in relation to dietary advice. This preliminary study aimed to assess the possibility of standardising dietary advice, determine which outcome measures could be considered clinically relevant and how many patients should be included.

In May 2000, researchers from two departments of general practitioners at university hospitals were commissioned to perform the requested preliminary study. The final report was presented in November 2000. According to the researchers, the major problem was the lack of evidence on the efficacy of any interventions for IBS

patients. Evidence from valid controlled trials was needed in order to develop a evidence-based guideline. Aim of their study was to assess determine the feasibility of the study. Their study included: the identification of an optimal outcome measure; the standardisation of dietary advice; and some specific IBS related problems, such as the inclusion of the relevant spectrum of patients and the identification of subgroups. To get an impression of the effectiveness of usual care, researchers reviewed the literature on dietary advice, held interviews with general practitioners, and conducted an inquiry with patients. To identify objective outcome measures, they reviewed literature and consulted both general practitioners and internal medicine specialists. They determined that the primary outcome measure should be a global assessment of patient judgement. Additionally, changes in symptoms of patients should be measured. The researchers concluded that standardisation of dietary advice would be difficult and discussed which design would be most feasible by referring to criteria for an adequate trial with IBS patients (Klein, 1988). In the end, a trial on the efficacy of mebeverine versus fibres and a placebo was proposed.

Although some questions remained unanswered, PAM employees proposed that a trial on mebeverine be commissioned. Despite this, the advisory committee decided, in January 2001, not to commission another trial on the effectiveness of mebeverine. They contended that the preliminary study failed to reveal a) the methodological problems of previous trials on mebeverine; and b) how these problems could be prevented in subsequent trials. As a result, they considered the feasibility of a methodologically sound trial on the effectiveness of mebeverine to be low due to potential placebo effects and other methodological problems (College voor Zorgverzekeringen, 2002).

Perspectives of policy makers, researchers, and target populations

As a part of the present study, interviews were held with relevant actors in order to reconstruct their perspectives. A summary is presented in Table 3.3.

According to an employee at the Ministry of Health, mebeverine should be removed from the health care package because of the lack of evidence on its effectiveness. The Ministry considered alternative interventions, such as dietary advice, to be more efficient. At that time, the Ministry was struggling with increasing costs of drugs and wanted to ensure that ineffective medicines did not impact the medical expenses carried by the community. They claimed that decreasing the cost of

drugs was necessary to prevent other problems in health care like waiting lists. They also claimed that the decision model was useful for deciding which interventions should be reimbursed and which interventions should not.

PAM employees considered an additional placebo-controlled trial on the effectiveness of mebeverine to be useless: *"[...] [A]fter a study on the effectiveness, we still might be unable to remove mebeverine from the public health package [...] [Alternatively,] you can try to affect prescription patterns by giving advice. Then, [interventions like] dietary advice and advice on a health regime become important".*

In order to change prescribing patterns, a different kind of information was needed and it was apparent that interventions other that drugs could be relevant. A social scientific approach in research was considered most appropriate. However, given that the pharmacy department at the HCIB is mainly involved in clinical trials, it was unlikely that they would accept a social scientific approach. Consequently, PAM proposed a study on the therapeutic value of mebeverine in comparison to other IBS treatments. Therapeutic value is also a criterion for drug reimbursement of drugs and PAM employees expected that a study on this aspect could provide useful information on prescribing patterns.

According to the researchers, the main problem was a lack of knowledge to support clinical practice. According to the researchers, the first step in the assessment of a drug is to define its efficacy. The effectiveness and therapeutic value of mebeverine could only be deemed relevant in subsequent phases. The researchers contended that dietary advice is very difficult to assess but that, if a diet is effective, this is because of an increased fibre intake. Additionally, they claimed that, before the effectiveness of dietary advice can be assessed, knowledge of whether or not fibres actually help was needed; *"[...] evaluating dietary interventions is complicated. It is easier to add only a bag of fibres [...] and it is easier to standardise [...]".* Researchers assumed that they would get the opportunity to perform the proposed trial.

Physicians acknowledged that mebeverine may not always help patients with IBS and that a placebo effect may be present. Nevertheless, they claimed to prescribe mebeverine because other effective treatments are lacking and because, in some patients, mebeverine appears to be successful. *"Mebeverine is easy in use and it is not harmful [...] it is the only therapy that can be given."*

Table 3.3.Reconstructed interpretive frames of actors involved

Actor	Judgement of solution	Problem definition	Background theories	Normative values
Ministry of health	No evidence on effectiveness mebeverine exclude mebeverine from health package.	increasing medicines costs	Decreasing costs of medicines is necessary to prevent other problems (waiting lists). Decision model adequate for reimbursement decisions	Only effective medicines at the expense of the community Affordable health care
PAM* staff	Affect prescribing practice A study on the effectiveness might be useless; Preferable, research on the therapeutic value of mebeverine as compared to a dietary advice.	Unlikely that mebeverine can be excluded from the package.	For good policy, it is important to know what is important to physicians and patients; A more social scientific approach might experience resistance, because internal traditions	Research that is relevant and useful for policy making
Researchers	Research on the efficacy of mebeverine compared to fibres is feasible	Evidence on the efficacy of interventions for IBS patients is lacking; Previous studies are methodologically flawed	Standardising diet is complicated If there's something in diet that is beneficial then these are fibres. Valid research provides relevant information	Research that is valid and feasible;
General practitioner 1	mebeverine is adequate, it might be effective in some patients; sometimes because the placebo effect more attention should be paid to psychiatric or mental causes by GPs and internal medicine	Some IBS patients visit physicians frequently; no effective treatment strategies available; counselling and reassurance take a lot of time. Compliance of dietary advice is low	Aetiology of IBS is unknown; Frequently, patients are anxious for severe illness (malignancies)	A good relationship with the patient
General practitioner 2	Sometimes medicines used from discomfit. Preferably, advice on healthy lifestyle & healthy food; Information flyers from the patients association;	Some IBS patients consult general practitioner frequently Not always possible to use other interventions besides medicines;	Patients have pain and wonder what the cause might be, including serious diseases; Some people are not acquainted with functioning of their own body; Complaints can result from unhealthy lifestyle	Inform and reassure patients
Gastro-entero logist	mebeverine is a sop; its effect is limited; reassurance is difficult and can take a lot of time; a solution should be to manage the anxiety; behaviour therapy is labour intensive (expensive);	During consultations, time is lacking; patients do have complaints but we don't have a solution for them	Patients sometime expect that a drug is prescribed; IBS is related to anxiety Several subgroups of IBS patients	To spend time meaningful

Actor	Judgement of solution	Problem definition	Background theories	Normative values
Patient (patients' association)	The patients association offers possibilities to talk to volunteers with IBS; Psychological care might be useful for listening; mebeverine might be effective in some type of spasms	Physician don't have enough time for talking and reassuring; Sometimes a physician has tried mebeverine in patients without success, the next patient won't receive it; current research does not take into account subtypes of spasms;	IBS can be due to general increased irritability Diet or stress can lead to complaints; Mechanism IBS is related to brain-gut-axis Currently, to much emphasis on scientific evidence, to less attention to patients experiences;	Recognition of complaints and disease

* PAM = department of Policy Analysis of Medicines

IBS patients visit general practitioners frequently but the treatment options are limited. Additionally, patient compliance with alternative interventions, such as dietary advice, is usually low. Thus, the effectiveness of these interventions is variable. According to some general practitioners, IBS is primarily caused by anxiety and not just a physiological abnormality of the colon. However, many patients are not open to psychological explanations of their illness. The presumption of physicians was that patients who receive mebeverine feel that they are taken seriously and will, as a result, visit the physician less frequently. A diagnosis and a prescription for a medicine are important for patients as they contribute to a sense of legitimacy and a feeling of recognition. The general practitioner aims to maintain a good relationship with patients and thus prescribes mebeverine. In the words of a gastroenterologist: *"I am convinced that if you talk to a patient for a long time, that patient can be helped and will need no medicines. However, that time is lacking."*

Patients with IBS mentioned that general practitioners only have a limited understanding of their complaints. For them, recognition of their complaints and being acknowledged is most important. Patients need to accept the existence of their complaints. Therefore, talking about their complaints is important. The patients' association established an IBS helpline that is maintained by volunteers who also have IBS. These volunteer workers can be contacted for answers to questions, advice, or just to tell one's story. According to the patient representatives, mebeverine, is effective in some patients.

Discussion

In this case on mebeverine, there was a lack of agreement on the kind of information that needed to be obtained. There was disagreement on which intervention should be included (diet or fibres) and on which outcome measures were most relevant (a decrease in symptoms, patient satisfaction, or number of consultations). Differences were found with respect to the criteria used to appraise the mebeverine situation. For physicians, establishing and maintaining a good relationship with patients was most important. They consider mebeverine to be helpful in the absence of other interventions. For the Ministry of Health, the increasing cost of drugs was important. It established criteria by which reimbursement decisions should be made and claimed that mebeverine should be excluded due to the lack of evidence needed to fulfil these criteria. In accordance with Hisschemöller's work, we contend that the problem was 'ill-structured': actors disagreed on the information needed and the norms at stake (Hisschemöller, 1993). Unfortunately, researchers and policy makers did not acknowledge that the problem was ill-structured. As a result, the problem definition shifted during the research project and the subsequent studies endeavoured to answer the wrong question.

The results from the preliminary study did not answer the HCIB's questions. It is important to note that the process of policy analysis can be seen as a series of successive rounds, each with its own problem definition: from initial indication to research problem; from research problem to research questions; from research questions to the provision of recommendations for policy. In every round, another actor was involved who redefined the problem, based on his or her interpretive frame and the contexts in which he or she worked. The end result is that the findings of the project were hardly useful for policy making. Similar problems have also been found in different settings (Hoppe & Grin, 2000).

The proposed research did also not correspond with the problems perceived by physicians and patients. Physicians did not prescribe mebeverine because they were convinced of its effectiveness. In fact, some physicians acknowledged a relatively high placebo effect in some patients. Physicians claimed to prescribe mebeverine for the following reasons: a) no other treatment strategies were available; b) too little time for counselling was available; and c) reassurance and acknowledgement of the patient was important. With this in mind, we can assume that the results of a trial indicating that mebeverine has limited effectiveness would not change physicians' current

prescribing practice. Excluding mebeverine from the health package could even limit their treatment options.

The error that was made in this case was that the policy maker failed to analyse the problem from the perspective of the target populations. A trial on mebeverine's effectiveness may have provided a solution to the Ministry's problem and, quite possibly, the evidence would have been sufficient to exclude mebeverine from the health package. However, it would not have solved the physicians' and patients' problems. To dissolve ill-structured problems, adequate problem structuring is essential (Hisschemöller, 1993). It is advisable to identify a policy's target populations and involve them during the developmental stages. The objective of involving these groups at an early policy development stage is to assess the degree to which policy implementation is dependent on their cooperation and to also estimate whether the requisite cooperation will be obtained. Obviously, in doing this, policy makers run the risk of that policy making will become the prisoner of its target populations. However, when analyses reveal that the requisite cooperation may not be obtained, policy makers can make a choice to apply additional measures so that the requisite cooperation is obtained (Grin & van de Graaf, 1996a).

To solve the health care problem of treating IBS patients, it may be important to acknowledge differences and variances between patients. Although physicians preferred to maintain the option by which they could prescribe mebeverine, it is questionable whether this drug provides the most optimal care for all patients. It is possible that the ambiguity of previous trials can be explained by an implicit assumption that patients groups are homogenous. The physicians we interviewed indicated that what can be considered the most effective treatment is different for different patients. Several theories on the underlying mechanisms of IBS exist. Some contend that complaints are due to physical abnormalities in the colon. Others claim that IBS is a physical expression of psychological factors, such as anxiety. At the same time, advocates for visceral hypersensitivity, for a neurotransmitter imbalance and for infection and inflammation exist (Horwitz & Fischer, 2001). Despite the varying theories, the previous studies on mebeverine were based on only one mechanism, namely the physical activity in the colon. Obviously, if a case is caused by psychological factors, the use of mebeverine will not eliminate the underlying cause. Patient education and/or behavioural therapy are then considered more appropriate (van Dulmen, 1996; van der Horst, 1997). Further, it is possible that the total IBS patient population is comprised of various subgroups. Unfortunately, professionals

have not come to any agreements on the criteria necessary to distinguish between sub-populations (Camelleri, 1999; Janssen et al., 1998; Ragnarsson & Bodemar, 1999; Wahnschaffe et al., 2001). In the absence of convincing evidence on specific therapies for identifiable subgroups of patients, physicians are likely required on to identify the best treatment for each patient individually. Evidently, this is the current practice of most physicians. It would, however, be advantageous to standardize this process as standardization can prevent a significant amount of bias. For example, n-of-1-trials can provide objective evidence that an individual patient is truly benefiting from a particular treatment rather than from the non-specific effects of treatment (Sackett et al., 1991). This approach appears to be promising despite its current lack of application in HTA studies. Additionally, as new drugs for IBS become available (Camelleri, 1999; Farthing, 1998), an approach whereby we identify the best treatment on an individual patient basis becomes important. These new drugs are not as inexpensive or harmless as mebeverine (Charatan, 2000) and it is likely that not all patients will benefit from their use. These suggestions correspond with the problems communicated by the physicians and are thus likely to be more successful. Further, our suggestions may stimulate increases in mebeverine use among those patients who are most likely to benefit while decreasing unnecessary prescriptions.

Our study has, like all studies, certain limitations. The first is that only a small number of respondents were interviewed. Although the study aimed to provide insight on the heterogeneity of the stakeholders' perspectives, given the small sample size, it is impossible to draw definitive conclusions on the views of all physicians or all patients. We cannot precisely gauge the generalizability of our findings. As a result, the recommendations provided with respect to the use of mebeverine among IBS patients are tentative.

Conclusion

The analysis of this case indicated that the health care problem was ill-structured which was not acknowledged by the stakeholders involved. As a result, the problem definition shifted and the resulting problem definitions failed to correspond with the problems perceived by the target populations. An argumentative approach in HTAs can help us to identify problems, to uncover the argumentation that underlies these problems, and to develop possible solutions. This, in turn, can not only make HTAs more relevant to decision-makers but also increase the effectiveness of future policy actions.

Policy implications

Policy makers should be more acutely aware of the possibility that problems may be defined quite differently by different stakeholders. Problem definitions critically determine the range of solutions that is taken into account. Hence, overlooking incongruencies in problem definition may lead to one-sided, partisan HTAs, with outcomes that are considered valid and relevant by only part of the target population. This, in turn, may severely hamper evidence-based policy making and resolution of the problem. We recommend that policy makers require a cogent analysis of the problem from a variety of perspectives, resulting in evidence of sufficient congruence in problem definition among stakeholders to ensure wider support for HTA outcomes and HTA-based policy decisions.

4

Non-compliance on the part of the professional community with a national treatment protocol: an argumentative policy analysis

In 1997, the National Health Insurance Board of the Netherlands introduced a guideline for the use of a new anti-epileptic drug, Lamotrigine. The goal was to limit the use of this relatively expensive drug to patients with difficult-to-treat epilepsy. A survey had shown that only a minority of neurologists were familiar with the guideline, and even fewer applied it in practice. In the present study, interviews were held with stakeholders to obtain a better understanding of why this policy measure failed. The results indicate that the problem definitions of policy maker and practising neurologists differed widely, and that the policy measure was conflicting with certain professional beliefs. In such cases, the theory of argumentative policy predicts that policy is unlikely to succeed, unless policy makers take actions to ensure a greater congruence in interpretive frames between them and their target population.

M Moret-Hartman, PD Knoester, YA Hekster, GJ van der Wilt. Non-compliance on the part of the professional community with a national guideline: An argumentative policy analysis. Health Policy 2006;78:353-9

The intractable problems associated with the implementation of public policy are well known (Lindblom, 1992; Netherlands Court of Audit, 2003; Trouiller et al., 2002). On the basis of an analysis of the nature and causes of these problems, policy scientists have argued that policy should be conceived as an instance of co-production between policy makers and the target audience (Fischer & Forester, 1993; Majone, 1989). A key feature of this argumentative policy theory is the recognition that different stakeholders may define policy problems quite differently, leading to different and sometimes opposing appreciations of proposed solutions. Differences in problem definition may, in turn, be related to differences in background theories and preferred ways of social organisation. Such ensembles are usually referred to as appreciative system or interpretive frame (Schön, 1983; Schön & Rein, 1994). In this concept of policy, it is crucially important to identify the target audience (Whose co-operation is necessary to make this policy successful?) and to identify their interpretive frames (How do they define the problem and how does this relate to other elements of their appreciative system?). Argumentative policy theory predicts that when there is evidence of insufficient congruence in problem definition between policy makers and target audience, implementation is likely to fail. In order to succeed, policy should also be directed towards achieving better congruence in problem definition. In other words, it should also be aimed at inducing a process of social learning. This may require adjustment on the part of policy makers, target audience, or both, and may entail reconsideration of policy options, evaluation criteria, or underlying assumptions and preferences (Grin & van de Graaf, 1998).

In this paper, we present the results of an argumentative policy analysis of a specific health care policy, enacted by the National Health Care Insurance Board (HCIB) in The Netherlands. This board is an advisory body to the Ministry of Health on coverage issues. In 1997 it issued a protocol for the use of a novel anti-epileptic drug, Lamotrigine in which the use of lamotrigine was restricted to difficult-to-treat patients. For reimbursement, Lamotrigine needs to be prescribed in accordance with the protocol. This initiative was taken since the costs of the new drug were substantially higher than those of conventional drugs, while there was no clear evidence that it had stronger anti-epileptic effect. The gist of the protocol was that the novel drug should be prescribed only to patients who show insufficient response or unacceptable side effects to (combinations of) conventional drugs. The protocol was issued to prevent that Lamotrigine would substitute conventional anti-epileptic drugs on a wide scale, with cost control as a major motive. The protocol was

distributed among all registered neurologists in The Netherlands. From a survey among neurologists, we found that the policy measure had been largely ineffective: only a minority (22%) of respondents knew the protocol, and an even smaller proportion endorsed its content and put it into practice (Tuinder et al., 2004). The aim of this paper was to identify the contents of the interpretive frames of policy maker and members of the target population. On the basis of this material, we discuss whether more congruence in interpretive frames should have been sought, and how this might have been used for successful policy making.

Methods

The method of reconstructing interpretive frames was used to elicit problem definitions, possible solutions, background theories, and preferences. (Grin et al., 1997) Data were collected by anonymous semi-structured interviews with a representative of the policy making institute and with members of the target audience (seven prescribing neurologists engaged in the treatment of patients with epilepsy). One neurologist was employed in a teaching hospital, one was employed in both a teaching and a general hospital, three neurologists were employed in a general hospital, and two neurologists were employed in a tertiary centre, specialized in treatment for patients with epilepsy. Two neurologists were also involved in the development of a broader guideline on the clinical management of patients with epilepsy, to be issued by the Dutch Society of Neurology.

In the interviews, questions focused on perceived problems and reasons for actions or decisions concerning care for patients with epilepsy. All interviews were taped, summarized, and coded, making distinction between four layers of interpretive frames: appreciation of solutions, definition of problems, background theories, and normative preferences. Respondent validation was conducted by sending a summary and interpretation of each interview to the respondent for correction. After validation, all respondents received an overview of results from all other interviews.

A summary is constructed of the key problem definitions, possible solutions, background theories, and preferences according to (1) the policy maker, (2) general neurologists working in a general or teaching hospital, and (3) epilepsy specialists working in a tertiary centre. Triangulation was conducted by checking findings from interviews in literature and documents.

Results

The reconstructed interpretive frames of the respondents are presented in Table 4.1.

Policy maker

To the policymaker, the protocol was a means to prevent neurologists from prescribing Lamotrigine in patients for whom a similar seizure control could be achieved with conventional anti-epileptic drugs, at a lesser cost. The problem stems from a fixed health care budget on the one hand, and the continuous development of novel health technologies on the other hand, for which funding is sought. Also, the problem is anticipated because more generally, physicians are held to be inclined to prescribe novel drugs to an extent that may not be supported by scientific evidence, so encouraged by manufacturers. A crucial aspect of the policy maker's background theory was the notion that the majority of patients with epilepsy can be adequately treated with conventional drugs. An important aspect of the policy maker's appreciative system was that an efficient use of public resources justifies restrictions on professional autonomy.

Neurologists

None of the respondents was aware that prescription of Lamotrigine should be in accordance with the HCIB protocol in order to obtain reimbursement from health insurance companies. As a key problem, neurologists reported to find, for the individual patient, the optimal (combination of) anti-epileptic drugs, and to find this optimal drug regimen as quickly as possible, without imposing undue harm. To them, there are two aspects that determine optimality: seizure control and side effects. The optimum may vary between patients, as patients respond differently to drugs, and because they value seizure control and the various side effects differently. Moreover, the optimum may not be stable over time: side effects may become apparent only after a prolonged period of time, acceptance of side effects or seizures may change, or drug effectiveness may decrease or may be affected by concurrent events such as pregnancy. This is the major challenge for the clinician, and anything that helps to achieve this objective will be welcomed. A protocol restricting the use of a novel anti-epileptic drug on the basis of its costs is not one of them.

Table 4.1. Interpretive frames of PAM staff and researchers

Actor	Proposed solution	Problem definition	Background theories	Normative values
Policy maker	Guideline is appropriate instrument to define reimbursement conditions for a new drug	How to control costs when novel, expensive technologies continue to be developed?	Health care professionals easily adopt novel technologies Conventional AEDs are sufficient for the majority of patients with epilepsy	Don't spend public money on services that perform marginally better at substantially higher costs
Neurologist general hospital	Guideline does not address problem of my practice	Little 'hands-on' experience with Lamotrigine Lamotrigine difficult in use	New drugs are not always better than existing ones Most patients can be treated successfully with conventional AEDs Relation between volume and quality of care	Safety Patients with refractory epilepsy should be treated in specialised centres
Neurologist teaching hospital	-	Little attention to long-term toxicity	Patients differ in their valuation of seizure control and side effects	Patients' quality of life
	Trials should be conducted that have greater relevance to daily practice	Evidence from trials not applicable to clinical practice	Knowledge on research methodology	Health care should be evidence based Guideline development is a professional responsibility Guidelines should be regularly updated
Epilepsy specialist	Therapeutic repertoire should not be unduly restricted	Selecting best therapy for each patient individually (search carefully by trial and error) conventional AEDs are far from optimal (toxicity)	incomplete seizure control incurs considerable costs to the patient and to society at large	Safety Professional autonomy Acting in the best interest of the individual patient Don't stop searching for better treatment modalities

Costs were not an issue, or, more accurately, costs were defined differently. To the neurologists, costs are incurred as long as no seizure control is achieved, without acceptable side effects. Here, costs are defined more broadly, in terms of unpredictability, interference with daily life, and costs of self-inflicted harm (resulting from seizures), to the patient and his family. Although not quantified, they are considered to outweigh the costs of drug treatment.

Apart from this commonality, there were certain differences between neurologists working in general hospitals, teaching hospitals, or specialised centres.

General Neurologists

Interestingly, the novelty of the anti-epileptic drug Lamotrigine was mentioned as a major problem by neurologists working in a general hospital. Inevitably, because of its novelty, relatively little is known, especially about the safety profile of the drug. Although trials have been published in the literature, the medical profession has had little opportunity to obtain experiential knowledge. This was considered particularly relevant, since earlier drugs that had been introduced on the basis of trial results, had turned out to be inferior to then available drugs in terms of safety and effectiveness. *"Some drugs promised to be very good, however, they appeared to have many side effects, which holds for vigabatrin, or were not as effective as promised, for example gabapentin"* (respondent N2, interview)

Respondents also considered Lamotrigine as relatively difficult in its daily use, since its dosage needs to be gradually increased in order to prevent rash. Generally speaking, neurologists thought that Lamotrigine was not more effective than conventional drugs. They also emphasized that in general hospitals, patients with uncomplicated epilepsy are treated. Generally speaking, they found no need to treat these patients with novel drugs, nor a justification to do so, in view of the limited knowledge of these drugs. Patients with more complicated epilepsy are referred to tertiary centres, specialising in care for epilepsy. As such, respondents considered it inappropriate to examine drug effects in this group of patients, and not part of their professional responsibility.

An additional problem was that evidence from trials does not always translate easily into clinical practice. The purpose of trials is to support registration of a new drug on the basis of its efficacy and safety. Because of differences in study populations and patients seen in daily practice, the generalisability of trial findings may be limited. According to this respondent, the HCIB should support the conduct of trials in a natural setting, to assist the professional community in finding the value of new drugs in daily practice.

With respect to costs, a respondent pointed out the arbitrariness of health care policy. Why should the use of Lamotrigine be restricted on efficiency grounds, while many treatments are covered that have never been assessed for their efficiency? Clearly, to this respondent, consistency in health care policy was an important element of his appreciative system.

Epilepsy specialists

In patients with more complex types of epilepsy, finding the best treatment is a process of trial and error: *'Predominantly, you follow the textbooks, but sometimes you try medicines or combinations of them as presented on conferences. Sometimes it is an improvement, sometimes not.'* (respondent N5, interview)

Respondents considered the toxicity profile of Lamotrigine an advantage, as well as its lower potential for interaction with other drugs. Furthermore, they had observed a positive psychotropic effect of Lamotrigine, which was sometimes an additional reason for its prescription.

An objection of these neurologists to the protocol of the national Health insurance Board was that it was too static. *'New anti-epileptic drugs are missing, the protocol is not updated [à] opportunities for evaluation of the treatment protocol should be incorporated.'* (respondent N4, interview)

Discussion

This study showed that policy makers failed to acknowledge how the target population defined problems or what argumentation gave direction to their prescription practices. The guideline might have seemed a solution for the problems of policy makers: to keep the prescription of lamotrigine low to control costs of anti-epileptic drugs. However, target populations did not cooperate and thereby policy failed.

Interviews with a limited number of neurologists allowed for the reconstruction of part of their interpretive frames that are relevant to the issue of the usage of novel anti-epileptic drugs. This resulted in information that is important to the policy maker in a variety of ways: Firstly, the target population appeared to be heterogeneous. Neurologists working in general hospitals differed from neurologists working in teaching hospitals or specialised centres in a way that should be taken into account when devising policy measures. Secondly, these general neurologists appeared cautious in prescribing new drugs. Little ground appeared to exist to assume that all neurologists would start prescribing the new anti-epileptic drug on a wide scale. Thirdly, other problems are experienced by neurologists in their treatment of patients with epilepsy. The HCIB might assist in resolving these problems, thereby realizing their own policy objectives: optimization of quality and efficiency of health care.

The gist of the protocol, issued by the HCIB was: try to achieve seizure control, without incurring serious side effects, using (combinations of) conventional anti-epileptic drugs. The rationale for this recommendation was cost control. Interestingly, the recommended strategy is common practice among general neurologists interviewed in this study. The rationale, however, is different: they consider the novel drug not particularly easy to use, and, more importantly, they have learned in the past that novel drugs, although approved by national agencies, need not always be better than existing ones. Neurologists considered conventional drugs such as valproate safe and effective, although recently the toxicity of these drugs has been discussed (Kaplan, 2004). Part of their professional ethics prohibits neurologists experimenting with new drugs when the annual number of patients seen in their practice is too small. This may be especially true of patients with epilepsy, where achieving seizure control without incurring side effects is notoriously difficult.

For neurologists working in teaching hospitals or specialised centres, however, the protocol was largely irrelevant. They only treat patients with refractory epilepsy; attempts to achieve seizure control with conventional drugs have already been made and unsuccessful. The type of problems that are experienced in their management of patients with epilepsy are related to the unpredictability of responses of individual patients to various treatments. The challenge, then, is to find the optimal treatment for each individual patient as quickly as possible. Data from published trials are relevant to this purpose, but to a limited extent: study populations may, and often do, differ from patients seen in daily practice, and treatment protocols may be atypical (Black, 1996). It would be helpful, therefore, to conduct more naturalistic studies[1] (Tunis, 2003) and to conduct N of 1 trials[2] (Sackett et al., 1991). Also, the setting up of central registries where unexpected events can be reported when treating patients with anti-epileptic drugs would help to identify possible side-effects at an earliest possible time, since trials have not always been found to constitute a reliable source for this type of information (Derry et al., 2001).

[1] A naturalistic trial is a trial in a natural setting

[2] A multiple cross-over trial within one patient. The patient receives one period intervention A and another period intervention B. The sequence of treatment is randomised.

Alternative solutions

Although current prescription practice resembled the recommended strategy, in the guideline from the Dutch Society of Neurology, Lamotrigine was proposed as a drug of first choice. Unfortunately, little evidence was given that supported this recommendation. Neurologists' experiences appeared important in their judgment. However, individual experiences on effectiveness and toxicity might be biased. As long as evidence is lacking, a careful prescription and strict monitoring of patients is needed. It would be justified to stipulate that novel drugs are used exclusively in specialised centres. Furthermore, data on prescription and patients' outcomes should be collected systematically. Subsequently, knowledge from these data can be used for process of learning by neurologists on the value of lamotrigine or other drugs.

Alternatively to guideline development, the HCIB might consider ways of assisting or encouraging the professional community to enact such measures, e.g. by co-funding naturalistic trials (that are unlikely to be funded by manufacturers) or by covering the costs of setting up N of 1 trial facilities or central registries.

If policy institutes help funding naturalistic trials and setting up research facilities, it would not be unreasonable to demand that they are involved in deciding whether novel drugs should continue to be used in specialised centres or may be released for general usage.

Limitations of study

The study has, of course, certain limitations. A small number of respondents were interviewed. This might affect the generalizability of our results. Although the study aimed to provide insight in the heterogeneity of stakeholders' perspectives, conclusions about the viewpoints of all neurologists or the completeness of possible viewpoint on this subject cannot be drawn. It cannot be excluded that in other general hospitals, neurologists do not refer patients who fail to respond to conventional drugs to specialised centres. Also, there may be other areas where physicians are more likely to adopt novel drugs. Policy makers should, therefore, examine this aspect from case to case.

Conclusions

This case study supports the idea that it is important to establish how target populations of policy measures experience problems and which solutions appear sensible to them at an early stage of policy development. The investment that this requires in terms of money and human resources is almost certainly modest as compared to the costs of (repeated) health policy failures.

5

Participatory workshops are not enough to prevent policy implementation failures: an example of a policy development process concerning the drug interferon-beta for multiple sclerosis

A possible explanation for policy implementation failure is that the views of the policy's target groups are insufficiently taken into account during policy development. It has been argued that involving these groups in an interactive process of policy development could improve this. We analysed a project in which several target populations participated in workshops aimed to optimise the utilisation of an expensive novel drug (interferon-beta) for patients with Multiple Sclerosis. All participants seemed to agree on the appropriateness of establishing a central registry of Multiple Sclerosis patients and developing guidelines. Nevertheless, these policy measures were not implemented. Possible explanations include 1) the subject no longer had high priority when the costs appeared lower than expected, 2) the organisers had paid insufficient attention to the perceived problems of parties involved, and 3) changes within the socio-political context. The workshops in which representatives of the policy's target populations participated did not provide enough interactivity to prevent policy implementation failure.

M Moret-Hartman, R Reuzel, J Grin, GJ van der Wilt. Participatory workshops are not enough to prevent policy implementation failures: an example of a policy development process concerning the drug interferon-beta for multiple sclerosis. Health Care Analysis 2008;16(2):161-75.

There are numerous examples of health policy measures that have failed in their implementation. A review from the Netherlands Court of Audit indicated that no evidence of a full policy implementation can be found in previous audits (Netherlands Court of Audit, 2003). For example, the applied measures for reducing greenhouse gas emissions have not met predetermined targets, the legal rules for maintaining food safety have not been able to prevent targeted risks, and policy measures for cost control of medicines have not resulted in more structural cost control (Netherlands Court of Audit, 1992).

The theory of argumentative policy analysis offers a possible explanation for such policy implementation failures (Fischer & Forester, 1993; Hoppe & Peterse, 1998). The basic idea of this theory is that actors' behaviours can be explained by different views on a problem and the argumentation behind these views. According to the argumentative approach to policy analysis (Fischer, 1999; Fischer & Forester, 1993), action is driven by processes of problem setting (Schön, 1983) in which actors define coherent sets of problems and solutions that correspond to these actors' normative and empirical background theories. The way in which problems are defined depends on the assumptions the actors make about the situation and their beliefs regarding what is good practice (normative values).

Combining these findings with those of classical implementation theory, Grin and van de Graaf (1996a) have argued that a policy will only be effective if both implementers and target populations consider the proposed policy measure meaningful. (Target populations consist of persons who will experience the consequences of a policy when implemented.) This means that the proposed policy should (a) make sense in the light of problems perceived by the target populations and (b) be consistent with their normative and empirical background theories.

However, the fact that policy problems and associated solutions tend to shift over time, different actors with different background theories being involved successively, renders this rather complicated. In other words, a policy measure is not invented at a specific moment in time, but develops over time. Nevertheless, the challenge is to identify policy measures that cohere with the views of all actors involved.

If the argumentative policy theory is correct, then it is advisable to identify the policy's target populations and involve them in the process of policy development. An interactive process of policy making could thus ensure that policy coheres with and prevent that it diverges from the views of target groups.

The objective of this paper is to analyze a specific instance where target populations were involved in the process of policy development. The policy institution in this example was the Health Care Insurance Board in the Netherlands (HCIB). It is an advisory board to the Ministry of Health, particularly with respect to coverage and reimbursement issues. The Department of Policy Analysis of Medicines (PAM) is responsible for identifying developments that may jeopardise optimal medical care, analysing the nature and size of such threats and conducting further research that may provide a basis for policy decisions. The annual work programme in which topics are prioritised is submitted to the Ministry of Health for approval.

The policy objective discussed in this paper was to promote the appropriate use of a recently introduced drug for patients with multiple sclerosis. Multiple Sclerosis (MS) is a neurodegenerative disease characterised by neurological dysfunction. The drug interferon-beta (IFNß) appeared to be a promising treatment. However, its costs were high (€ 12,000 per patient annually) and the evidence of its long-term effectiveness was limited. Therefore, policy measures to guide the prescription of this medicine were considered necessary. Representatives of prescribing physicians (neurologists), health insurance companies, and patients were invited to participate in two workshops. During these workshops, participants discussed policy measures that could promote the appropriate use of the new drug. Although two concrete policy measures had been proposed, these have never been implemented.

In this paper we evaluated whether the theory of the argumentative policy analysis could explain the proceedings in this case study. It offers a description of the process of the policy development that included the two workshops, as well as an analysis of the views of policy makers and target populations in order to assess whether the proposed policy measures fitted the perceived problems and underlying background theories.

Methods

Relevant documents were analysed and semi-structured interviews were held with the various stakeholders. These documents included correspondence, reports from the Board, internal memos on this subject, a report from Health Council on IFNß, reports from meetings, research proposals, and research reports. Interviews were held

with three HCIB employees, two employees from the Ministry of Health, the organisers of the workshops, two neurologists, a patient, and a medical advisor from a health insurance company. Participants of the workshops were contacted and interviewed to reconstruct the proceedings of the workshops. MS patients were contacted via a Dutch MS patient organisation. Representatives of target populations were interviewed to assess their problem definitions and underlying argumentation.

Interviews were transcribed verbatim. A summary from the interview and a concept report were sent to respondents for verification and literature was used to check the findings from interviews or documents. In line with the theory of argumentative policy analysis, we used the method of reconstructing interpretive frames to analyse target populations' argumentation (Grin & van de Graaf, 1996b, 1997; Moret et al., 2007). The idea is to reconstruct (1) how problems are defined, (2) how solutions or policy measures are judged, (3) what theoretical and normative assumptions shaped them, and (4) what normative preferences underlie this all. Together, these four 'layers' of evaluation entail an individual's interpretive frame.

Results

Proceedings of project

The content of the workshops was, to a large extent, determined by two reports concerning IFNß. These had been issued by the HCIB and by the Health Council, respectively. In July, 1995, the Ministry of Health asked both institutions to provide advice regarding the introduction of IFNß on the Dutch market. The Ministry asked for recommendations with respect to whether or not restrictions could be imposed on the reimbursement of the drug.

In May 1996, a committee from the Health Council reported to the Ministry of Health (Gezondheidsraad, 1996). This committee was comprised of three neurologists, a chemical technologist, and a HCIB staff member. It concluded that IFNß could be a promising new drug, but also cautioned against unrealistic expectations. Clinical trials had shown that IFNß could decrease the rate and severity of exacerbations, but there was no evidence of IFNß preventing the onset of disability. The Health Council emphasised that the drug should be prescribed only to patients who meet eligibility criteria for the trials. These criteria were: a) clinically definite Relapsing Remitting MS; b) at least two exacerbations in the two previous years; c) mild to moderate disability; and d) age of 18 years or older. The Council also

proposed the following policy measures: 1) develop a guideline for treatment of MS patients with IFNß; 2) ensure that the drug is prescribed by neurologists who have sufficient experience with the diagnosis and treatment of MS patients; 3) properly instruct and guide MS patients; 4) conduct a systematic follow up of patients using IFNß in order to evaluate side effects; 5) provide clear indications on which patients should be treated with IFNß; 6) establish a national registry of MS patients in order to conduct further research on the effectiveness of IFNß; and 7) conduct a re-appraisal after a number of years.

June 1996, the HCIB issued recommendations that corresponded partially with the Health Council's report. The HCIB recommended that: a) patients fulfil the criteria described in the Health Council's report; b) a treatment protocol be developed; c) health insurers approve reimbursement requests before starting treatment; d) a prospective registration be established; and e) the use of IFNß be re-assessed after three years (College voor Zorgverzekeringen, 1996). Restricting prescription to a limited number of experienced neurologists was considered impossible, because IFNß had already been included in health care package.

In 1999, the PAM department started a project on IFNß. Initially, they considered evaluating other, less expensive interventions that could be effective, but are of no interest to the industry. Indications had been received from the field that a much cheaper drug, namely methotrexate (used in oncology and rheumatoid arthritis), could be as effective as IFNß. Furtheremore, PAM considered developing a national database of MS patients. PAM later decided that conducting a clinical trial was not its responsibility. PAM staff, in collaboration with neurologists, decided to initiate the development of a treatment protocol and a national registry of MS patients. They proposed the organization of two workshops.

Workshops

In 2000, these workshops were organised by an external institute for policy research. The first workshop aimed to provide an overview of the criteria that are used for prescribing IFNß. This workshop was preceded by an inventory study among health insurance companies into current reimbursement practices with respect to IFNß. This inventory study was performed by a HCIB department. The results from the inventory study on current reimbursement practices indicated that health insurance companies did not assess reimbursement requests against medical content. Assessments were limited to an administrative review of the completeness of data (College voor

Zorgverzekeringen, 2000a). Nevertheless, health insurance companies considered the pre-utilization approval effective, because it functioned as an administrative barrier. The authors of this report recommended the following: a) assess whether a pre-utilization approval by a central committee could improve the assessment of reimbursement requests; b) develop a clear protocol and design application forms.

Neurologists, medical advisors from health insurance companies, policy makers, and a representative from a patient organisation participated in the workshop. During the workshop, vignettes with patient descriptions were used to discuss which patients should be treated with IFNß. The organisers concluded that only a few problems were perceived with respect to criteria for IFNß use. Health insurers questioned whether the situation was indeed problematic given the small size of the patient population and IFNß's status as an essential medicine. Neurologists argued that problems arose incidentally. Most often, these problems were related to reimbursement requests for continued use of IFNß. All participants agreed that subjectivity in decisions on prescription should be minimised. The participants claimed that clear distinctions should be made between the criteria for initializing treatment and the criteria for continuing treatment. The development of a guideline by neurologists was considered relevant. According to the participants, small adjustments of the current criteria would be sufficient (B&A, 2000).

The second workshop aimed to obtain advice on how existing databases could be improved so that the effects of IFNß or other new drugs for MS patients could be evaluated. In preparation for this workshop, a neurologist was asked to make an inventory of available databases that contained data on IFNß users. This inventory showed that data on MS patients were no longer being collected systematically in the Netherlands (B&A, 2000). Former local databases were no longer up-to-date. The participants in the workshops agreed that a national database could be relevant and could serve the following goals: a) policy making (such as financial surveys, planning health care capacity); b) research (effects of treatment; monitoring for side effects); and c) clinical practice support for neurologists (reflection on clinical practice, improvement of expertise). Participants agreed on the inclusion of general static data in the database, such as demographic details, diagnosis (type MS), the criteria used to make the diagnosis, and the initial treatment. Participants disagreed on the kinds of dynamic data, such as treatment details and information on physical functioning. A database including a broad range of data could be useful for research, but this was considered expensive. A small database that includes static data only could be used as

a sample frame for selecting potential participants in further studies. In their report, the organisers concluded that additional research was needed to ascertain which variables should be included in the database.

Proposed policy measures

Based on the organizers' report, the HCIB proposed the following policy measures to the Ministry of Health (College voor Zorgverzekeringen, 2000b):

- Have neurologists develop an evidence based guideline for the treatment of patients with multiple sclerosis.
- Have reimbursement requests for IFNß appraised on medical grounds.
- Possibly, implement a central pre-utilisation approval by experts (neurologists).
- Define clear criteria for (dis)continuing IFNß reimbursement
- Establish a national registry of MS patients. The database should enable scientific research, the improvement of treatments, monitoring, and the evaluation of efficiency and therapeutic value of new interventions.

In 2003, a medical advisor from a health insurance company mentioned that some small changes had been made to the procedures established for judging reimbursement requests (requests for continued use). A neurologist mentioned working on a guideline for diagnosis and treatment of MS patients. According to a PAM staff member, neither a central pre-utilisation approval nor a national database to prospectively register patient data had been established.

Reconstructed interpretive frames

In June 2003, we conducted interviews to reconstruct interpretive frames of policy makers and policy's target populations. The actors' views are summarised in Table 5.1 (IFNß guideline) and Table 5.2 (national database).

Ministry of Health

Initially, the Ministry contended that the long term effectiveness of the drug was not established sufficiently, while the acquisition costs were high. Trials had shown that IFNß decreased the number of exacerbations, but it was unknown whether IFNß would prevent disability.

-

Table 5.1. Judgement of guideline or central pre-utilisation approval for IFN and underlying problem definitions

Actor	Judgement of solution	Problem definition	Background theories	Normative values
Ministry of health (1996)	+ Clear criteria for reimbursement needed	IFN is an expensive medicine uncertainty on meaning of effects Possibly inappropriate use because high expectations	Only medicines of which evidence of effectiveness is available at the expense of the community If no evidence on effectiveness, use is inappropriate	efficiency
Ministry of health (2003)	- Central committee not efficient; too expensive Guideline should be set up by neurologists	Actual costs were lower than expected costs. Health insurers did not perceive problems	Finances (also) influence the Ministry's agenda.	cost control efficiency Quality improvement is neurologists' task
Health care insurance board (also in advice Health Council)	+ clear criteria for start & discontinuation needed; treatment protocol should be developed	Health insurance companies judged reimbursement requests differently, Expensive medicine; while effectiveness uncertain; A large population of potential users	IFN belonged to new type of medicines, of which impact was unclear Impact = does use lead to savings elsewhere	Task: to assess new medicines & defined relative position towards alternative interventions
Staff member departmen t PAM	+ possibly, a central pre-utilisation approval a clear treatment protocol should be available	expected increase in use IFN; health insurance companies judged reimbursement requests differently Individual health insurer lack expertise to judge reimbursement requests	Situation IFN resembles situation growth hormones, for which a central pre-utilisation approval in combination with central registration was set up Central pre-utilisation approval improves efficient use of medicines, prevents inappropriate use	cost control; efficiency equal accessibility
Staff member departmen t PAM (2003)	- no central judging committee (expensive/ additional value is limited)	no problem concerning reimbursement requests Situation concerning IFN was relatively stable (only discussion in case of new medicines or extension of indication)	Criteria for central pre-utilisation approval: - medicine is expensive - inappropriate use expected - decisions on treatment are complex	efficiency
Medical advisor of health insurance company	+ / - guidelines needed for decisions reimbursement Central pre-utilisation approval no longer meaningful	unclear when IFN is (still) meaningful neurologists are prescribing IFN;	Pharmaceutical companies benefit from extending criteria Expensive drugs are financed from public funds, on a solidarity base ;For accessible health care, fundamental choices should be made ;No clear guideline on acceptable outcome measures for expensive medicines	Clear vision from the ministry of health on pharmacy Physicians are responsible for guideline development;
Neurologis t	+ / - No IFN guideline but MS guideline (diagnosis & treatment); summarise available evidence	In the past, differences in reimbursement decisions between health insurers	If a medicine is available at the market and health insurers reimburse it, it should be possible to prescribe the medicine.	be able to offer something to patients

Actor	Judgement of solution	Problem definition	Background theories	Normative values
Patient	+ Guideline might be useful	care differences between hospitals (IFN not offered in every hospital); Advises (on lifestyle) are contradictory	No common knowledge on MS Prognosis differs between individuals	Follow own experiences

The potential target population, namely patients with MS, is large and the expectations of both patients and physicians may also have been high. As a result, the risk that the drug could be used inaccurately was considered to be high. Consequently, additional policy measures to control IFNß use in clinical practice were requested.

According to employees from the Ministry of Health, the HCIB report (College voor Zorgverzekeringen, 2000B) had revealed that only health insurers perceived few problems with respect to judging reimbursement requests. The costs of a committee for a central pre-utilisation approval were considered relatively high, while the costs of the interferon-beta prescription were less than expected, namely € 18 million (in 1999) versus the € 90-180 million estimated by the Ministry. Therefore, both interventions were considered to be no longer relevant.

PAM staff

According to the PAM staff, the main problem was that the use of IFNß was expected to extend to other subgroups of patients. A central pre-utilisation approval in combination with a national registry of MS patients could provide a solution. This contention was made because the situation concerning IFNß resembled the situation concerning other drugs for which a central pre-utilisation approval had already been established. In that case, requests for reimbursement of drug use were judged by a central committee of medical experts. For this purpose, a clear protocol including criteria for IFNß use is needed. Simultaneously, patient data were recorded in a national registry.

At the time that PAM staff started their project, it was undesirable to evaluate why the proposed policy measures had not been implemented until then. Reason was a change in the relationship between the HCIB and the Ministry of Health. In 1999, the HCIB, an independent advisory board, was established as the successor to the Sickness Funds, which was a politically involved advisory board comprised of actors from the field.

Table 5.2: Perspectives towards a national registry of MS patients or IFN users

Actor	Judgement of solution	Problem definition	Background theories	Normative values
Ministry of health (1996)	+	Uncertainty on (meaning of) effects	Only medicines of which evidence of effectiveness is available at the expense of the community If no evidence on effectiveness, use is inappropriate	efficiency
Ministry of health (2003)	- Central committee no longer efficient	Expectations on costs and problems did not come true. Health insurers do not perceive many problems (anymore)	Finances (also) determine the Ministry's agenda. MS is complex disease; difficult to measure effects, to know what should registered	cost control efficiency
Health care insurance board (in Health Council)	+ database for 1) patient data, to take samples for research, or 2) register all relevant data for research	Insufficient evidence concerning effectiveness of IFN (disease progression)	Aim of national database: to monitor use of IFN in practice (which patients (indications), volume, costs) and to evaluate effectiveness INF	Task: to assess new medicines
PAM staff	+ National database linked to central judging committee	discussion on effectiveness of IFN;health insurance companies judged reimbursement requests differently, leading to difference in health care	central pre-utilisation judgement provides a minimal registration	efficiency
PAM staff (2003)	- registration without central judgement committee difficult	central committee no longer relevant No problems perceived	A database can be a means in policy research, not a goal /objective	Efficiency setting up databases not their task / responsibility
Medical advisor of health insurance company	+ prospective database only way to assess long term effectiveness database only useful if clear which data should be registered	Unclear when IFN is meaningful Judgement of relative value as compared to alternative interventions not systematically but in the neurologists head no clear guideline on acceptable outcome measures for expensive medicines	Expensive drugs are financed from public funds Pharmaceutical industry's power is underestimated For accessible health care, fundamental choices should be made. What data should be registered is a political question	clear vision from the ministry of health on pharmacy quality of care
Neurologist	+ / - depends on aim and who is responsible	aim of national database was unclear registration takes time	If a medicine is available on the market and health insurers reimbursement If registration for policy evaluation, Dutch government should pay and initiate it	be able to offer something to patients
	+/- Neurologists not responsible for data collection. Possibly a role for patients in data collection	Aim behind database was unclear National coverage might be difficult	MS neurologists have their own "kingdoms" Distrust among neurologists	

After completion of the project, the PAM staff considered the option of having reimbursement requests judged by a central committee no longer meaningful. They considered the implementation of policy measures difficult because the drug had already been introduced several years earlier. Limiting or discontinuing the reimbursement of IFNß would have been practically impossible. Furthermore, the scale of problems relating to the prescription and/or reimbursement of IFNß and the costs related to IFNß use turned out to be much smaller than expected. PAM considered the initiation of a national MS patient registry without the central pre-utilization judgment to be infeasible.

Medical Advisor Health Insurance Company

According to a medical advisor from a health insurance company, the problem was attributable to a lack of clarity on exactly what kind of patients benefit from IFNß. IFNß is an expensive drug and the costs are either carried by the community (public health insurance) or reimbursed on an individual basis (private health insurance). Preferably, physicians should develop a guideline on the relative position of IFNß in relation to alternative interventions.

At the time of IFNß's introduction, a central pre-utilisation approval could have been relevant. However, time had passed and the prescription of IFNß had become common practice. Unfortunately, new drugs are often introduced and included in the health care package long before all medical specialists agree on criteria for treatment. Professionals cannot develop guidelines quickly. In contrast, the time that passes between a drug being introduced into the market and its inclusion in the health care package is regulated by law and, in most cases, quite short. As a temporary solution, the drug could be included into the health care package under certain conditions, such as the registration of patient data.

Neurologist

A neurologist mentioned working on a guideline for diagnosing and treating MS patients as part of an initiative from the Dutch Society for Neurology. A guideline for diagnosis and treatment was considered relevant as it could support evidence-based practice. As in other medical fields, knowledge and treatment options have increased. The development of a guideline is a time-consuming endeavor as, often, it has to be done alongside the professional's usual activities. A central pre-utilisation approval was no longer relevant. Prescribing IFNß was considered common

practice, also in small general hospitals. Most neurologists were already familiar with the indications for treatment. IFNß was proven to be effective in one type of MS that is characterised by invalidating exacerbations (relapsing remitting MS) through clinical trials that demonstrated IFNß's ability to decrease the severity and frequency of these exacerbations. A national database could still be useful. However, its purpose must be clear. An appropriate goal could be to evaluate the long-term effectiveness of IFNß. The neurologist questioned whether participants would have come to agreement on specific recommendations, such as which data should be collected and who should become the owner of the database. Most neurologists do have 'their own kingdoms' and do not want to share these with others.

Patient

The patient claimed that agreement among neurologists about MS treatment could be useful. For him, the main problem was that MS care differs between hospitals. He conveyed his experiences with receiving contradictory answers to questions from numerous health professionals and the option of using IFNß had not been discussed in the hospital where he was initially treated. For this patient, the exacerbations are highly invalidating. Obviously, his preference is to lead a normal life, in so far as that is possible. MS has a high impact on his life.

Discussion

Although results from the workshops showed that all respondents agreed that developing a guideline and a national database could be meaningful, these policy measures have never been implemented. A number of factors can provide an explanation for the proceedings in this project.

Firstly, the costs of IFNß appeared to be lower than expected. Perhaps the information campaign on IFNß had had this effect (Coolen, 1997). It may also have been that the calculation by the Ministry of Health, which differed from calculations made by the Health Council and rested on the assumption that all MS patients are treated with IFNß, were unrealistic. In any case, the subject of IFNß no longer had high priority. From the point of view of policy makers, the expected high costs of IFNß in combination with its uncertain effectiveness were the main problems behind the proposed policy measures. Apparently, however, the problem of the high costs had still higher priority than the limited knowledge on IFNß's effectiveness. As a

result, policy makers left it to the medical profession to implement policy measures. Neurologists, however, considered cost-containment not their problem and not their responsibility,

This also explains why the initial questions about the long-term effectiveness remained unanswered, without this being perceived as a problem. The effectiveness of IFNß and the possible role of the industry was an issue of international debate. Clinical trials have shown that IFNß decreased the number and severity of exacerbations in patients with relapsing-remitting MS (IFN MS Group, 1993; PRIMS, 1998). Probably, this is enough reason for neurologists to claim the right to prescribe IFNß. That evidence of long-term effectiveness, in terms of preventing disability, and safety was still limited (Rice, 2002) is not of great concern to them. Neither is the possible influence of the industry on prescription practices through the financing of clinical trials that had been discussed (de Haan & Vermeulen, 1999; Pieters, 1998; Polman, 1999; Vermeulen & de Haan, 1999). Some authors have even challenged the hypothesis of MS as an inflammatory auto-immune disease (Behan et al., 2002). A national registry could have been relevant from this point of view. In the UK, the Department of Health initially refused to reimburse IFNß use for reasons of costs (Lie, 2004). In 2002, however, agreements were made with the industry concerning the funding of IFNß. The Department of Health announced that it would reimburse IFNß for MS patients who agreed to participate in a monitoring program on the effectiveness of IFNß.

Secondly, the policy development process from the beginning centred around two specific policy measures, whereas it could have started with a broader scope, actors involved first eliciting the problems they perceive and adequately structuring the problem. Now, the proposed measures appeared not to be the most optimal solutions for the problems as perceived by neurologists, and therefore represented right solutions to the wrong problems. This typically concerns what Hischemöller and Dunn have coined an 'error of the third kind' (Hisschemöller, 1993; Dunn, 2004). We have found that on a first order level, all actors agreed that a national guideline and a registration could be meaningful. But as it was, on a second order level they disagreed as to what goals such policy measures should serve. For example, neurologists considered a national guideline relevant, as long it was a broad guideline on the diagnosis and treatment of MS in general. All actors agreed that a national registry could be relevant, but defined different goals. Each goal came with different variables to be included in the registry.

Thirdly, during the workshops, the proposed policy measures remained rather vague and were not elaborated in detail. As a result, participants could easily consent without violating their background theories and preferences. They had not felt the urge to (re)consider these theories and preferences. This, however, was necessary, as is clearly illustrated by the respondents who questioned whether participants would have come to agreement on specific recommendations, such as which data should be collected and who should become the owner of the registry. Such an agreement requires that policy measures are made sufficiently substantive for every actor to understand what are the consequences. Moreover, it requires that background theories and normative preferences are elicited and scrutinized.

Fourthly, changes in the socio-political context affected policy development. Opinions with respect to which kind of policy measures are the responsibility of the HCIB had changed over time. The PAM staff considered conducting a clinical trial to obtain relevant information that would not be provided by the industry. Simultaneously, their position towards another Dutch institute involved with subsidising clinical research changed. Putting out clinical trials was then no longer their responsibility. Furthermore, the relationship between the HCIB and the Ministry of Health changed over the course of this project. These changes strongly affected ideas with respect to the Board's responsibility on initiating a national registry of patient data.

On this basis we conclude that the workshops failed to meet the objectives set for interactive processes in policy development, namely to prevent policy implementation problems as a result of diverging views among target groups. Moreover, we concede that they could not have met those objectives. From the case of IFNß, we infer that an interactive process should meet the following criteria:

- The interaction should cover the whole process from problem structuring to policy implementation, in order to be able to deal with problem shifts and changes in the socio-political context. It is not enough to reduce interactivity to workshops at one or two moments in time.
- Actors involved should resist the temptation to think that policy problems can be understood at a first order level. That is, background theories and normative preferences should be explicated in the problem definition phase. Only if problem structuring is taken serious in this sense can one think of developing solutions that could meet with the approval of all target groups.

- Developing widely endorsed solutions requires that actors involved are willing to learn from one another and adapt their views if necessary. A process of interactive policy development should include room for such learning processes (Guba & Lincoln, 1989; Reuzel, 2001; Reuzel, 2004; Yanow, 2000). Grin and van de Graaf (1996a) have argued that learning is likely to occur only if external events urge a revision of background theories and preferences, or if repeated failures show actors that their background theories are not functional.

- Proposed policy measures should be sufficiently elaborated, as to enable target groups to assess their consequences and constructively engage in the interactive process. As the case of IFNß shows, actors involved are not willing to reconsider their background theories and preferences, should proposed policy measures not be sufficiently elaborated and actors involved not understand what is at stake.

In sum, this study has shown that the organisation of workshops in which target populations participate does not qualify as an appropriate process of interactive policy development. From the beginning, emphasis had been put on a limited number of interventions aimed to control the expected increase in treatment costs. Although target populations participated in policy development, perceived problems and which interventions could provide a solution had been discussed insufficiently. For policy development to be successful, interactive methods are needed, in which problem definitions and assumptions are explicated and discussed, providing an opportunity for mutual learning between actors involved.

6

Importance of identifying interventions that are congruent with perspectives of actors involved: Lessons from a case study on new analgesics

The Dutch Health Care Insurance Board, an advisory body to the Ministry of Health commissions research to guide their policy recommendations. However, the studies conducted did not always yield relevant information which might result from differences in problem definitions held by the Board and by target populations. To compensate for these problems, the Board made a number of changes in their commissioning procedures. Researchers were asked to analyse problems from the perspective of the policy maker and the policy's target populations. We monitored and analysed this new approach when used in a European tender that was issued for research proposals on a new type of analgesics, selective cox-2 inhibitors (COXIBs). Main reasons for commissioning this research were its frequent off label prescription and high costs. Our analysis showed that researchers seemed to be reluctant to adopt the Board's problem definition, questioning whether it would be shared by 'the field'. At the time of finishing the study, however, one of the COXIBs was withdrawn from the market because of serious side effects. The results of this case study emphasize that involving target populations in policy research should not imply that target populations control the problem definition and identification of solutions. The primary aim of involving target populations is to identify solutions that are congruent with both the perspectives of policy makers and policy's target populations.

M Moret-Hartman, GJ van der Wilt, J Grin. Importance of identifying interventions that are congruent with perspectives of actors involved: Lessons from a case study on new analgesics (submitted)

As in most Western societies (Kamke, 1998; Hurst, 1991; Rodriguez, 1999), a crucial challenge to Dutch health policy makers is to ensure that the quality of and access to health care is maintained or improved while keeping costs within acceptable limits. Striving for an optimal balance between quality, accessibility, and efficiency requires new and innovative policy making strategies. In order to be considered as legitimate and effective in bringing about the requisite re-orientation of health care practices, it is crucial that policies are designed such that a sufficient level of congruency is achieved in problem definition between policy makers and target groups (Grin & van de Graaf, 1996a). This task is far from trivial. It involves a shift from a supply-driven system to a system that seeks to optimally resolve health care problems. It also demands a shift from a system dominated by medical rationality to a system that is considerate of the wide variety of rationalities possessed by all parties, including providers of care, recipients of care, and insurers (Grin, 2004).

One of the organisations charged with this task is the Dutch Health Care Insurance Board (HCIB). Within the Board, the Department of Policy Analysis of Medicines (PAM) responsible for identifying developments that may jeopardise optimal medical care, analysing the nature and size of such threats, and ensuring that additional research that may provide a basis for resolution through policy is conducted. This department consists of both staff members employed by the Board and an advisory committee consisting of external experts from the field. Annually, PAM staff invites professionals, such as physicians, health insurers, and policy makers, to list (potential) problems involving the use of medicines. Following this consultation, PAM selects relevant topics which are then presented for approval to the Ministry of Health. Once approval is provided, PAM staff investigates the problem in zfurther detail and determines the kind of research needed to support policy making. The department then commissions the research.

In 2000, PAM identified a number of shortcomings in its activities programme. Research projects did not always generate clear answers to their questions. In addition, policy measures based on research results appeared to have only limited impact in practice. Consequently, PAM requested that we monitor and analyse their programme and provide recommendations for improvements. The results of two case studies suggested that when commissioning research, the problem definition guiding the research repeatedly shifted over time. This shift could be largely explained by differences in the interpretive frames of the various stakeholders (Moret-Hartman et al., 2007). Based on these findings, the Board implemented a number of changes in

the programme. Firstly, the Board decided that, when tendering research projects, it would state its objectives and problem definition more explicitly. Secondly, the Board asked researchers to take into account the problem definitions of policy's target populations. In this way, PAM aimed to involve researchers more closely in the process of structuring the problem and developing policy options.

In this paper, we report on our experiences with this novel approach to commissioning policy research. The case study provided here involves research that was commissioned to develop policy to achieve an optimal introduction of a new class of analgesics on the Dutch market. Research questions were: Was the new approach to commissioning research successful in terms of increasing a) the cooperation of researchers; b) the relevance of the results; and c) the effectiveness of subsequent policy measures?

Methods

Theoretical framework

First and foremost, policy is considered as a matter of co-production. This means that the development and implementation of policy is shaped by interactions between implementers of policy, members of target groups, and policy makers (Derthick, 1972; Mazmanian & Sabatier, 1983; Pressman & Wildavsky, 1973). As Elmore (1985) has argued, this implies that policy design should iterate between forward mapping (translating the social problem into a policy problem and then generating policies to deal with the problem) and backward mapping (analysing ex-ante the extent to which implementers and target groups will likely respond to these policies in ways that help resolve the social problem). According to the argumentative approach of policy analysis (Fischer, 1999; Fischer & Forester, 1993), action is driven by processes of problem setting (Schön, 1983) in which actors iterate between forward mapping and backward mapping in order to define problems and solutions that correspond with each other and the actors' normative and empirical background theories. Combining these findings with those of classical implementation theory, Grin and van de Graaf (1996a) have argued that a policy will only be effective if implementers and target populations consider the proposed policy measure meaningful. This means that the proposed policy should also make sense in the light of problems perceived by the target populations and should be consistent with their normative and empirical background theories. This implies that it may be

crucial to policy making to analyse the interpretive frames of policy's target populations.

Data collection and analysis

We monitored the process of policy analysis conducted by PAM staff, the European tender for proposals, and the PAM appraisal of the research proposals that were submitted. Data were collected by document analysis of policy analytical documents at various stages and of research proposals. Semi-structured interviews were held with PAM staff and researchers who submitted proposals. We also attended PAM meetings. Lastly, we examined the first draft of the research report. Employing triangulation (Stake, 1995) between these sources, we reconstructed and interpreted the process of policy analysis. We additionally analysed the views expressed in documents and interviews, according to a method validated elsewhere (Moret et al., 2007). In accordance with Grin et al. (1997), we distinguished participants' interpretive frames as being comprised of four layers, namely judgement of solutions, problem definitions, normative and empirical background theories, and normative preferences.

Results

Consultation of the field

In 2001, the PAM staff received indications that the use of selective cox-2 inhibitors (COXIBs) was less than optimal. COXIBs were, at that time, a new class of analgesics (non-steroidal anti-inflammatory drugs, NSAIDs) which were claimed to cause less gastrointestinal side effects than conventional NSAIDs. According to some health care professionals, however, the added value of COXIBs was rather limited. Despite this, a rapid increase in the number of prescriptions was observed shortly after the drug was introduced on the Dutch market.[3] Furthermore, in the majority of cases, high dosages were prescribed, probably because only high dosages were reimbursed. Another source in the field questioned whether these new drugs actually yield a therapeutic benefit as compared to the combined use of conventional NSAIDs and appropriate gastric protectors. The new COXIBs were considerably more expensive

[3]This was also found in data provided by the Foundation Pharmaceutical Key Numbers, which is a foundation that collects data on the use of drugs. (Farmaceutisch weekblad, 2001,136:49).

than conventional NSAIDs. Additional urges to explore this issue came from a Dutch publication which revealed that, in the Netherlands, COXIBs were frequently being prescribed for medical indications other than those for which the drug had been registered[4] (Jabaaij et al., 2001).

Policy analytic process

In January, 2002, PAM staff conducted an analysis of this issue. The employees identified three problems, namely off-label use, channelling (use of medicines in patients at high risk for side effects), and perverse incentives (the reimbursement of high dosages only). The research questions proposed by PAM staff were:

1. How are COXIBs prescribed in practice and to what degree is this prescription (in)efficient?
    - For which medical indications and patients are these drugs prescribed? What is the extent of off-label use and what arguments are used to justify off-label prescriptions?
    - Compared to both conventional NSAIDS and other patient groups, how often is COXIB being prescribed and what are the costs of co-medication (gastric protectors)?
    - What developments can be observed in patterns of side effects, especially among high risk patients?
    - Which mechanism(s) can explain the financial inefficiency, namely the reimbursement of high dosages only?
2. How do actors judge the COXIB use in practice? In order to realise an efficient prescription of new drugs, is control considered useful and necessary?
3. Which policy measures are available or should be developed to limit the inefficient use of (new) drugs in the Netherlands?

During the working group meeting we attended, we observed that emphasis was placed on the first question in the document. For PAM staff, however, the policy related component (third question) was considered most important. Furthermore, we discussed what level of detail in research questions was most appropriate. We

[4]By 2000, rofecoxib had been registered as useful for pain management in patients with osteoarthritis and celecoxib had been registered as useful for pain management in patients with osteoarthritis and rheumatoid arthritis.

suggested replacing the detailed research questions with more general questions while placing emphasis on the problem definition(s) underlying these questions.

The advisory committee discussed a modified version of this policy document (January, 2002). According to the committee, the document was ambiguous. They claimed that the focus was on the COXIBs, while the introduction of the document which was written by the PAM staff, focused on policy measures for new drugs in general. The advisory committee argued that more cases should be analysed to allow for more general conclusions on promoting the efficient use of drugs. They recommended distinguishing the last question from the preceding questions.

PAM staff preferred, however, to restrict the project to the case of COXIBs as the inclusion of more cases was likely to make the project impracticable and infeasible. They considered a detailed analysis of a case essential as it could assess the motivations that underlie prescribing practice. The final policy document served as the basis for a European tender for research (April, 2002).

Research proposals

By June, 2002, the Board had received four research proposals in response to its tender. These four proposals differed considerably with respect to the problem definition and proposed research methods. Details of the four proposals are summarised in Table 6.1.

In proposal A, the researchers proposed reconstructing, through qualitative research methods, the argumentation that underlies current COXIBs prescriptions practices. These researchers contended that insight with respect to these arguments would likely provide an answer to the question of whether COXIBs are being used (in)efficiently. In their proposal, an explicit distinction was made between off-label use and inefficient use. These researchers claimed that a drug can be efficient even when it used for other indications. The researchers considered acquiring support for the research project from stakeholders to be imperative.

Proposal B suggested using quantitative research methods to investigate whether COXIBs decrease the risk of side effects and direct medical costs. This team did not wish to assume from the onset that policy measures were necessary. In fact, they doubted whether policy measures were all necessary.

Table 6.1. Summary of research proposals

Proposal	organisation that submitted proposal	Proposed research methods
Proposal A	non-profit institute for supporting policy and decision making	1) survey among physicians, questions about their prescribing practice and motivations, 2) group interviews to investigate views on possibility of controlling the use of COXIBS. 3) collect and analyze data from available registers on the use of medicines.
Proposal B	institute for pharmaco-epidemiology and pharmaco-economics	1) Indicators for off-label use to obtain information on the argumentation behind prescriptions. 2) Multiple regression methods for comparing 'off-label prescriptions' with 'on-label prescriptions' using data from databases. 3) cost-effectiveness analyses to determine the degree of inefficient use.
Proposal C	institute for research on pharmacy	1) Sent questionnaire to patients and physicians to obtain data on considerations behind prescription (patients' complaints, previous medication, indication, etc). 2) data on previous medication from the pharmacy database. 3) ask opinion leaders among health insurers, patients, physicians, and pharmaceutical industry for their judgements on the efficiency of the prescriptions.
Proposal D	research and consultancy organisation for companies and government	1) quantitative data on the current use of COXIBS; 2) interviews with physicians, health care insurers, manufacturers, associations of physicians, pharmacists, and patients about their judgements on current prescription practice and the desirability of measures; 3) organize workshop

In proposal C, researchers offered to perform a prospective cohort study with patients who have been prescribed COXIB or conventional NSAID. This team recommended obtaining data on the reasons that underlie these prescriptions. They also suggested contacting opinion leaders among health insurers, patients, physicians, and the pharmaceutical industry to acquire insight on the beliefs regarding the efficiency of these prescriptions. It was argued that these judgements would be critical for drawing conclusions regarding utilization of COXIBs.

In proposal D, researchers proposed a project in which quantitative data would be collected on the current use of COXIBs; in addition, qualitative data would be collected by interviewing relevant actors about their opinions on current prescription practices and the need for policy measures. The main objective was to identify policy measures for preventing inefficient use of new medicines. The COXIB case was considered by this team to be a case study that could shed light on the more general issue of (in)efficient use of new drugs.

Judgement of proposals

The four proposals were assessed by the PAM staff and the PAM advisory committee. According to the advisory committee, the efficiency problem had been elaborated insufficiently in proposal A. An additional problem with this proposal was its strong reliance on support for the research project from stakeholders. According to PAM staff, this dependence had the potential to evolve into a non-decisive practice of policymaking by consensus. The PAM advisory committee further questioned whether the results would be accepted by the medical profession. Research proposal B seemed to be strongly affected by an internal research program. The first two questions were elaborated more comprehensively than had been requested. Furthermore, the researchers' perspective seemed to correspond more with medical professionals than with policy makers. In proposal C, the focus was placed more on clinical practice than on policy making. A further criticism was that the proposal defined efficiency in a way which was inconsistent with the tender. The final option, proposal D, corresponded best with the Board's intentions. The researchers had paid attention to creating sufficient support in advance. However, the Board considered the researchers' definition of the term efficiency too broad. Furthermore, the Board considered an investigation of the determinants of prescription critical for defining intervention target groups; this aspect was not included in proposal D. As a result, the researchers modified proposal D to include the aspects mentioned by the Board (October, 2002). Upon completion, the study was commissioned.

Research report

In May, 2005, the researchers reported their results. Firstly, data on the number of COXIBs prescriptions were provided. Prescription was defined as efficient if it was on-label, if use was long-term (meaning that it was not the first NSAID the patient received) and if the patient was at risk for gastro-intestinal side effects. Less than three percent of all prescriptions fulfilled all three criteria. The interviews revealed that physicians considered their prescription practices as rather restrained. The report indicated that, on average, a physician prescribes COXIBs once or twice a week. Justifications for prescribing COXIBs included: a) COXIBs are comparable to NSAIDs (and thus have broad indications for prescription) but cause less side effects; b) the prescription of one medication only increases therapy adherence; and c) the costs of a COXIB are considered equal to the costs of combining a conventional NSAID with a gastric protector. Health insurers considered off-label prescriptions problematic but lacked the data necessary to distinguish off-label from on-label prescriptions. Neither

physicians nor health care insurers were defined as 'real problem holders'. However, the researchers added that, from a societal perspective, inefficient and off-label prescription is indeed problematic.

Table 6.2. Interpretive frames of PAM staff and researchers

Actor	Proposed solution	Problem definition	Background theories	Normative values
PAM staff	Instruments for signalling and controlling prescription of new medicines by health insurance companies	Inefficient use of new analgesics (coxibs) off-label use limited evidence on effectiveness and safety	Prescription efficient if criteria for reimbursement are met; if evidence is available. New medicines prescribed in patients in whom no effects will be found Policy measures can be used to realize adequate prescription from the beginning.	Health insurance companies have important role in achieving efficient health care
Proposal A	Qualitative research on motives behind prescription. Important role in study design for supporting committee of medical professionals	HCIB* considers prescription of coxibs inefficient, but physicians may have good reasons for off label prescription	Off-label use can result from unnecessary prescriptions (no benefit) or a lack of evidence (at that time). Trust among participants needed to reconstruct motives behind prescription	Supporting power among professionals in the field
Proposal B	Quantitative research: Indicators for off label use and cost-effectiveness.	According to HCIB is off label prescription of coxibs problematic Research question was complex and not objective.	To answer HCIB's questions extensive research on a high level of expertise is needed. Theory of HCIB: too much influence of pharmaceutical industry Pharmaco-epidemiology is their expertise.	Perform objective research. Independence from government or pharmaceutical industry.
Proposal C	Quantitative research: reasons for prescribing coxibs Judgements of actors on the efficiency of current prescribing practice	Coxibs are frequently used off-label, but people can have reasons	For physicians, prescription freedom is important. Indication for prescription change over time. Insight in arguments behind prescriptions needed for policy development	Judgement on efficiency of prescription should be left to professionals in the field
Proposal D	Qualitative & quantitative research. Indicators for prescription actors' judge-ments of use and policy measures.	Inefficient prescription new medicines. HCIB's problem definition is narrow: only off label	Physicians don't consider prescribing practice problematic. Participation target population essential in policy development. Prescribing affected by pharmaceutical industries. Expertise in policy research	Answer question from commissioner

The proposed policy measures were derived from the interviews and a workshop that was incorporated into the research project. The potential policy measures generated were placed in one of the following categories: a) formal regulations; b) guidelines; or c) the provision of information and advice. All potential measures were judged based on whether the potential measure was: a) orientated to physicians; b) expected to improve quality; c) feasible; d) suitable for new drugs; e) evidence based; and f) acceptable in terms of the costs. Two policy measures in the category 'provision of information and advice' were considered most appropriate, namely including a note (warning sign) in electronic patient records when an off-label prescription is provided; and academic detailing, which is a process by which a trained health care professional/educator visits a physician at his/her own office and provides an educational intervention on a topic. The aim was to provide physicians with objective information to counterbalance information provided by pharmaceutical industries.

Although relevant information was provided, the PAM staff considered the proposed policy measures to be of limited relevance.

Interpretive frames of PAM staff and researchers

The main elements of the interpretive frames of PAM staff and research teams who submitted the various proposals are summarised in Table 6.2. According to PAM staff, the primary problem was that COXIBs were prescribed excessively and inappropriately (thus to the wrong patients) shortly after their introduction on the Dutch market. PAM contended that policy measures that are capable of signalling these excessive and/or inappropriate prescriptions patterns were necessary. Alternatively, measures, applied by health insurance companies, were needed to influence prescription practices in clinical health care practice. This reflected their view that health insurers should play a role in improving the efficiency of health care.

The researchers that submitted proposal A emphasised the need for support from medical professionals. This emphasis was derived from previous negative experiences with a former research project in which the participants, mostly physicians, did not sufficiently support the selected approach. The researchers claimed that creating an atmosphere of trust with respondents is imperative. This was considered especially important, since asking respondents to convey motivations underlying off-label prescriptions may be seen as an infringement of professionals. The researchers that submitted proposal B indicated that they did not share the problem of the Board. They claimed that *"off-label use simply occurs. The Board may*

not like it but this is how things usually go." They also considered the tender for proposals biased: *"The Board asked for confirmation of their own ideas, namely the inefficient use of COXIBs."* They emphasised the importance of a more 'objective' approach. The researchers who had submitted proposal C also had doubts about the Board's problem definition: *'According to the Board, all prescriptions that deviate from the registration of indications are inefficient. However, in practice, this issue may be more complicated.'* They further contended that physicians would consider the freedom to prescribe in whatever manner they see fit to be important. The researchers from proposal D challenged, similarly to the others, the problem definition put forth by the commissioner. Nonetheless, their primary objective was to answer the PAM staff's questions.

Discussion

The researchers appeared to be reluctant to conduct the requested research project. This reluctance was linked to a) doubt regarding whether the Board's problem definition would be shared by 'the field'; b) the desire that medical rationality should override policy objectives; c) the desire to maintain a relationship with medical professions that is based on respect for the professional autonomy and the expertise of medical professionals; and d) the conviction that any policy intervention not supported by medical practitioners will fail.

Lack of congruence in perspectives

Candidate researchers expected that 'the field' would contest the proposed policy measures. This expectation was in fact borne out by the commissioned research project. The PAM staff considered off label use of these new drugs to be problematic because of the high costs and also because they were uncertain about the long term effects and safety of the drugs. Stakeholders in the field (policy's target populations) did not share the problem definition of inadequate use claimed by the Board. Most physicians did not consider their own prescription patterns to be problematic. In addition, they suggested that the proposed solutions conflicted with the normative values of physicians and could be perceived as an attack on physicians' professional autonomy (prescription freedom).

Although most physicians did not consider their COXIB prescribing patterns to be problematic, the manufacturer of one of the COXIBs had to withdraw the drug from

the market because of severe side effects in 2004 (Merck, 2004). The fact that the risk of myocardial infarction had not been indicated earlier was subject to significant international debate (Hawkey, 2005). Moreover, the role of both the industry and the FDA in the assessment of the safety of new drugs has also been discussed (Avorn, 2006; Topol, 2004; Waller et al., 2005).

In retrospect, the Board's concerns with respect to the potential long term effects and safety of COXIBs may very well have been justified. However, we contend that the problem is attributable to the way in which the PAM staff defined the problem. PAM staff emphasised the off label use of the COXIBs. Additionally, the PAM's pilot study defined not only the problem but also the underlying intentions and the direction that needed to be taken to solve the problem. Although researchers were asked to take the target populations' problem definitions into account, as evidenced by the pilot project, it is apparent that the PAM staff was not particularly open to alternative perspectives. Finally, the procedures used for the European tender for research also affected the proceedings. The administrative process complicated contact between the commissioner and researchers. As a result, a discussion of the problem definition, the underlying argumentation, and the preferred solutions was practically impossible. Instead of focussing on off-label use, the focus could have been on the lack of evidence on the safety of these drugs, especially with respect to off label use. This kind of problem definition would have resulted in other research questions that may have been very relevant. In fact, such studies have by now been conducted (Hippisley-Cox & Coupland, 2005; Lévesque et al., 2005)

Need for interventions that are congruent with perspectives of all actors involved

The results indicated that the perspectives of both the target populations and the PAM staff were relevant. It is, however, important to note that involving target populations in policy development does not imply that the target populations control the way in which the problem is defined, nor which solutions are preferred. Involving target populations does not mean that consensus on the problem definition and on the interventions that can provide a solution has to be reached. The primary aim of involving target populations is to identify solutions that do justice to the perspectives of all actors involved.

Seeking congruence between the perspectives of actors demands another approach to research. Problems and possible solutions should be explored through dialogue and interaction (Denzin, 1989; Grin et al., 1997). This kind of process can help to: a)

assess the level of congruence between the perspectives of the actors involved; and b) stimulate a process of learning. Furthermore, the process can generate procedures by which differences in interest and power can be taken into account. An analyst can help to create conditions for mutual trust, which is essential for processes of learning and reaching agreement (Schön & Rein, 1994).

Researchers' role

In this project, researchers were not only asked to provide specific evidence but also to take the problem definitions of the policy's target populations into account. Implicitly, it was assumed that researchers should be policy advocates in order to have a policy impact. Our results suggest that researchers are not necessarily willing to adopt the role of policy advocate. In fact, researchers often prefer to maintain a neutral position towards all parties involved in the project. Koch et al. (2003) has also recognised the reluctance of researchers to take on the role of policy advocate. On the other hand, Kemp & Weehuizen (2005) have argued that policy analysts and researchers can play an important role in innovation and policy learning With this in mind, involving independent policy analysts to assess actors' perspectives and advise policy makers may be a worthwhile endeavour. That is, if one follows Jennings' (1987) distinction, while it may be misguided to treat them as 'advocates', they may play a role as 'counselor' to and between the various parties involved.

Conclusion

Involving target populations in policy research should not be equated with target populations being in charge of the problem definition and identification of solutions. The results of this case study emphasise the importance of seeking solutions that are congruent with both the perspectives of policy makers and policy's target populations. So rather than acting as an advocate, researchers may act as 'counselors'. Although it can be difficult to attain congruence between actors involved, an interactive methodology might provide opportunities to effectively reconcile all actors' interests and the power differences that may be present.

7

Diverging views on relevant policy research: a case study on improving the efficient use of drugs for pulmonary hypertension

A possible explanation for limited knowledge utilization is a shift in problem definition during policy research. The process of commissioning policy research was studied, with special emphasis on the management of shifts in problem definition. The case study consisted of policy to improve efficiency of the use of medicines for a particular patient population. We conclude that it is not so much a shift in problem definition per se that is the main problem. A shift in problem definition may be an indicator of a social learning process; however, the participants should be aware of this, and should endorse the validity of shifts in problem definition, if they occur.

M Moret-Hartman, R Reuzel, J Grin, GJ van der Wilt. Diverging views on relevant policy research: a case study on improving the efficient use of drugs for pulmonary hypertension (submitted)

Many theories on knowledge utilization have been described and discussed. Explanations for use or non-use of knowledge are sought in communications, in the action of the rational actor, or in the product of bureaucratic politics (Rich, 1991; Rich & Oh, 2000; Jacobson, 2007). Originally, knowledge utilization was regarded as a rational process, assuming that a specific decision can be attributed to the use of specific information. More recently, theories on knowledge utilization have changed from unidirectional models to interactive models, from one-time one-direction dissemination towards an ongoing relationship between producers and users (Jacobson, 2007). Alternative to the viewpoint of knowledge utilization as a rational process, knowledge utilization may be seen as a series of events affected by the type of available information and the problem-solving area (Rich, 1991). Weiss demonstrated that the direct use of evaluation (instrumental use) in decision rarely takes place (Weiss, 1980; Weiss & Bucuvalas, 1980). More often, 'use' includes processes such as the understanding of programs and gaining of new ideas and insights (conceptual use). Moreover, many policy actions do not imply concrete decisions, but involve a set of incremental, uncoordinated steps of decisions.

In the Netherlands, the Health Care Insurance Board (HCIB), an advisory body to the Ministry of Health, commissions research to obtain knowledge that is deemed necessary for decision making on the reimbursement or use of drugs. HCIB found, however, that results of projects that they had commissioned were of limited relevance to their policy tasks. A possible explanation for this might be that a shift occurs in problem definition, such that the problem that is addressed by researchers differs from the problem as originally defined by HCIB. Specifically, the problem may be redefined in such a way that it becomes amenable to available research techniques, and fits with the interests of a specific research community (Moret-Hartman et al., 2007). This suggests that awareness of this phenomenon and management of problem definition could be a means to improve relevance of policy research.

Empirical evidence on explanations for non-use and the effectiveness of interventions to enhance knowledge utilization is limited. We present the results of a case study where commissioning procedures had been adjusted to prevent shifts in problem definition. Research questions were: (1) Did research proposals correspond with the problems as perceived by both those who commissioned the study and the stakeholders in the field? (2) Were the research results translated into policy measures that provide a solution to perceived policy problems?

Case Description

This research project was commissioned by the HCIB's Department of Policy Analysis on Medicines (PAM). The HCIB is an independent organisation between the Ministry of Health and the actors in the field (physicians, health insurance companies, etc.). Its task is to coordinate the implementation and funding of health care insurance provisions. In order to support that contribution, the HCIB performs policy analysis on optimization problems. The Department of PAM is responsible for identifying developments that may jeopardise optimal pharmaceutical care, analysing the nature and size of such threats, and ensuring that further research is conducted that can provide a basis for policy decisions. An annual plan in which research topics are prioritised is submitted to the Ministry of Health for approval. The research findings provide the basis for policy advices to the Ministry of Health. A committee consisting of experts from the field, mostly in health services research, has been established and is regularly asked to provide recommendations on PAM's policy analyses, submitted research proposals, and reports to the Ministry of Health.

The Board's project dealt with two novel, expensive drugs for patients with pulmonary hypertension. Pulmonary hypertension (PH) is a rare disease of the small pulmonary arteries that results in an impoverished heart function (Gaine & Rubin, 1998). It can be caused by other diseases (secondary PH) or occur on its own (primary PH). The main symptom is shortness of breath. The disease presents in relatively young people (35-45 years old), and the life expectancy after diagnosis is short (one to two years). In November 1999, the Ministry of Health asked the HCIB for advice on the reimbursement of a new drug, epoprostenol. Concerns included its high costs and the possibility that its prescription might extend to non-licensed indications. Evidence on the drug's long-term effectiveness or its effectiveness in specific subgroups of patients was limited. Estimates of the annual costs of epoprostenol ranged from €225,000 to €450,000 per patient. As arrangements between physicians, health insurance companies, and the industry were in place, the actual charges were approximately €50,000 per patient, regardless of the amount needed. The HCIB advised the Ministry of Health 1) to limit the reimbursement of epoprostenol to PH patients with moderate to severe disease; 2) that patients should be treated by physicians specialised in PH treatment; 3) to evaluate the effects of treatment; 4) and to designate three or four hospitals as PH expert centres. In May 2002, a second drug, bosentan, was introduced into the market. The advantage of bosentan was that it

could be administered orally, unlike epoprostenol, which required intravenous administration. In February 2003, the Ministry of Health decided to restrict the reimbursement of epoprostenol and bosentan to a) patients with moderate or severe PH, b) for whom the prescription was provided by a specialist experienced in treating these PH patients c) after approval by health insurance companies.

In 2003, PAM staff prepared a policy analysis of current policy on maintaining the conditions for reimbursement of epoprostenol and bosentan and the effectiveness of these drugs.

Methods

We (JG & MM) advised PAM staff during policy analysis, provided feedback on the policy analysis draft, and analysed the received research proposals.

Advice on policy analysis was based on lessons from previous case studies and insights from the theory of the argumentative policy analysis. The argumentative approach contends that the actions of professionals can be explained by the views these actors have on a given problem and the argumentation that underlies these views (Fischer & Forester, 1993). A policy may be more effective if it is congruent with problem definitions and background theories of actors involved (Grin & van de Graaf, 1996a). Our results revealed that the problem definition guiding the work shifted repeatedly throughout the research process (Moret-Hartman et al., 2007). We contended that the shift in problem definition could be largely explained by differences in the perspectives of the actors involved, who acted on the basis of a professional frame that is shaped by education and earlier experiences (Schön 1983, Schön & Rein, 1994). Researchers redefined the problem based on their expertise and experiences. As a result, the studies failed to generate the knowledge sought by the commissioner. Furthermore, problem definitions of researchers and policy makers failed to correspond with the problems as perceived by the target populations in the field, the physicians and patients. A policy may only be effective if target populations consider the proposed policy measure as meaningful (Grin & van de Graaf, 1996a; Yanow, 2000). This implies that the cooperation of these target populations with proposed policy depends on the extent to which this policy makes sense in the light of problems perceived and does not violate normative background theories held. Thus, it seems advisable to take into account the viewpoint of a policy's target populations during the development of new policy measures.

We recommended that PAM staff a) explicate the board's problem definition and argumentation in order to prevent misinterpretations and a subsequent shift in problem definition; and b) take the problem definitions of the policy's target populations into account in order to increase the likelihood of acquiring their cooperation on policy measures. Confronted with these lessons, the PAM staff 1) explicated the purpose and the background of their research requests in their documents; and 2) asked researchers to make an inventory of problems and judgements of solutions according to the policy's target populations.

In providing feedback on the policy analysis draft, we observed that emphasis had been placed on one solution. We recommended openness to alternative solutions derived from the problems perceived by the actors involved. We also suggested that the researchers compile one inventory of problems rather than two separate inventories among health insurance companies and physicians. We argued that interaction between these target populations could be advantageous in that it could stimulate a discussion about problem definitions and widely supported solutions.

We analysed documents (policy analytical documents in different stages and research proposals) and held semi-structured interviews with the PAM staff and the researchers who submitted research proposals. All interviews were tape-recorded, summarized, and coded. We analysed the problem definition and proposed solutions for each stage of the policy analytical process and the research project that followed. In the interviews, questions focused on the perceived problems concerning care for patients with pulmonary hypertension and the underlying argumentation, in conformity to the methodology of reconstructing interpretive frames (Grin et al., 1997; Moret et al., 2007). A person's interpretive frame consists of appreciation of solutions, problem definitions, background theories, and normative preferences. Careful analysis of interpretive frames helped distinguish between what was agreed and what was disagreed. Respondent validation was conducted by sending a summary of each interview to the respective respondent for correction.

In January 2006, PAM staff was contacted by telephone to determine how far the implementation of the proposed policy measures had progressed at that point in time.

Results

Policy analysis

In 2003, the HCIB's Department of Pharmaceutical Care (CFH) recommended that PAM staff evaluate the effectiveness of both epoprostenol and bosentan. PAM staff had also received indications from health insurance companies who had experienced problems with respect to appraising reimbursements requests. The health insurance companies suspected that bosentan was being prescribed to patients who did not meet the criteria for reimbursement.

In May 2003, PAM employees prepared a policy analysis on this subject. PAM staff proposed that research be commissioned to provide the following information: 1) current policy on maintaining the conditions for reimbursement of epoprostenol and bosentan and the problems encountered by health insurance companies or physicians; 2) the possible need for adjusting the criteria for reimbursement; and 3) the therapeutic value and relative position of these drugs, their efficiency, and their actual use in clinical practice.

Interpretive frame: According to PAM staff, the problem was uncertainty about the effectiveness of these very expensive drugs (Table 7.1). A normative assumption held by PAM staff was that use in specific patient-groups is adequate only when evidence of its effectiveness in these groups is available. The most appropriate solution was to analyse the effectiveness of these medicines in subgroups of patients.

PAM staff decided to commission two projects simultaneously: Project 1 to obtain specific information on the effectiveness of the drugs epoprostenol and bosentan; Project 2 to obtain information about the problems as perceived by physicians and health insurance companies concerning the reimbursement of these drugs. Both projects were commissioned using a non-European tender. Projects with expected costs above certain levels need to be tendered in an open procedure that fits strict European criteria. The costs of these projects were estimated to be lower, providing flexibility in the organisation of the tendering process and allowing for meetings between the commissioner and researchers. In June 2003, three institutes were invited to submit their proposals. In August 2003, the HCIB received drafts of three research proposals. One proposal dealt with the evaluation of the effectiveness of the drugs and two proposals dealt with the inventory of reimbursement problems and possible solutions. Researchers were invited to present their proposals at the HCIB office. After these meetings, the researchers adapted and elaborated their proposals.

Proposal A

The proposal on the evaluation of the drugs' effectiveness (Proposal A) was submitted by clinical experts from three hospitals. The authors doubted the appropriateness of current prescriptions. They proposed to investigate the following: the effectiveness of the two drugs, the efficiency of the diagnostic and treatment processes, and the effects of the treatment on patients' quality of life. To evaluate the effectiveness of epoprostenol and bosentan, the researchers indicated that they would conduct a literature review and data analyses based on patients' records in their respective hospitals' databases. The researchers also proposed the inclusion of other novel medicines in order to assure that the findings were up to date.

Interpretive frame: According to the researchers, the problem was that the diagnostic process in other hospitals was frequently inadequate (Table 1). Diagnosis of PH and its underlying causes is difficult. Prescription of epoprostenol was limited to a few specialised centres because of complicated, intravenous administration. Bosentan, however, could be administered orally and was thus more widely accessible. As a result, more hospitals started to administer the drug themselves without referring the patient to a clinical expert centre. They claimed that it is known which subgroups of patients are likely to benefit from the new medicines. Background theories included that, given the rarity of the disease and the wide range of mechanisms that can cause it, experience and expertise is needed to diagnose PH and its underlying causes adequately. The researchers claimed that the best solution would be to limit the number of clinical centres for diagnosis and prescription of these drugs.

Proposal B

The first proposal concerning reimbursement problems (Proposal B) was submitted by a Dutch knowledge organisation. The researchers proposed an assessment of the current state of affairs on 1) guideline development by physicians and 2) the identification of a selected number of hospitals for PH diagnosis and treatment. In this proposal, both quantitative and qualitative research methods were proposed. The researchers suggested an analysis of data from registries to determine the current use of the two drugs. They also suggested conducting interviews to compile an inventory of a) the current policy according to physicians and health insurance companies; and b) the problems they perceive. Lastly, they suggested organising a workshop to discuss problems and possible solutions.

Table 7.1. Interpretive frames of PAM staff and researchers

Actor	Proposed solution	Problem definition	Background theories	Normative values
PAM staff	•Clear criteria for whom prescription should be reimbursed •Research on effectiveness of medicines •Adaptation of criteria for reimbursement	•Lack of evidence on effectiveness of the medicines in subgroups of patients •Insurance companies' difficulties in assessing reimbursement requests •No centres for PH care designated by the Ministry of Health	•Criteria for assessing the value of interventions useful for decision on what to reimburse •Use of an intervention is efficient when there is evidence of its additional value (effectiveness and costs). •A limited number of expertise centres is sufficient for a small population.	Efficiency
Proposal A	•Research on the adequacy of diagnosis and effectiveness of medicines •Diagnosis and treatment in limited number of hospitals to increase expertise	•The medicines are frequently prescribed to the wrong patients. •Diagnosis of subgroup of PH patients (underlying cause) in hospitals frequently inadequate	•Due to rarity of disease and high variability, many patients need to be seen for building up expertise. •PH is not one single disease. •Only a minority of PH patients benefit from the available medicines.	•An intervention should be available for patients who will benefit from it.
Proposal B	•Inventory of current practice & perceived problems among actors involved •Identify solutions during workshop •Assess whether current criteria for reimbursement need to be adapted	•Problems with decisions on reimbursement of two medicines •No unequivocal criteria concerning experience of prescribing physicians •No centres for PH care designated by the Ministry of Health	•Workshops will create support among actors involved.	•Reliable and robust research which answers commissioner's questions
Proposal C	•Design of solutions based on suggestions from actors involved •Solution: one clinical expert centre and a central committee of clinical experts who judge reimbursement requests	•Possible inappropriate use of new PH medicines; problems of insurance companies in assessing the reimbursement requests •Self regulation of PH expert centres by hospitals unsuccessful	•Policy research to catalyse policy implementation	•Adequate orientation of an intervention in medical practice

Interpretive frame: According to the researchers, the reason for the requested research was that health insurance companies perceived problems indicating that they lacked the knowledge to assess these reimbursement requests properly. Furthermore, the HCIB had been informed that bosentan was being prescribed to patients with problems that were not specifically linked to the therapeutic indications for epoprostenol and bosentan. Previous recommendations (developing a guideline and selecting a limited number of hospitals for PH treatment) had not been implemented.

Proposal C

The second proposal for the evaluation of reimbursement problems (Proposal C) was submitted by a university department. This proposal aimed to provide information on a) the use of the drugs and the associated costs; b) the current regulation of use and the perceived problems; and c) possible solutions to these problems. The researchers proposed to obtain data on the use of the drugs directly from health insurance companies. Questionnaires would be used to acquire data on the criteria for use, the characteristics of the clinical centres health insurers work with, and the perceived problems. Supplementary interviews were also planned. Lastly, the proposal indicated that meetings would be arranged between representatives of all actors involved.

Interpretive frame: The researcher defined the problem as a problem of two expensive drugs being licensed while the evidence on their effectiveness was still limited (Table 7.1). This researcher claimed that the optimal solution would be to establish one clinical expert centre and a central committee of clinical experts who would appraise requests for reimbursement. Preferably, the government should regulate the prescription of expensive medicines. The researcher considered self-regulation of such expert centres unlikely. Thus, the requested policy research might serve as a catalyst for implementing policy measures.

PAM's judgement of the research proposals

According to PAM staff, proposal A had not been sufficiently elaborated. Furthermore, they considered that the researchers had placed too much emphasis on the adequacy and efficiency of the clinical diagnosis. As a result, PAM staff determined that the proposal did not fulfil their needs and thus decided not to commission this project. With respect to the proposals on reimbursement problems, PAM concluded that the two proposals were very similar. Because the first proposal (Proposal B) was more detailed than the second (Proposal C), PAM decided to commission Proposal B.

Research report

In September 2004, the research report was submitted to the HCIB. The researchers had evaluated the existing policy on the use of epoprostenol and bosentan. Self-regulation by means of designating certain hospitals as expert centres had been unsuccessful. A guideline on PH care had been developed but was not yet implemented. Health insurance companies had difficulty assessing whether reimbursement criteria were met, especially with respect to the condition that only prescriptions provided by specialists be reimbursed. Because of this difficulty, health insurance companies took this condition to be met when the physician was affiliated with one of the hospitals specialising in PH treatment. At that time, multidisciplinary teams for PH care had been established in five hospitals (four university hospitals and one general hospital). The researchers recommended that these hospitals be designated as clinical expert centres. They further recommended that this designation last for a limited period (for example, three years) and that renewal should require an evaluation. Lastly, the researchers recommended that the condition that drugs should be reimbursed only if prescribed by a specialist be replaced with the condition that links reimbursement to the hospital in which the treatment was initiated.

Policy measures

In November 2004, the HCIB advised the Ministry of Health on policy measures concerning the reimbursement of epoprostenol and bosentan. Most of the recommendations provided by the researchers were included. Whereas the researchers had recommended to limit treatment to a selected number of hospitals with expertise in PH treatment, PAM staff decided not to adapt the conditions for reimbursement as such, but rather, to restrict the prescription of these drugs to all teaching hospitals (n=7) and the one general hospital that had expertise in this field by means of the Special Medical Procedures Act. Additionally, these clinical expert centres should participate in a national multidisciplinary working party that would commit them to treatment according to the protocol and to collecting and analysing prescription data at a national level. By January 2006, however, regulation by means of the Special Medical Procedures Act had not been implemented.

Interviews revealed that PAM staff preferred regulation of the prescription of these medicines by law. Changes in the Dutch health insurance system and in the HCIB's function were expected. Since January 2006, maintaining the conditions for reimbursement had become the task of health insurance companies. The HCIB's role

would be limited to decisions concerning the inclusion of medicines in the benefit package.

Discussion

In this paper, we reported on a case study in which changes in the commissioning procedure had been implemented to prevent a shift in problem definition. We found that a shift in problem definition could not be prevented in all proposals. The researchers who wrote Proposal A redefined the problem, because they considered PAM's problem definition inadequate. The researchers and the PAM staff disagreed on what kind of information was needed to improve PH care because they held different background theories. The problem definition did not shift during the tendering of Project 2. Both proposals resembled the problem as perceived by PAM staff and the research report of the elected project appeared to answer PAM's questions. Based on the findings in the research report, PAM staff advised the Ministry of Health on policy measures. Concrete policy measures were formulated although the required evidence on the effectiveness of the drugs (Project 1) had not been obtained. The proposed policy measure was to restrict reimbursement to a limited number of hospitals. These measures have, to date, not been implemented, leaving the initial problems unsolved.

Problem shifting not always undesirable

PAM staff considered the redefinition of the problem in Proposal A to be undesirable. However, the problem as defined by the researchers corresponded well with the problems mentioned in Project 2's research report. The researchers who wrote Proposal A emphasised the relevance of an adequate process of diagnosis above information on the effectiveness of the drugs in certain subgroups of patients. Proposal A was focused on the efficiency of diagnosis and selection of patients for treatment. This was rooted in their background theory, which included that expertise is needed for adequate and efficient diagnostic process.

While the rejection of this proposal reflected PAM's discontent with its focus on diagnosis rather than the effectiveness of interventions, HCIB thus also missed some useful insight on the latter. The researchers' background theory made them aware that, given the limited number of patients and high variability, it might be desirable to concentrate PH care in a limited number of clinical centres. As early as 2001, the HCIB had advised the Ministry of Health to limit the number of hospitals that may diagnose and treat PH patients. Unlike in the UK, for instance (Evans et al., 2002;

Peacock, 2003), such centres had not been designated in the Netherlands. The proposed research on the efficiency of diagnosis could have provided evidence to support the relevance of limiting the number of PH centres. This finding raises the question of whether the shift in problem definition should be defined as undesirable.

Need for learning processes

The greatest barrier in this case study was that neither the commissioner nor the researchers were open to changing, or even discussing, their divergent viewpoints. In general, it is difficult to know which problem definitions should be followed. As Schön and Rein remarked, there are no objective criteria for judging whether viewpoints are right or wrong (Schön & Rein, 1994). Alternatively, they argued, one should appeal to a shared perception for reaching agreement on a specific situation. If congruence in perspectives is lacking, a process of learning should take place (Grin & van de Graaf, 1996a). The aim is not to reach an agreement on the problem definition but on interventions that can provide a solution for the perceived problems. Solutions need to be identified that are congruent with the perspectives of all actors involved: provide a solution for their problems and do not conflict with their normative values. Therefore, problems and possible solutions should be explored through dialogue and interaction. Guba and Lincoln (1989) provided a methodology for interactive processes and considered the occurrence of learning within and between actors to be two main criteria for success. One of the conditions for this kind of learning is the existence of opportunities for actors to exchange ideas and arguments openly (Grin & van de Graaf, 1996a).

In the project reported here, the meeting between policy makers and researchers could have provided an opportunity to come to some sort of agreement on measures for improving the adequate use of epoprostenol and bosentan. Previous reports from case studies stressed the need for communication and close cooperation between researchers and policy makers to improve the policy process (Moynihan, 2004; Milbank, 2001). PAM staff decided to commission two projects simultaneously. In Project 1, researchers were asked to obtain specific information. In Project 2, researchers were asked to assess the viewpoints of the policy's target populations. Physicians specialised in PH care wrote research Proposal A and thus were members of the policy's target population. Because of limited time and expertise, PAM staff asked researchers to take into account the viewpoints of actors in the field in a parallel project (Project 2). Ideally, prior to commissioning research, the problems of

stakeholders in the field should be explicitly incorporated during the process of policy analysis. Interactive methods for policy analysis can be useful for this purpose.

Knowledge utilization

HCIB employees considered that facts about the effectiveness of an intervention would be sufficient for decision making on this subject. However, policy makers have to deal with a complex decision-making process, which is a challenge for both policy makers and researchers who provide knowledge to support those decisions. Researchers need to anticipate different interpretations of research needs. Policy makers need to anticipate the different perspectives of actors whose cooperation is needed. As Weiss noted (1993), evaluation takes place within a political environment. When translating the results into policy measures and implementing these policy measures, the cooperation of several actors is needed. Results of an evaluation are not used by changing the judgment of interventions, but utilization mostly involves a change in the problem definition or background theories. Changes in these deeper levels of one's interpretative frame are more difficult to realize (Grin & van de Graaf 1996a). Because information need to compete with experiential knowledge as well as pathophysiological insights, instrumental use, in terms of the direct application of new evidence, is unlikely to occur. In this light, knowledge transfer should be regarded as a process of learning.

Conclusion

In hindsight, too much emphasis was put on preventing a shift in problem definition. Assumption was that preventing a shift in problem definition could enhance the utilization of knowledge from policy research. The main problem was that neither the commissioner nor the researchers were open to changing, or even discussing, their divergent viewpoints. For identifying interventions that are congruent with the perceived problems and normative values of all actors involved, problems and possible solutions should be explored through dialogue and interaction. A shift in problem definitions may be necessary for solving problems.

8

Interactive problem structuring to develop meaningful interventions and research questions: a case study on emergency hospital care for patients after auto-intoxication

Many notable researchers have emphasised the importance of usefulness in policy research. The quality of an evaluation largely depends on the quality of the underlying problem definition and the quality of the problem definition often improves as stakeholder involvement increases. By means of a study on the management of attempted suicides by drug overdose, we explored whether an interactive methodology could be adequate for problem structuring. Despite the fact that a high level of care is often unnecessary, these patients are often admitted to the internal ward or intensive care unit. To solve the efficiency-problem, some physicians proposed a pilot study to evaluate the effectiveness of a six-hour observation unit within the emergency department. Although evaluating such an unit was technically feasible, we felt uncertain about the appropriateness of this intervention and the way the underlying problem was structured. Results of the study demonstrated that the use of interactive problem structuring made the divergent problem definitions and underlying normative values transparent. Using this model, the information resulting from our research efforts is expected to be more useful and available to the stakeholders involved and therefore more effective. Moreover, the study provided us with a deeper understanding of potential resistance from important stakeholders to the implementation of such a unit.

M Moret-Hartman, R Reuzel, J Grin, GJ van der Wilt. Interactive problem structuring to develop meaningful interventions and research questions: a case study on emergency hospital care for patients after auto-intoxication (submitted)

Scholars such as Weiss, Rossi, and Dunn have emphasised usefulness as an important aspect of policy research. In their view, evaluation research should be useful for policy making (Dunn, 2004; Rossi et al., 1999; Weiss, 1993). Following this reasoning, the quality of evaluation largely depends on the quality of the underlying problem definition and the quality of the problem definition often improves with the increasing involvement of the stakeholders. William Dunn even goes so far as to define problem structuring as "a recurring phase of policy inquiry in which analysts search among the competing problem formulations of different stakeholders" (Dunn, 2004: 72). Involving stakeholders in problem structuring is important because:

- Involving stakeholders leads to a more comprehensive view of the problem situation and therefore diminishes the risk of an error of the third type: solving the wrong problem.
- It enables the policy analyst to design policy solutions that conform to stakeholder views and therefore increases the chances of successful implementation.

The idea that stakeholder involvement increases successful implementation has been confirmed by Frank Fischer and others, who argue that a policy is likely to fail if it does not represent a solution to problems as perceived by stakeholders and if it violates these stakeholders' normative preferences (Fischer & Forester, 1993; Fischer, 1999; Grin & van de Graaf, 1996a; Grin & van de Graaf, 1998). The evaluation driven development and implementation of a policy is shaped by interactions between all stakeholders, including those who create the policy, those who implement the policy, and those targeted by the policy (Derthick, 1972; Mazmanian & Sabatier, 1983; Pressman & Wildavsky, 1973). However, difficulties arise when stakeholders have different, perhaps even incompatible, views. Moreover, these views may be misguided, poorly supported by evidence, or otherwise less valid.

In theory, fourth generation evaluation serves the dual role of involving stakeholders and eliciting their views, and provides a means to deal with misguided or incompatible views. As conceived by Guba and Lincoln (1989), its backbone is the so-called hermeneutic-dialectic circle of interviews. Participants are interviewed successively. In each interview, the evaluator elicits views and exchanges claims, concerns, and issues. Having interviewed the last participant, the evaluator then returns to the first interviewee, presenting them with the other interviewees' views, thus moderating a continuous process of 'vicarious learning' in which alternative

views are negotiated, misconceptions are filtered out, and a consensus develops about problem definitions and solutions.

In this article, we will not address the many intangibles of fourth generation evaluation such as participant selection, closure, power differences, and so forth. Rather, we regard fourth generation evaluation as a problem structuring process and, on the basis of a case study involving the management of attempted suicides by drug overdose; we will answer the following questions:

1. What problem definition emerges from a fourth generation approach to problem structuring, as compared to the problem definition as first presented by the person initiating the problem investigation?
2. What difficulties are encountered that specifically pertain to the fourth generation approach to problem structuring?

In this article, we will report on the use of an interactive methodology to structure a problem concerning the clinical treatment of patients who attempted suicide by auto intoxication. Our aim was to explore, on the basis of this case study, whether an interactive approach offers an improvement in problem structuring, as compared to conventional approaches to evaluation in health care that are usually restricted to assessing an intervention's effectiveness and efficiency. Initially, we will present the results of this preliminary study. This will be followed by contrasting our results with the expected outcomes of a more conventional approach to problem structuring that does not feature an interactive problem structuring process.

Case description

The hospital involved in this study is a university-based teaching hospital. Each year, approximately 150 persons present themselves at, or are brought to, the emergency department following an attempted suicide using drugs. Some of the medical specialists, particularly those in internal medicine, felt that the care provided to these patients was inefficient. They perceived that while these patients are often admitted to the internal ward or intensive care unit it is rarely medically necessary. As a result, these patients occupy hospital beds that other patients may need more urgently, and thus hospital resources are used inefficiently. Meulendijks et al. (2003), after analysing the type and amount of drugs used, physiological parameters of the patients, and the disposition of hospital admissions, confirmed that 60% of auto

intoxicant patients are admitted to an internal medicine ward or intensive care unit. Of these patients, only 40% receive treatment. In most cases, this treatment is initiated within one hour of the patient's presentation. In the remaining patients, treatment is started within 4.5 hours after their presentation. Based on retrospective data, the criteria that predict treatment courses based on the clinical signs were identified (Brett et al., 1987; Meulendijks et al., 2003).

To solve the efficiency-problem, some physicians suggested that patients be observed in the emergency department in order to determine whether admission to the general or intensive care ward is necessary. In the case of an adverse event, the patient would be transferred to the internal medicine ward or intensive care unit. If no such events occur within a few hours' time, the patient would be discharged.

Our department (Medical Technology Assessment) provides general methodological support to clinical departments. We were asked to assist in conducting an assessment of a six-hour observation unit. Although evaluating such an assessment unit was feasible, we felt uncertain about the appropriateness of the six-hour observation unit and the way the underlying problem was structured. We doubted that all stakeholders involved would find the results of such an assessment - whatever they would be - compelling enough to make them co-operate with whatever changes the results suggested.

In order to properly structure the problem regarding the care of patients after auto intoxication, we suggested that a preliminary study be conducted. This involved an interactive evaluation where various stakeholders participating in the project discussed their views. Our aims were:

1. To reconstruct stakeholders' views towards hospital care for auto intoxicated patients, and on this basis establish a shared problem definition,
2. To identify solutions to problems perceived, as well as the conditions under which these solutions would be met with widespread support; and
3. To identify research questions that appear to fit the problem definition and are amenable to research.

Methods: Interactive evaluation

We performed an interactive evaluation based on the fourth generation methodology elaborated by Guba and Lincoln (1989). Interactive methodology involves a multiple series of open-ended interviews with stakeholders, where the interviewer exchanges

claims, concerns, and issues between the respondents (Grin et al., 1997). We first interviewed two actors (an internist and a psychiatrist) who had been involved in the retrospective study of the treatment of auto-intoxicated patients. At the end of the interviews, we asked them to suggest the names of other stakeholders, preferably with opposing views towards the auto-intoxication policy. This 'snowball' led us to interview psychiatrists (n=2), internists (n=2), nurses (n=2), intensive care specialists (n=3), general practitioners (n=2), a psychiatric critical care worker (n=1), a clinical psychologist (n=1), an ER physician (in training for internist) (n=1), a pharmacologist (n=1), the head of the emergency room (n=1), physicians from a general practitioner centre (medical emergency outside normal working hours) (n=2), a representative of an organisation for (ex)suicidal patients (n=1), and a patient (n=1).

During the interviews, which lasted approximately one hour, respondents were invited to elaborate on the auto-intoxication policy, particularly the problems as they perceived them, and solutions they considered appropriate. In the second part of the interview, the interviewer asked them to respond to the viewpoints of other stakeholders that the interviewer introduced to them anonymously. We sent interview summaries to the respondents for validation. After the first series of interviews, an overview of all the interviews was made and circulated among the respondents. In this summary, we emphasised the criteria for adequate care for patients after intoxication. During the second series of interviews, we asked respondents to respond to this summary.

Interviews were transcribed verbatim, and then coded and summarised. We used the reconstructing interpretative frames method to analyse the stakeholders' reasoning (Grin et al. 1997; Moret-Hartman et al., 2007). An interpretative frame is the interviewer's reconstruction of a respondent's view, featuring four 'layers' of problem definitions, proposed solutions, empirical and normative background theories, and normative preferences (Grin & van de Graaf, 1996b; Schön, 1983). This method is a suitable tool for analysis, particularly in cases where it is important to associate a respondent's problem definitions and solutions with his or her background theories and preferences. Interpretative frames helped us to assess whether proposed solutions matched problems as perceived by stakeholders and did not violate these stakeholders' normative preferences. In this way, we merged Fischer's theory of argumentative policy analysis with the fourth generation evaluation paradigm.

Furthermore, we scrutinised the responses provided by the interviewees against the literature available on the subject. Using this data and research, we wrote a draft of the final report that we also circulated among the respondents for comments.

Results: Interactive evaluation

We approached 25 stakeholders for interviews. Two persons declined the invitation and three persons could not be reached to make an appointment. Twenty stakeholders participated in this study. Thirteen of them were interviewed twice. Of those who were interviewed once, three respondents replied to the overall summary by mail or telephone. Two respondents indicated that they felt less involved with the situation being addressed. Two respondents could not be reached for a second interview. A summary of the respondents' interpretative frames is presented in Table 8.1. Below, we present a summary of the discussions between the most important respondents.

Internist

According to an internist, the main problem was that many intoxicated patients were being admitted to a general ward or intensive care unit, while, in retrospect, admission often appeared unwarranted. These hospital beds are scarce and expensive. Unnecessary admissions might result from the unreliability of patients' anamneses. It is often unclear what patients have ingested, how much, and how long ago. Patients are sometimes somnolent when they arrive at the hospital, or under- or overestimate the amount of drugs ingested, sometimes deliberately. As a result, it is difficult to assess the risk of complications. However, most auto-intoxications are considered non-life-threatening. Furthermore, most drugs reach peak blood levels within a few hours. On the basis of this, the internist argued that the solution to the efficiency problem is to monitor the patient for a few hours in the emergency department.

Table 8.1. Stakeholders' judgement towards a 6-hour observational unit, their problem definitions and underlying argumentatio

Actor	Judgement of solution	Problem definition	Background theories	Normative values
Internist	+	Patients' risk is uncertain Admission of intoxicated patients is not always necessary Current hospitalisation policy is inefficient.	Current policy on admission is too defensive. Usually no severe complications after intoxication. Complications can be predicted using clinical signs Hospitalisation at somatic unit adequate if physical complications can be expected	Internist should assess physical risk and initiate needed therapy Risk assessment based on actual data
Physician at ED**	+ if central monitoring is available	Anamnesis is unreliable. Patients risk uncertain Observation in general ward difficult at night: few nurses and no technical apparatus to monitor patients.	Suicidal act is often a cry for attention Many patients have psychiatric disorder Some patients deliberately over- or underestimate ingested amount. Patient with psychosocial problems not at internal ward.	assess physical risk and initiate needed therapy
Intensive care specialist at teaching hospital	+ also other interventions needed	Sometimes admission to intensive care unit unnecessary. Risk estimated by relatively inexperienced physicians Intoxicated patients not seen by senior physicians.	Physicians in ER are in training for internist Intensive care specialist is specialised in care for critically ill patients Lists of possible complications from national expertise centre too theoretical; might not occur in practice.	Use available expertise. Only patients whose vital signs are threatened need to be admitted at IC. Risk assessment based on experiences in clinical practice.
Intensive care specialist at general hospital	+/- medium care unit available	Uncertainty about risks; both somatically and in terms of repetition of a suicidal act. No problems perceived in current situation (patients at medium care unit)	Patient can over- or underestimate amount of ingested medicines For assessing risk of complications, patients should be strictly monitored	Safety Exclude risks as much as possible
Nurse at ED	± if separated and have desinated team of nurses	Patients can be aggressive or agitated. Care for these patients is laborious; patients cannot be left unattended. In an observational unit, more routine patient care tasks need to be done.	Suicidal act can be a cry for attention. Speaking with patients about problems not considered as nursing task.	Working in specialised care and high turnover.
Head nurses ED	± if at separated unit	Patient flow at ED is suboptimal. Physicians at ED have little experience.	High turn over of patients at ER. Usually, intoxications are of comparatively short duration. Sometimes observation can avoid hospitalisation.	Efficiency and adequate patient flow.

Actor	Judgement of solution	Problem definition	Background theories	Normative values
Pharmacologist	± measurements of ingested drugs more adequate	Anamnesis is unreliable; uncertain what was ingested Hospital has no clear structure for the disposition of intoxicated patients.	For assessing the risk for complications one should know what and how much has been ingested Knowledge available on toxicity and complications for specific medicines	Use available expertise. Risk assessment based on evidence of the toxicity of drugs.
Psychiatrist at teaching hospital	+	If patients are hospitalized they are seen the next morning; patient less willing to talk about the event and family and friends usually not present	Hospitalisation not optimal for patients' social situation.	Offer patients the care they need. Aim of consultation: to understand what happened
Psychiatrist at general hospital	- no opportunity for psychiatric consultation at convenient moment	Patients not only have somatic problem but also may have psychosocial crisis. Patients will not remember things discussed shortly after event. Psychosocial or psychiatric care not well arranged.	After intoxication amnesia can occur. Consultation also a means to arrange aftercare. A suicidal event provides the opportunity to evaluate patient situation and therapy. Patients have limited problem solving ability	Prevent repetition of suicide attempt Aim of consultation: arrange adequate mental care
Clinical psychologist	- Other interventions needed: Introduction of a casemanager First appointment aftercare inside the hospital. Active outreach.	There is little attention given to patients who attempt suicide. Connection between acute care in the hospital and aftercare is suboptimal.	Continuity in care and care giver can prevent repetition of suicide attempt. Active approach from hospital and care givers needed in preventing repetition Patients who attempt suicide have increased vulnerability Discussing suicidal tendency can be difficult and threatening for care givers.	Prevent repetition of suicide attempt. Continuity of health care
General practitioner	- Other interventions needed:Initiation of care in co-operation with general practitioner	Co-ordination of mental after-care is suboptimal	Psychiatrists don't have an adequate overview of care patients are receiving. Insufficient communication between primary health care and hospital care	General practitioner is central dossier holder. Continuity of health care
Employee from organisation supporting (ex)suicidal patients	- Other interventions needed: Central triage on acute (psycho)social problems	Insufficient attention towards patient problems. Accessibility of psychosocial care in acute situations is limited.	Suicidal act results not only from psychiatric disorder but also often due to psychosocial problems. If problems are not solved, repetition of suicidal attempt. Patients more likely to talk about problems with fellow sufferers Psychiatrist not most adequate caregiver to talk to these patients.	Provide solutions to patients' problems. Give patients the ability to speak about their problems.

* '+' supporting a 6-hour observational unit; '±' supporting if conditions are met; '-' opponent

** ED = emergency department

ER physician

The ER physician agreed that an observational unit could be useful, not only for patients suffering from auto-intoxication, but also for other classes of patients, such as those with head traumas. In fact, she argued that the general ward is an inadequate location for monitoring a patient's physical state due to the lack of apparatus to measure blood pressure, heart rate, and respiration. Moreover, most auto-intoxicant patients arrive during the evening or at night when only a few nurses are present in the general ward. However, the ER physician pointed out another aspect surrounding the care of suicidal patients. Some patients become aggressive, and do not remain in their beds. Furthermore, the ER physician claimed that many patients who intoxicate themselves have psychiatric disorders. Therefore, she argued, patients who are somatically safe should be admitted to a psychiatric ward. An observational unit for this class of patients requires additional measures to ensure the safety of both patients and medical personnel.

Intensive care specialist

An intensive care specialist agreed that the existing hospitalisation policy concerning auto intoxication was inefficient. Due to the uncertainty with respect to drug intake and risk, patients might be admitted to an intensive care unit too readily. However, an observation unit in the ER does not necessarily solve this problem. In the ER, patients are commonly seen by newly graduated physicians who have limited expertise and experience. Therefore, the intensive care specialist emphasised that, besides the introduction of an observational ward, persons involved in the development of hospital policies should develop a clinical protocol concerning diagnostic interventions to be performed, and experts to be consulted.

Emergency department nurse

ER nurses rejected the implementation of observation unit, should this unit be located in the ER. They acknowledged the existing efficiency problem, but raised several objections against the proposed solution. First, they argued that some patients who have intoxicated themselves behave aggressively or become agitated, thus disturbing the care of other patients. Secondly, most ER nurses are specially trained to work in a dynamic, acute care environment. They prefer not to perform routine nursing activities such as washing and feeding patients. For these reasons, an observational ward should be located in a separate ward and a separate team of nurses should be established to care for these patients.

Clinical pharmacist

Thus far, all respondents agreed that the existing care for auto intoxicated patients was inefficient, framing the problem similarly, but rather arguing that an observational unit could be appropriate only under certain conditions. The first interviewee to frame the problem differently was a clinical pharmacist. According to him, the problem was that patients' anamneses were unreliable and that referring physicians lack the knowledge to adequately assess a patient's prognosis. A better solution would be to measure the type or amount of drug ingested by the patient in the patient's blood or urine. Using this information, a clinical pharmacist could estimate the risk of complications. As well, the clinical pharmacist emphasised that additional modifications were also needed to improve the care for intoxicated patients. First and foremost, a multidisciplinary team should be established to deal with these patients. Furthermore, physician knowledge regarding the diagnosis, and treatment of auto intoxicants should be enhanced through education.

Psychiatrist

A psychiatrist also considered the implementation of an observational ward useful, but like the pharmacist, he departed from another problem definition. If patients have been hospitalised, the psychiatrist is called in for an evaluation the following morning. The psychiatrist prefers to visit the patient as soon as possible after the event, because he considers the patient more willing to speak about the situation shortly after the event. Secondly, family members or friends are more likely to be present in the evening, when most auto intoxications occur. Family or friends often have useful information regarding the patient. The aim of the consultation is to assess the suicidal risk and to understand what happened. According to the psychiatrist, an observational unit would provide an opportunity to talk to the patient soon after the intoxication.

However, a psychiatrist from a general hospital considered it useless to have a psychiatric consultation shortly after the event. Patients can be somnolent upon arrival and later will not remember what has been discussed during the consultation. Many patients have ingested benzodiazepines which may cause amnesia. According to this psychiatrist, the aim of a psychiatric consultation is to assess the suicidal risk and arrange adequate care. In his view, it is best to admit patients to a medium care ward, where they are visited by a psychiatrist the next morning. He agreed that the general ward was unsuitable, because patients could not be monitored adequately.

The patients' physical state may deteriorate and there is a risk of repetition of the suicidal act.

Clinical psychologist

According to a clinical psychologist, many caregivers have little affinity with these patients and get irritated with them. Caregivers do not always understand when relatively small problems result in suicidal acts. A negative attitude exhibited by a caregiver affects the patient's willingness to talk about his or her problems with a psychiatrist. Another problem that the psychologist raised was that the connection between hospital care and after-care for these patients is inadequate. Some patients will not show up at outpatient consultations or community mental care clinics, which is problematic as continuous care is considered necessary to prevent a repetition of the suicidal act. In order to improve the continuity between acute hospital care and after-care, active outreach measures should be used to encourage patients to keep appointments and caregivers could come to the hospital to meet the patients.

General practitioner

The general practitioner also perceived problems concerning the co-ordination of after-care. Often, the psychiatrist arranges a new therapy, which sometimes interferes with therapy a patient is already receiving. Furthermore, the psychiatrist is not always adequately informed about previous, often unsuccessful, interventions. Mental after-care should be arranged in co-ordination with other caregivers, including the patient's general practitioner, who keeps the general patient file.

Patient

An employee from an organisation providing support to (post)suicide attempt patients agreed that not all patients receive adequate care. Many patients do not want to talk with a psychiatrist about their problems. They do not want to be labelled as a psychiatric patient, or have grown disillusioned with psychiatric care. Although some patients have a psychiatric disorder, the suicidal act usually results from psychological and social problems.

A patient who has been admitted to a psychiatric ward after treatment in the emergency department does not remember anything from the emergency department. The psychiatric ward provides him with a feeling of safety and security.

At this ward, patients have the opportunity to talk about problems with some fellow-sufferers, who empathise with the patients' situation.

Conclusions

Based on the interviews, we drew the following conclusions:

1. The introduction of a six-hour observation unit was well supported within the university hospital, under the condition that it is located in a separate unit and equipped with its own, well-trained staff.
2. In order to improve patient care, arrangements should be made between departments in the hospital and between hospital and primary (mental) health care. These arrangements should be documented in (local) guidelines.
3. Currently, the psychiatric consultation targets different goals: acute care and non-acute care. An observational unit might cause problems with non-acute care, such as the arrangement of after-care. Possibly, the psychiatric consultation could be split up with a consultation shortly after the attempted suicide, and a later consultation, perhaps on the next day. A disadvantage of this is that it is time consuming to see patients twice. Furthermore, it is uncertain whether a patient will show up for the second consultation once discharged.
4. Relevant research questions were determined that concerned the efficiency of an observational unit, as well as several issues concerning psychiatric care. Would the implementation of an observational unit decrease the number of unnecessary hospitalisations? Would the implementation of an observational ward decrease the number of patients who are consulted by a psychiatrist? Would patients remember or keep appointments that are made at the observational ward? Would the implementation of an observational ward decrease the number of patients who repeat a suicide attempt?

Outcome: Interactive evaluation versus conventional approach

To assess whether our interactive problem structuring process had been useful, we needed to compare our results with the results we would have obtained, had we used a conventional approach. While the actual results resulting from a conventional approach are based on conjecture, according to the internist who first presented the problem to us, it is safe to assume that the problem would have been framed in terms

of efficiency and unnecessary hospitalisations. Research would have focussed on the efficiency of a six-hour observation at the ER. Already, we believe that we may point out a few important differences.

Efficiency research improved

Despite the fact that efficiency remains an important issue, and efficiency research remains relevant, the interactive process led us to understand how to improve the validity and usefulness of the outcomes. Efficiency research, especially when it is comparative, rests on the assumption that both comparators are optimised and stable, i.e. that they are no longer 'moving targets'. As the interviews have clearly shown, in our instance, this was not the case. Therefore, efficiency research is as yet premature.

The interactive problem structuring process yielded a broader insight into the conditions under which a six-hour observation unit could be successful. In particular, the arrangements between the departments involved, and the role of the psychiatrist likely contribute to optimised care. The discussion on these issues has led to research questions that, at this stage, appear to be more relevant than the question of efficiency.

The first question raised through our research concerned the uncertainty surrounding the patients' anamneses and the associated risks. To ameliorate this problem, two possible strategies were proposed: a) the introduction of an observation unit; and/or b) a laboratory analysis and titre of patient blood and urine to determine the kind and amounts of the drugs ingested. Few studies have been performed that have followed a group of consecutive auto-intoxicant patients post emergency treatment to calculate the risk for complications (Hollander et al., 1999; Meulendijks et al., 2003). A six-hour period for monitoring patients is often mentioned in articles or books on clinical toxicology based on pharmacological data. Most patients have ingested benzodiazepines that are relatively non-life threatening when taken alone (Meulendijks et al., 2003). In intoxicated patients, complications occurred within a few hours after ingestion. (Arranto et al., 2003; Liebelt et al., 1997). In other medical fields, observational units resulted in a reduction in the number of hospitalisations (Martinez et al., 2001). Information from blood and urine drug screening can also affect the decision regarding hospitalisation of patients after intoxication (Fabbri et al., 2003). Identification of intoxicants, drug levels and patient observation are considered to be complementary approaches (Kellerman et al., 1988; Perrone et al., 2001). In only a few cases, however, did drug measurements result in a change in

clinical practice (Kellerman et al., 1987). Apparently, in most cases, observation of patients will suffice. Nevertheless, some drugs or combinations of drugs are associated with a high risk of complications (tricyclic antidepressants), or have complications that might occur long after ingestion (acetaminophen) (Neeleman & Wessely, 1997). It is important to ascertain whether these medicines have been ingested.

The second question dealt with the most effective time for a psychiatric consultation to take place. Previous studies have demonstrated lower immediate and delayed recall in patients or healthy volunteers who had ingested benzodiazepines (Verwey et al., 2000; Verwey et al., 2005). These data suggest that after 12 hours, and even after 24 hours, patients remember little (Verwey et al., 2000). Should patients be seen shortly after arrival (a few hours after ingestion) or the next day (less than 12 hours after ingestion)? Are the recall tests used in the above-mentioned studies comparable to the information discussed during a psychiatric consultation, and are the results transferable to the situation in our hospital? However, even patients who have used different methods for their suicide attempts have had problems recalling what they have discussed with their psychiatrist (Kerkhof, 1985).

Different framing and resistance from stakeholders

Another significant finding from our research is that a six-hour observation in the ER would meet with resistance from the ER nurses. This is a critically important insight, as the nurses' resistance would have most likely led to poorer quality care. Acknowledging this factor and adjusting the plans for a six-hour observation unit before an efficiency research project is conducted will not only ensure a better quality of patient care, but ensure the validity and usefulness of the research.

Our study supported the theory of argumentative policy analysis in that different stakeholders frame the situation with respect to auto-intoxicated patients differently. Importantly, the interactive process not only revealed differing problem definitions and perceived solutions, but has also enabled us to uncover and describe deeper levels of the respondents' interpretative frames, i.e. normative and empirical background theories, and preferences. Understanding why a stakeholder frames a problem the way he or she does is important for determining if these stakeholders will cooperate with implementing particular solutions. For example, factors such as the lack of affinity with the patients felt by some stakeholders, or the normative assumption that the hospital should assume full responsibility for the care of suicidal patients (not only for the treatment of their physical intoxication), determined the

view of stakeholders toward appropriate and acceptable changes in current practice. It is difficult to imagine how failing to acknowledge these factors would not lead to conflict, as the theory predicts, and how research that inevitably conforms to only one particular view could be useful in such a context.

Auto intoxication in broader perspective

The interactive process has lead us to recognise that in our study there was a marked difference between the hospital and patient perspectives regarding the care for patients after auto intoxication. The patient perspective dealt with offering integrated care to optimise the pathway of care. These perspectives do not necessarily contradict each other, but the patient perspective does lead to different research questions such as those concerning psychiatric consultations and aftercare, aside from the question of efficiency. Interviews revealed that in addition to the uncertainty surrounding the ingested drugs and the risk of complications, other problems might threaten patient care. Besides the treatment of the somatic problem, consideration should be paid to the psychosocial or psychiatric problems underlying the suicide attempt. Many studies on the management of these patients deal with either somatic (Jones, 1999) or psychosocial care (Kapur et al., 2004; Kerkhof, 1985). Although these problems can be discussed separately, our study demonstrated that the solutions to these problems may interfere with each other. Some respondents emphasised the need for measures to improve the continuity between acute care in the hospital and (mental) after-care. Others discussed the possible consequences of a six-hour observation unit on the usefulness of the psychiatric consultation.

Discussion

The theory of argumentative policy analysis explains that a policy fails when it does not match stakeholder views. We have embedded this theory in fourth generation methodology, using the reconstruction of interpretative frames as our chief analytical tool, to develop an interactive problem structuring process. We have used this process to re-structure the problem of managing patients after auto intoxication, which we assumed to be an ill-structured problem. That is, we assumed that the actors disagreed on what information was needed and what norms were at stake. In such cases, the analyst has an important role in defining the problem (Dunn, 1994; Hisschemöller, 1993).

The interactive methodology was useful in that it effectively made the divergent problem definitions and underlying normative values transparent. Compared to the traditional approach (evaluation in health care mostly assumes that the problem concerns efficiency), problem structuring resulted in the identification of alternative interventions and research questions. Even though efficiency remained an important issue, the interactive process demonstrated that efficiency research was premature. Questions that appeared to be more relevant were linked to the assumptions actors had about the proposed intervention.

Only a few studies were found where the outcomes of interactive methods were compared to the results of conventional methods. One problem with such comparison is that the methodologies are based on different paradigms, each having its own criteria for quality. Although it might be difficult to define which outcomes are best, the relevance of the outcomes may differ. An interactive methodology has been used in the evaluation of cochlear implants (Reuzel, 2004). This interactive evaluation resulted in concrete policy recommendations supported by a wide range of stakeholders. A previous study assessing the efficiency of cochlear implants did not provide all the relevant information needed for policy decisions on this subject and was strongly contested. As well, it is difficult to compare the outcomes of studies done in different situations. Grin and Hoppe (2000), compared the structuring of car mobility problems in three technology assessment projects. They concluded that not only the methodology used (interactive methods or not) affected the outcome, but also institutional characteristics and context.

Validity of the interactive approach

With regard to the validity of our approach, we would like to address the following issues.

First, although we believe that the problem concerning auto intoxication is now better structured than it was, we cannot tell with certainty whether the process has been exhaustive and the problem structure may be further enhanced, perhaps by involving still more participants.

Secondly, the role of the interviewer\moderator is critically important. Without the interviewer's creative role, the process probably would not have had enough impetus to lead to a desired problem definition. However, in order to play this role, the interviewer assumes the power to influence the process. As the responses to our report are fairly favourable, we feel that the role of the evaluator has not been problematic in

this case. However, the issue deserves more attention. This has been addressed it in greater detail elsewhere (Reuzel et al., 2007).

Thirdly, it was difficult to bring the process to a closure, that is, to define the boundaries of the problem. For example, was it valid to leave issues such as the aftercare for patients and arrangements between caregivers inside and outside the hospital out of the problem definition? We feel that it was. However, this is far from self-evident, especially if the establishment of an observation unit has consequences for the issues mentioned, as some participants argued.

Fourthly, the process has been too short to leave sufficient room for 'vicarious learning' (Guba & Lincoln, 1989). Moreover, learning processes were difficult to trace, as during the first series of interviews some respondents anticipated the viewpoints of other stakeholders, which were partially known to them. Usually, participants in an interactive process learn, and on this basis adjust their views at all levels of their interpretative frames. Guba and Lincoln trust that vicarious learning eventually leads to congruence between the views of participants, or even consensus. Regarding this subject, it is important to note the disagreement between the two psychiatrists on when best to consult with the patient. We have tried to settle this issue using the evidence available in the literature, but it remains uncertain whether both psychiatrists could have lived with the outcome, should they both have worked in our hospital. In this instance, the time allotted for the study has been too short to reach congruence or consensus. Therefore, we have marked the issue as a further research question and a relevant part of the problem. We considered this lack of congruence valuable as it allowed us to identify this research question and demonstrate the usefulness of our approach. We have not used the interactive process to reach agreement about solutions for the care of auto intoxicated patients, but for problem structuring.

Conclusion

Despite questions of validity, we believe that the interactive problem structuring process has been a valuable project. It has significantly broadened and improved our insight into the subject under investigation. As a result, our research efforts will be better geared to information needs from persons involved and therefore more useful. Moreover, we have gained a deeper understanding of potential resistance from important stakeholders and now have the opportunity to acknowledge their reasoning.

Acknowledgements

We would like to thank all respondents who participated in this research. This project was funded by the institute of Science and Society of the Radboud University Nijmegen. Thanks to Renske Wentzel for her analysis of the literature on the risk of complications and the amnesic effects of benzodiazepines.

9

Discussion: Problem based Health Technology Assessment

Currently, the means by which the limited usability of Health Technology Assessments (HTA) in policy making and clinical practice can be improved is sought in strategies that advocate the active implementation of research findings. In this thesis, an alternative hypothesis for improving the limited usability of HTA was proposed. A possible explanation for the limited impact of HTAs is that they do not sufficiently answer the questions deemed important by their potential users, namely policy makers and health care professionals. This hypothesis, based on the theory of argumentative policy analysis, contended that policy makers and professionals often have different perspectives on the situation and therefore assess interventions differently. If an evaluation of an intervention is to be relevant for policy making, it should not be limited to the mere assessment of the intervention's outcome. Rather, it must also explore and discuss possible differences in the relevant actors' problem definitions, their underlying assumptions, and their normative values (Fischer, 1999). In this thesis, a number of case studies have been analysed in an effort to explore whether the insights put forth by the argumentative approach to policy analysis can provide an explanation for the successes and failures of not only the proceedings that are part and parcel to HTA but also the policy making that follows an HTA.

The research questions included:

1. In cases where HTA research or the implementation of subsequent policy measures failed, were the research findings or policy measures congruent with the views of all actors involved?
2. How does involving the policy's target population's influence the design of policy research? Does their participation result in research findings or subsequent policy measures that are congruent with the views of all actors involved?
3. What are the outcomes of an interactive process of problem structuring in policy research, whereby the problem definitions, underlying assumptions, and norms are analyzed and discussed? How does this impact the interventions themselves, the criteria for success, and the research questions considered relevant?

Analyses of the Case Studies

Ill Structured Problems

Analyses of the case studies indicated that the actors involved, namely policy makers, researchers, and policy's target populations, did not agree on the kind of information that was needed nor did they agree on which intervention was most appropriate. They defined the problems for which solutions were sought differently and therefore used different criteria to assess the intervention in question. These differences in problem definition were the result of differences in the theoretical assumptions (background theories) and normative values held by each actor. In accordance with Hisschemöller (1993) and Dunn (2004), we contend that this resulted from the fact that the problems were ill structured. There was no agreement with respect to what kind of information was necessary and the actors held divergent normative values.

In the case study on the use of mebeverine for the treatment of irritable bowel syndrome (Chapter 3), policy makers requested information on the effects of the drug and on the effects of a diet. Attempts to exclude the drug from the benefit package failed due to discussions on the evidence relating to the drug's effectiveness. A new trial was expected to generate the knowledge that was lacking. A normative value of the policy makers was that costs should be carried by the community only when the drugs that had proven to be effective. For physicians, however, new evidence would not change their prescribing patterns. They considered the frequent visits of unsatisfied patients to be their greatest problem. They claimed that, mebeverine is safe and effective in some patients and could therefore be prescribed, even if the effects were a placebo effect. According to the physicians, prescribing a medicine was important as it could help to maintain a good and cooperative relationship with the patient.

In the case study investigating the use of epoprostenol and bosentan for the treatment of pulmonary hypertension (Chapter 7), policy makers commissioned research to generate the necessary information on the effectiveness of these two drugs. A normative value in this case was that the use of medicines in specific patients groups is only adequate if evidence on its effectiveness in those groups is available. Assumption was that criteria related to the characteristics of an intervention, such as efficacy, effectiveness, therapeutic value, and efficiency, are useful criteria for determining whether or not an intervention should be reimbursed. However, according to the researchers (in this case, the researchers were physicians), the

problem was that the medicines were frequently prescribed to the wrong patients. They claimed that the diagnosis was inadequate due to lack of experience and expertise. Background theories that emerged were the condition pulmonary hypertension is rare and the condition is often the result of other underlying pathology. The researchers thus claimed that experience is needed to make an adequate diagnosis. The researchers considered an adequate selection of patients necessary to yield adequate prescription practices of the drugs.

In the case study on intoxication (Chapter 8), it became evident that two psychiatrists valued the observational unit differently. The psychiatrists had different perspectives with respect to what defines an optimal situation and what the aim of the psychiatric consultation should be. According to one psychiatrist, an observational unit for intoxicated patients can create an opportunity to talk to the patients shortly after the event. This psychiatrist also claimed that the psychiatric consultation should be used to, firstly, assess if, and to what extent, the patient is suicidal and, secondly, understand what happened. His colleague in another hospital considered short admission at an observational unit impractical. He claimed that the patient would not remember the agreements that had been made. Additionally, he indicated that relevant caregivers from outside the hospital can rarely be contacted when an intoxication incident takes place as this most often occurs outside of regular work hours. The psychiatrist emphasised that, although a suicide assessment is important, arranging adequate post-incident care for the patient must also be made a priority. In the same case study, internists and clinical pharmacists agreed that the uncertain anamnesis of intoxicated patients is a problem. However, their preferences with respect to solutions varied. The clinical pharmacists claimed that measuring the type and amount of ingested drugs is the only way to determine the risk of complications while the internist claimed that observing the patient's symptoms and physical state is an adequate method of predicting the actual risk of complications.

Answering the Wrong Question

The fact that the health care problems were ill-structured was insufficiently acknowledged in the case studies presented in this dissertation. As a result, the problem definitions shifted during the research project and the implementation of subsequent policy measures failed. Although research provided an answer to a research question relevant from one point of view, this question was irrelevant to either the commissioner of the study or the actors who need to cooperate on the

subsequent policy measures. In essence, the research conducted answered the wrong question. This is called a type III error (Dunn, 2004; Hisschemöller, 1993) and, evidently, due to the type III error, the subsequent policy measures endeavoured to solve the wrong problem.

Due to a shift in problem definition, the research results generated by a preliminary study in the mebeverine case (Chapter 3) were not considered useful by the commissioner. As a result, the intended clinical trial was not commissioned. In the case study on the use of epoprostenol and bosentan for the treatment of pulmonary hypertension (Chapter 7), the researchers redefined the problem because they considered the initial definition inadequate. Once again, this redefinition was not considered useful by the commissioner and the study was not commissioned. In our case study on Cox-2-selective inhibitors (Chapter 6), the problem shifted in three out of the four research proposals after a European tender for research proposals was presented. This case study showed that the researchers were strongly oriented towards standard medical rationality and practice and, at the same time, were inspired by the methodologies they were used to apply.

In some other case studies, the focus was placed on why the implementation of some policy measures based on the research results failed. It became evident that policy failed because the target population defined the problem in such a way that the applied interventions could not be considered a solution to the problem. Alternatively, the interventions conflicted with the target populations' normative values. In the study on interferon-beta (Chapter 5), the actors agreed that interventions could be useful. However, they did not agree that the implementation of policy measures was needed to optimise patient care. Neither the Ministry of Health nor the policy's target populations, namely the physicians and health care insurance companies, considered the situation to still be problematic at the time the policy was proposed. In the mebeverine case study (Chapter 3), the failure to withdraw mebeverine was the immediate impetus for the research project. However, whether the selected means, namely providing new evidence, could have prevented a second policy failure remains questionable. It is unlikely that new evidence would have provided a solution to the problems perceived by the target populations. In the case study on lamotrigine (Chapter 4), a protocol on the use of lamotrigine had not been successfully implemented. The protocol did not correspond with the normative values of the neurologists in specialised centres on searching for adequate and new therapies.

In many cases, the evaluation was limited to the level of what Fischer (1999) terms verification. These evaluations focused only on measuring whether the intervention met a set of defined criteria. In practice, however, we observed that actors disagreed not only on what criteria or outcome measures were relevant but also on which problem needed to be solved.

Diversity

Variability in problem definitions was related to: a) differences in interpretive frames; b) differences with respect to the context in which actors worked; and c) the characteristics of the topic being investigated. In all case studies, differences in interpretive frames were found between actors from the various disciplines. Relevant actors included policy makers, physicians, and researchers. Differences were not only found between actors from various disciplines, they were also found within a given discipline. In the case study on care for intoxicated patients (Chapter 8), the psychiatrists had different views with respect to the aim of the psychiatric consultation and the relationship between the physician and the patient. These differences in problem definitions were shaped by differences in theoretical assumptions and normative values.

Differences with respect to the context in which actors worked were also found. In the case study on the use of Lamotrigine for the treatment of epilepsy (Chapter 4), neurologists working in a general hospital and neurologists working in a tertiary level epileptic centre treated different spectrums of the epilepsy patient population and also held different treatment goals. The neurologists that worked in the general hospital were restrictive in their prescription practices when it came to prescribing the relatively new drug Lamotrigine. Most epilepsy patients receiving treatment in the general hospital could be effectively treated using conventional drugs. The neurologists therefore decided that the patients should not be exposed to new drugs that may have unknown effects. In contrast, the neurologists at the epileptic expertise centre treated patients with complex epilepsy and their views thus differed from those of the neurologists working in the general hospital. Additionally, in the case study on intoxication (Chapter 8), the available equipment differed. In the general hospital a medium care unit was available. This affected the proposed interventions to arrange adequate care for intoxicated patients. Furthermore, the cases demonstrated that changes within the work environment affected the actors' ideas regarding their responsibilities and their appraisals of the interventions. In both

the case study on Interferon-beta to treat multiple sclerosis (Chapter 5) and the case study on the new drugs for pulmonary hypertension (Chapter 7), the interventions that had been proposed earlier were no longer considered meaningful by the actors' after reorganisation because the proposed measures no longer belong to their task.

Lastly, many of the case studies dealt with topics that were difficult to define for one or more of the following reasons: a) the projects dealt with indistinct diseases whereby multiple theories about the underlying mechanism of the disease exist (i.e. irritable bowel syndrome and multiple sclerosis); b) the diagnostic process for the disease was complex (i.e. pulmonary hypertension and the risk assessment for intoxicated patients); c) finding the optimal treatment for an individual patient was difficult (i.e. irritable bowel syndrome, epilepsy, adequate care for patients who have attempted suicide by intoxication); d) the disorders demanded multidisciplinary care (i.e. in the case on intoxication, the patient required both somatic and psychiatric or psychosocial care), or e) patients' complaints could be attributed to multiple causes or diseases (i.e. new analgesics, pulmonary hypertension, the underlying reasons for the suicide attempt). These difficulties related to the topic in need of investigation are frequently dealt with as uncertainties. In many cases, research was commissioned in an effort to decrease the uncertainty. However, disagreement among experts will remain existent, because they are arguing from different premises (Schwartz & Thompson, 1990).

Problem Structuring

In order to avoid answering the wrong question, it is imperative that we engage in a more explicit and extensive process of problem structuring whereby the different perspectives are taken into account. As suggested by Fischer (1999), a full evaluation should include an assessment of actors' problem definitions and their underlying argumentation. In the case studies presented in Chapters 5, 6, 7, and 8, the target populations had been involved so that their perspectives would be taken into account. However, the results of these studies suggest that participation of stakeholders is insufficient for estimating the likelihood of successful policy implementation (interferon-beta, chapter 5). Compiling an inventory of the target populations' problem definitions parallel to commissioning specific research was also insufficient. In the studies presented in Chapter 6 and Chapter 7, the target populations were consulted in a part of the study. However, they were not given the opportunity to influence the problem definition nor were they able to impact the

direction in which solutions would be sought. Although the fact that the actors may define the problem differently was acknowledged, our findings indicated that the PAM staff's definition of the problem had the most influence on the kind of knowledge sought through research.

In Chapter 8, we found that an interactive methodology based on the model for fourth generation evaluation (Guba & Lincoln, 1989) was useful in that it effectively made the divergent problem definitions and underlying normative values transparent. In the study reported in Chapter 8, the interactive process of problem structuring yielded a much broader insight on the conditions under which an intervention, a six hour observation unit, could be successful. In comparison to the traditional approach, problem structuring resulted in the identification of alternative interventions and research questions. Even though efficiency remained an important issue, the interactive process demonstrated that efficiency research was premature. Questions that appeared to be more relevant were linked to the assumptions actors had about the proposed intervention with respect to the uncertainty surrounding patients' anamneses and associated risks and/or the most adequate moment for psychiatric consultation.

If an interactive approach had been applied in the other case studies, the proceedings of the studies could have been quite different. In the case study on the use of Epoprostenol and Bosentan for the treatment of pulmonary hypertension (Chapter 7), an interactive approach could have made it possible to obtain evidence that was relevant to both perspectives. For example, both the effectiveness in certain groups of patients and the efficiency of the whole process of diagnosis and treatment could have been evaluated. With respect to the health care problem concerning treatment for patients with the irritable bowel syndrome (Chapter 3), acknowledging the differences between patient subgroups and understanding the potential mechanisms underlying the complaints could have been highly advantageous. In the absence of convincing evidence on specific therapies for identifiable subgroups of patients, identifying the best treatment for an individual patient in a standardised way, for example by means of N-of-1 trials, may be necessary. These suggestions correspond with the problems as defined by the physicians and could have potentially led to prescription practices whereby patients who are likely to benefit from Mebeverine are given the prescription and patients who are unlikely to benefit are not given the prescription.

For a full evaluation, an analyst should aim to identify the various problem definitions and make the underlying assumptions and normative values transparent. If actors' interpretive frames are insufficiently congruent, the underlying assumptions should be discussed. It is even worthwhile to make the nature of this incongruence the topic of research. The methodology of reconstructing interpretive frames (Grin et al., 1997) is adequate in not only elucidating the viewpoints of individuals but also assessing the degree of congruence in perspectives (Chapter 2). Additionally, the classification of problems according to Hisschemöller (1993) can help to determine what kind of research is needed. If a problem is well structured, current HTA methods can provide evidence that the actors can agree upon. In these cases, researchers assess whether or not the defined criteria are met by the intervention (verification). When a problem is ill structured, problem definitions and underlying argumentation should be discussed (as will be further discussed below). This corresponds with the levels of validation and vindication (Fischer, 1999). In other words, assessing the degree to which a problem is structured can enable effective decision-making and helps to decide at which level of Fischer's model the evaluation should take place. In this alternative approach to evaluation, it is not the specific intervention that is the main objective. Rather, it is the problem for which a solution is sought. By applying an interactive methodology, we make it possible to pay attention to diversity in perspectives and situations. In many schemes on evaluation or policy research, assessing whether a programme meets criteria is only a part of the whole process or approach. In the context of policy analysis, Dunn (2004) has placed significant emphasis on the problem structuring phase. Conventional HTA methods are limited to one of the phases of policy research (methods used during the 'recommendation' phase). In conducting a problem oriented HTA, we can learn from the examples of argumentative (also called interpretive) methods put forth by Grin & van de Graaf (1996b), Grin et al., (1997), Yanow (2000) and Reuzel (2001). The elaboration of this problem oriented approach to HTA is provided in paragraph 'recommendation' following a short reflection on the case studies and a brief discussion of some relevant aspects characteristic to this argumentative approach.

Reflections

Limitations of this Study

The aim of this study was to explore whether the theory of argumentative policy analysis could provide an explanation for the limited applicability of HTA results. In order to explore the hypotheses in different contexts, the case studies were chosen so that they would differ on several variables. The case studies dealt with either failed policy research, failed implementation of policy measures, or cases in which lessons learned were incorporated to prevent failure. The cases also included different forms of medical expertise such as primary care, hospital care or both. Additionally, the cases involved problems arranged within a regional context (i.e. one hospital) or at a national level.

Unfortunately, no case studies in which policy research resulted in the successful implementation of policy measures were included. In two prospective case studies (new analgesics in Chapter 6 and the drugs Epoprostenol and Bosentan in Chapter 7), some of the lessons learned from previous studies were incorporated. However, the time frame of these studies was too short to thoroughly assess the consequences of the implemented changes. Furthermore, some comments can be made on the approach taken in these cases (see above, paragraph 'problem structuring'). Additionally, in the case on care for intoxicated patients, the proposed observational unit has, to date, not yet been implemented. At the time of the evaluation, it was already clear that the implementation of an observational ward would not be implemented prior to the realisation of a new hospital building. As a result, clear evidence on the surplus value of the argumentative approach in terms of the successful implementation of policy measures cannot yet be determined. Nevertheless, this study did provide us with several insights that suggest that the argumentative approach is a promising means of improving the link between an HTA and its users. As discussed above, an interactive process of problem structuring was successful in assessing the degree of congruency. It also generated outcomes other than those yielded by a more conventional approach (Chapter 8). Some lessons can be learnt but further research is needed so that the limitations and difficulties of this approach can be dealt with effectively.

Reflections on the Theory of Argumentative Policy Analysis

Although the theory provided an adequate framework for explaining both the proceedings of research projects and the subsequent policy measures, we identified

some difficulties that manifest when an argumentative approach to policy development is applied.

Firstly, we observed that changes in the policy context negatively affected the project on Interferon Beta (Chapter 5) and the project on drugs for pulmonary hypertension (Chapter 7). Shifts in the Health Care Insurance Boards responsibilities created a situation in which the proposed policy measures were no longer considered relevant. As a result, the projects ended without solving the problem. In these research projects, it appeared difficult to anticipate these political changes. Possibly, if the approach to the problem had been broader, it may have been possible to deal with the changes in the political context. However, time and resources are all too often limited thereby making it difficult to redefine the problem when some steps have already been taken, as was the case in the study on pulmonary hypertension (Chapter 7). In both case studies, the actors' reactions to the changes in context were dependent on the actors' interpretive frames. Each actor assessed the consequences of the actual changes and took action according to this assessment. Obviously, predicting the way in which changes within the context will affect the proceedings of policy development is difficult. It is thus imperative that the analyst learn how to handle these changes effectively.

Secondly, we noticed that some difficulties presented when changes in the process of commissioning research were implemented (Chapter 6 & 7). The main assumption of the argumentative approach is that policy making is a process of co-production. The adoption of an argumentative approach strongly depends on the willingness of the actors to learn from one another and the various perspectives put forth by other actors. The results indicated that policy makers were not always open to the suggestions and perspectives of the target populations. When policy makers noticed that target populations or the researcher had not "understood" what they wanted, they often decided not to commission the project or they discontinued the research project (Chapter 3 on Mebeverine, Chapter 6 on new analgesics, and Chapter 7 on pulmonary hypertension). In the case study on the use of new analgesics (Chapter 6), researchers questioned the acceptability of the research project's outcome. While the staff members of the Health Care Insurance Board requested that the researchers analyse target groups problem definitions, they also requested that the researchers generate the knowledge they considered most relevant. Evidently, these two requests conflicted. Apparently, it is not obvious to apply a problem oriented approach. This is partly due to the fact that actors are often

suspicious of the perspectives of others. Schön and Rein (1994) referred to Habermas who stressed that discussion can only take place when actors have freedom, openness and equality in dialogue. Clearly, domination of any one perspective inhibits dialogue. Unfortunately, having freedom, openness and equality in dialogue is an ideal situation that is very difficult to create in practice. For example, we observed strategic behaviour among stakeholders in the project on policy measures for Interferon Beta use (Chapter 5). Eliminating strategic behaviour is practically impossible. A procedure is needed that makes it possible to deal with this strategic behaviour in a sound way.

Evidently, the difficulties relating to the application of an argumentative approach, namely changes in the context, limited willingness to learn from one another and strategic behaviour, demand that conditions be placed upon the procedure and the role of the analyst. This is discussed further below (paragraph 'The role of the researcher')

The Need for Interaction and Learning

Although the reconstruction of interpretive frames through one interview can assess the degree of congruence in perspectives (Chapter 2), it cannot predict the actions of the actors. When congruence in the perspectives (problem definitions, underlying assumptions and norms) is insufficient, a process of learning needs to take place. This learning may include considering another solution, redefining an actor's problem definition, or adapting an actor's background theories. Learning between actors that hold divergent interpretive frames aims to construct congruent meanings. Congruence is necessary for the identification of interventions that are broadly supported. It is also necessary for the elicitation of cooperation from all actors involved. The results of the case studies have demonstrated that involving actors and target populations without discussing their problem definitions and underlying assumptions prior to the selection of interventions is insufficient (Chapter 5, 6, and 7). Reaching for consensus by creating shared problem definitions or shared values is not necessary (Grin & van de Graaf, 1996a). Grin and van de Graaf (1996a) have provided some conditions for learning between actors. They distinguished between first-order learning (learning related to assessments of solutions and problem definitions) and second-order learning (learning related to assumptions and background theories). Discussing problem definitions and underlying assumptions during interaction (frame reflection) can result in a redefinition of problem

definitions and their underlying assumptions (reframing) (Schön & Rein, 1994). According to Guba and Lincoln (1989), the occurrence of learning within or between participants during an interactive process is one of the main criteria for its success. Learning does not only imply a change in perspectives, but also the acknowledgement of each others point of view (Reuzel 2001). Learning can take place between respondents in an interactive process. Alternatively, measures can be implemented to in an effort to attain learning in certain (target) populations (Grin & van de Graaf, 1998).

Interaction between stakeholders enables us to judge the validity of the claims that play an important role in the interactive process. When actors disagree, deciding which perspective should be changed is difficult. Many questions arise: Should policy makers change their perspective when interviews reveal that the target population holds different views towards the situation? What must be done when the target population does not perceive problems for which changes are needed? Should target populations then change their perspective? In the case study on the use of Epoprostenol and Bosentan for the treatment of pulmonary hypertension (Chapter 7), policy makers asked if there was enough evidence on the effectiveness of the medicines while physicians believed that an adequate diagnosis was also needed. In the case study on the use of new analgesics (Chapter 6), policy makers defined the problem in a one sided fashion and this resulted in dismissive reactions from the field. Physicians did not consider their prescription practices of the new analgesics to be problematic. Nevertheless, at a national level, many off label prescriptions were observed and the drugs appeared to be less innocent than previously thought thereby leading to the withdrawal of the drugs from the market. Schön and Rein (1994) remarked that there are no independent criteria that can be used to judge whether frames are right or wrong. Judging an actor's assumptions is limited to assessing its coherence and verifying the theoretical assumptions. Alternatively, we can endeavour to reach a shared perception and agreement on a specific situation. Attempting to understand the perspectives of others by translating one frame to another is one potential method by which this can be done. Actors themselves can seek mutual understanding and thereby create shared frames (Schön & Rein, 1994). One should be able to place the perspective of another on a given situation in one's own frames. This idea corresponds with Habermas' concept of 'communicative action', which is an act of speech that is oriented towards mutual understanding (Outhwaite, 1996, p. 160) The goal of communicative action is to enable one or more persons to reach an

understanding. However, frame reflection does not always lead to a reframing of the problem and reframing does not always lead to a resolution of the problem (Schön & Rein, 1994). Sometimes controversies can be resolved at a practical level. Interventions that provide a solution for quite different problems and different norms can be identified. In the case study on intoxication (Chapter 8), the six hour observational unit was considered useful for numerous problems relating to various normative values. For internists, the unit was seen as an opportunity to monitor patients and therefore a means by which the risk of complications could be decreased. For psychiatrists, the unit provided an opportunity to talk to the patient shortly after the event.

Lastly, it is important to note that sometimes external factors can help to define the scope of the analysis. In general, the level of freedom with which the commissioner can act strongly affects how the analysis is defined. For example, in the projects initiated by the Health Care Insurance Board, its mission determined the range in which the definition of problems and the identification of solutions could be discussed

The Role of the Researcher

A problem oriented approach impacts the role of HTA researchers. This approach requires the analyst to take an active part in defining the nature of the problem (Dunn, 2004). HTA researchers should enlighten and inform policy by helping surface the diversity of socioethical issues that may affect individuals and society (Lehoux et al., 2007). It is the researchers' task to analyse the problem and the degree to which actors agree. If actors disagree, the analysts must structure the problem in such way that it is manageable. To do this, researchers should reconstruct arguments and moderate the discussion. As a mediator, the analyst can create a climate in which actors with different perspectives are stimulated to rethink their interests (Schön & Rein, 1994). The researchers can then verify the arguments put forth by either consulting other sources, such as the literature, or confronting the actors with each other's arguments. As demonstrated in Chapter 6, this role is often difficult. Additionally, many researchers are incline to resist this role as they often hold strong views on how they should relate to other actors. Consequently, applying an open, problem oriented approach can be difficult. In the project on intoxication, one specific intervention, namely the six hour observational unit, received significant

attention right from the start. The underlying assumptions were made transparent but these assumptions could have been discussed more explicitly.

Alternatively, policy makers can play a role in problem structuring. This, however, can be problematic as policy makers are often stakeholders with vested interests. In the case studies on the use of COXIBs and Epoprostenol, the commissioned research was strongly affected by the HCIB's preferences with respect to the intervention. Unfortunately, HCIB's preferences led to dismissive reactions from both researchers and health professionals.

Recently, many authors emphasise a dialogue between researchers and users to improve research utilisation (Millbank, 2000; Moynihan, 2004; Hivon et al., 2005; Elliot & Popay, 2000; McGregor & Brophy, 2005; Levin et al., 2007). Schön and Rein (1994) have argued that collaboration between academics, (policy) designers, and practitioners is imperative. They proposed consultative policy research in which the researcher helps practitioners to generate usable knowledge. The researcher can help practitioners in two ways. Firstly, the researcher can aid the process of frame reflection by reconstructing the interpretive frames. Secondly, he or she can help practitioners to create conditions for mutual trust, which is essential for processes of learning and reaching agreement.

Several difficulties that the analyst must deal with during the interactive process have been mentioned (paragraph 'reflections on the theory of argumentative policy analysis'). Nonetheless, the interactive research methodology does provide opportunities to deal with power differences and strong identities. Criteria for a sound interactive process have been presented by Guba and Lincoln (1989) and elaborated upon for interactive Technology Assessment by Reuzel (2001). Important elements of an interactive analysis include: individual interviewing, discussing respondents' utterances anonymously, and engaging the analyst as a moderator between all participants. Clearly, the analyst can play an active role in redressing power differences (Reuzel et al., 2007). In order to increase the rigor of the analysis, the analyst's role should be made transparent and ownership should be left to the participants of the interactive process.

The Role of Scientific Evidence

Evaluation should not be limited to the assessment of the intervention's outcomes but should also discuss the differences in problem definitions and underlying assumptions. The kind of information needed and how that information will be used

depends largely on the degree of problem structuring and the kind of changes that are needed. Sometimes it is more appropriate to assess the assumptions behind certain interventions or problem definitions. Although, in the context of the project on intoxication (Chapter 8), prior research was concerned with the evidence relating to the cost effectiveness of a six hour observation, the interactive analysis indicated that evidence on the effects of drugs for amnesia was equally important for implementing the unit. Empirical research is thus not restricted to measuring the outcome of interventions in terms of defined criteria. The theoretical assumptions on how an intervention can provide a solution to a problem and how the intervention can contribute to the final goal must also be considered. The outcome of evaluations should be used more frequently for interpreting the relationship between the intervention, the mechanisms that make an intervention work, and its context (realistic evaluation; Pawson & Tilly, 1997).

Another consequence of applying an problem oriented approach is that scientific evidence is no longer the only source of knowledge that governs the direction of evaluation and subsequent policy measures. Current ideas on the use of research results promote the direct application of specific outcomes to a specific intervention, change or implementation. However, Weiss demonstrated that the direct use of evaluation (instrumental use) in decision rarely takes place (Weiss, 1980; Weiss & Bucuvalas, 1980). More often 'use' includes processes such as the understanding of programs and gaining new ideas and insights (conceptual use). Moreover, many policy actions do not imply concrete decisions, but involve a set of small uncoordinated steps of decisions. It is more likely that new evidence from research is used conceptually by redefining problem definitions or assumptions that are part of an actor's background theories. Because information competes with experiential and epidemiological knowledge as well as pathophysiological insights, the direct application of new evidence is unlikely to occur.

Recommendation: More Emphasis on Problem Analysis in HTAs

Conventional HTA methods need to be embedded in a more problem based approach, whereby an extensive process of problem analysis must take place prior to the design of research projects. The aim of an HTA would then be to assess the level of problem structuring and the type of research needed to deal with this problem. When problems are well structured, conventional HTA methods are adequate for

providing the necessary knowledge. When problems are ill structured, interactive research methods are needed to structure the problem and identify broadly supported interventions and changes. Thus, by taking into account the problem definitions of actors involved and their underlying arguments, assumptions and values, the approach becomes problem based. This kind of approach results in an evaluation that takes contextual considerations and the heterogeneity of the situation and perspectives into account.

However, there may be a gap between the theoretical ideal of the argumentative approach and the practical possibilities in HTA research. A number of factors do affect the actual potential of using HTA in a problem based manner. These factors include the amount of time available to conduct the study, the expertise needed, and the willingness of actors involved to have an open attitude to alternative problem definitions and solutions. Furthermore, the current structure for funding HTA studies may also serve as a limitation. Attaining funding for a problem analysis prior to conventional research is difficult. Furthermore, a problem based approach may result in uncertainty regarding the results of the study. Therefore, the following practical method for preventing type III errors (solving the wrong problem) is recommended: The process of HTA research should be divided into three phases, namely 'problem analysis', 'problem structuring' and a 'knowledge provision'. Following this, all researchers should assess the level of structuring needed to improve the usability of their study results. This approach is comparable to the 'usable knowledge' approach posited by Hoppe & Grin (2000).

The question remains whether a process of problem analysis is sufficient to reconstruct the perspectives (problem definitions, underlying assumptions, and values) of the actors involved. In order to reconstruct a problem definition and the assumptions that underlie that definition, interaction between actors may be necessary. The concept of "reconstructing" suggests that an actor's problem definition is not a given fact that can easily be measured. Sometimes people will give the issue additional consideration when confronted with another perspective. Alternatively, an actor may redefine his or her perspective when confronted with alternative views. As discussed above, an important goal of interactive research is changing perspectives. This too needs to be explored further. In conclusion, despite the possible limitations of the approach, I contend that an explicit process of problem analysis that takes the perspectives of relevant actors into account prior to actual research is imperative.

Argumentative Problem Analysis

Prior to conventional HTA research, one should identify which actors are involved, who needs to cooperate on possible interventions and who is likely to experience the consequences of the intervention. The following step is the assessment of how these actors define the problem and judge the interventions. Based on this information, an estimate of the level of problem structuring (ill, moderately, or well structured problems) should be made. The level of problem structuring will then determine the kind of research or intervention needed to solve the problem(s). When the problem is well structured, conventional HTA methods can be used to generate the requested information. When problems are ill structured, the problem should be structured actively so that agreement can be reached on the criteria for success and the kind of interventions that should be implemented. For this purpose, the interactive HTA method is appropriate. This method should be performed by analysts with sufficient experience in using these methods so that the results generated are valid. The question is, how can we, with minimal effort, get a good impression of the level of structuring? The case studies in this thesis revealed a number of variables that can effectively serve as an indicator for ill structured problems. Factors that were associated with diversity included (mentioned above) the involvement of actors from various disciplines, the existence of different perspectives and theories on the underlying causes of the disease being subject to evaluation, and differences in the contexts in which actors work.

Yanow (2000, p.22) has provided the following steps for interpretive analysis: a) identify artefacts that are carriers of meaning for a given policy issue; b) identify communities of meaning/practices that are relevant; c) identify the (discourses(; d) identify the points of conflict and their conceptual sources; e) develop interventions, such as showing the implications of different meanings or mediating and intervening to bridge the differences. Grin et al. (1997) have developed a handbook in which steps for an interactive Technology Assessment are presented. Using these models as a basis, the steps to be followed in a problem based, argumentative approach in HTA are identified and presented in Table 9.1.

Recommendations for Practice and Future Research

Although a number of aspects require further exploration, some recommendations can be made. All HTA researchers should be able to perform an extensive process of problem analysis. Problem analysis should involve the identification of potentially divergent definitions of the problem for which a solution is sought. In HTA

education, courses on how people frame problems and on how this can lead to different perspectives should be provided. Furthermore, the methodology of reconstructing interpretive frames should be taught as a tool for problem analysis and assessing the degree of congruence in perspectives.

Expertise on interactive methodology is needed for adequate problem structuring when problems are ill structured. So, in addition to including experts on economics or quality of life, multidisciplinary groups in HTA departments should also include experts on interactive methodology. Alternatively, multidisciplinary working groups can be formed around research projects in which experts in interactive methodology are also included.

Commissioners, policy makers, and funding institutes should demand a solid problem analysis in research which should then be extensively described in research reports. In order to generate adequate reports, the criteria for policy issue papers developed by Dunn (2004) could be applied. Possibly, one could organise a limited number of pilot projects.

Based on the discussion above concerning the distinction between policy analysis and problem structuring, some questions remain unanswered. A comparison between the outcome of an 'argumentative problem analysis' and 'interactive problem structuring' using the methodology described by Guba & Lincoln (1989) should be made. Relevant questions could include: a) is it possible to reconstruct actors underlying background theories and normative values to an acceptable level during problem analysis?; b) when a problem analysis is conducted with a limited number of respondents, what are the chances that some relevant perspectives will be missed?; c) will this complicate implementation?; d) is it possible to assess the level of structuring during a problem analysis in which only a limited number of stakeholders participate without interaction and discussions regarding the underlying arguments? Another important research question is whether HTA projects in which argumentative problem analysis have been performed actually result in the successful implementation of policy measures that indeed solve the perceived problem(s).

Table 9.1. Steps in problem oriented HTA

1.	**Identify stakeholders:**
	a. actors who's co-operation is needed
	b. actors who might experience consequences
2.	**Identify their definitions of the problem**
	a. Include a limited number of stakeholders that might have different perspective (based on expertise or context in which they work)
	b. Identify which criteria are considered as relevant
	c. Identify what should be changed in current practice
3.	**Identify underlying assumptions, arguments and values**
	a. Assumptions on mechanism behind disease and intervention
	b. Arguments concerning own task and aim of treatment
4.	**Assess the degree of structuring**
	a. is the problem well-structured, moderately structured or ill-structured
5.	**Action:**
	a. Show implications of different interpretations
	b. Identify the type of research that is needed
	i. well-structured problem: identify solution and possible research questions (conventional methods, such as cost-effectiveness analysis; RCT)
	ii. ill-structured problems: reach for agreement on problem definition and interventions (interactive methods, interactive Technology Assessment)
	c. Apply or commission requested empirical or interactive research

10

Appendices

References

Arranto CA, Mueller C, Hunziker PR, Marsch SC, Erikson U. (2003) Adverse cardiac events in ICU patients with presumptive antidepressant overdose. Swiss medical weekly 133: 479-483.

Avorn J. (2006) Evaluating drug effects in the post-vioxx world. There must be a better way. Circulation 113:2173-2176.

B&A. (2000) Gebruik van interferon-beta bij multipele sclerose. Den Haag. B&A groep.

Banta D.(2003) The development of health technology assessment. Health Policy 63(2):121-132.

Barker M. (2003) Assessing the 'quality' in qualitative research. The case of text-audience relations. European Journal Communication 18:315-335

Behan PO, Chaudhuri A, Roep BO. (2002) Pathogenesis of multiple sclerosis revisited. J R Coll Phys Edinb 32: 244-265.

Berg M, van der Grinten T, Klazinga N. (2004) Technology assessment, priority setting, and appropriate care in Dutch health care. International Journal of Technology Assessment in Health Care 20(1): 35-43.

Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. (1998) Getting research findings into practice. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. BMJ 317:465-468

Black N. (1996) Why we need observational studies to evaluate the effectiveness of health care. BMJ 312(7040):1215-1218.

Brett AS, Rothschild N, Gray R, Perry M. (1987) Predicting the clinical course in intentional drug overdose. Implications for use of the intensive care unit. Archives of Internal Medicine 147(1):133-137.

Camelleri M. (1999) Therapeutic approach to the patient with irritable bowel syndrome. American Journal of Medicine 107:27S-32S

Charatan F. (2000) Drug for irritable bowel syndrome taken off the market. BMJ 321:1429.

Chernichovsky D. (1995) What can developing economies learn from health system reforms of developed economies? Health Policy 32:79-91

Clavarino AM, Najman JM & Silverman D. (1995)The quality of qualitative data: two strategies for analyzing medical interviews on analyzing interviews. Qualitative Inquiry 1: 223-242

College voor Zorgverzekeringen (2000A) Onderzoek handelswijze ziekenfondsen bij aanvragen interferon beta bij multipele sclerose. Amstelveen: College voor Zorgverzekeringen.

College voor Zorgverzekeringen. (2000B) Evaluatie van het gebruik van interferon-beta bij multipele sclerose. Amstelveen: College voor Zorgverzekeringen

College voor Zorgverzekeringen. (2002) Mebeverine. Samenvatting onderzoeksrapport "irritable bowel syndrome in primary care". Amstelveen: College voor Zorgverzekeringen.

Connell AM. (1965) Physiological and clinical assessment of the effect of the musculotropic agent mebeverine on the human colon. BMJ 2: 848-851.

Coolen R. (1997) Patiëntenvoorlichting interferon beta 1b beïnvloedt keuzeproces. Pharmaceutisch Weekblad 132: 1394-1396

de Haan RJ, Vermeulen M. (1999) Interferon bèta-1b nu ook voor de secundair progressieve vorm van multiple sclerose? Ned Tijdschr Geneeskd. 143(14): 709-711.

Denzin N. (1989) Interpretive interactionalism. Newbury Park: Sage

Derry S, Loke YK., Aronson JK. (2001) Incomplete evidence: the inadequacy of databases in tracing published adverse drug reactions in clinical trials. BMC 1(7).

Derthick M. (1972) New Towns In-Town. Washington D.C.: The Urban Institute

Draborg E, Gyrd-Hansen D, Poulsen PB, Horder M (2005) International comparison of the definition and the practical application of health technology assessment. International Journal of Technology Assessment in Health Care 21(1):89-95.

Drummond M, Cooke J, Walley T. (1997) Economic evaluation under managed competition: evidence from the U.K. Soc Sci Med 45(4): 583-595.

Drummond M, Weatherly H. (2000) Implementing the findings of health technology assessments. If the CAT got out of the bag, can the TAIL wag the dog? Int J of Technology Assessment in Health Care 16:1-12.

Dunn W. (2004) Public policy analysis. An introduction. 3rd edition. Upper Saddle River, NJ: Pearson/Prentice Hall.

Duthie T, Trueman P, Chancellor J, Diez L. (1999) Research into the use of health economics in decision making in the United Kingdom. Phase II. Is health economics 'for good or evil'? Health Policy 46: 143-157.

Elliot H & Popay J. (2000) How are policy makers using evidence? Models of research utilisation and local NHS policy making. J Epidemiol Community Health 54:461-468.

Elmore RF. (1985) Forward and backward mapping. In: Hanf K and Toonen D, eds., Policy Implementation in Federal and Unitary Systems. Dordrecht: Martinus Nijhoff, 33-70.

Evans TW, Gatzoulis MA, Gibbs JSR.(2002) A national pulmonary hypertension service for England and wales: an orphan disease is adopted? Thorax 57:471-472.

Fabbri A, Marchesini G, Morselli-Labate AM, Ruggeri S, Fallani M, et al. (2003) Comprehensive drug screening in decision making of patients attending the emergency department for suspected drug overdose. Emergency Medicine Journal 20: 25-28

Farthing MJ. (1998) New drugs in the management of the irritable bowel syndrome. Drugs 56: 11-21

Fischer F.(1980) Politics, values, and public policy : the problem of methodology. Boulder, Colo. : Westview Press

Fischer F & Forester J. (1993) The argumentative turn in policy analysis and planning. Durham, N.C.,: Duke University Press

Fischer F. (1999) Evaluating Public Policy. Nelson-Hall Publishers, Chicago, USA.

Fischer F. (2003) Reframing Public Policy. Discursive Politics and Deliberative Practices. New York: Oxford University Press Inc.

Frankish CJ, Kwan B, Ratner PA, Higgins JW & Larsen C. (2002) Social and political factors influencing the functioning of regional health boards in British Colombia (Canada). Health Policy 61: 125-151.

Gagnon MP, Sánches E, Pons JMV. (2006) From recommendation to action: psychosocial factors influencing physician intention to use Health Technology Assessment (HTA) recommendations. Implementation Science 1:8

Gaine SP and Rubin LJ. (1998) Primary pulmonary hypertension. Lancet 352:719-725.

Gezondheidsraad (1996) Toepassing van interferon-beta-1b bij patiënten met multiple sclerose. /05

Gezondheidsraad (1997) Preoperatief onderzoek. Rijswijk: Gezondheidsraad.

Gezondheidsraad (2000) Van implementeren naar leren; het belang van tweerichtingsverkeer tussen praktijk en wetenschap in de gezondheidsraad. Gezondheidsraad Den Haag; publicatie no. 2000/18.

Graf von der Schulenburg JM. (2000) The influence of economic evaluation studies on health care decision making. A European survey. IOS Press, Amsterdam.

Greenwood B & Mandel KG. (1992) Effect of mebeverine hydrochloride on jejunal motility and epithelial transport in the anaesthetized ferret. European Journal of Pharmacology 211: 143-147.

Grin J & van de Graaf H. (1996a) Implementation as communicative action. An interpretive understanding of interactions between policy actors and target groups. Policy Sciences 29 (4):291-319.

Grin J & van de Graaf H. (1996b) Technology assessment as learning. Science, Technology and Human Values 21(1): 72-99.

Grin J, van de Graaf H, Hoppe R. (1997) Interactieve Technology Assessment. Een eerste gids voor wie het wagen wil. Den Haag: Rathenau Instituut.

Grin J, van de Graaf H. (1998) Van bovenaf, het beleid van onderaf bezien: de argumentatieve wending en de werking van beleidsinstrumenten. In: R. Hoppe & A. Peterse. Bouwstenen voor argumentatieve beleidsanalyse.

Grin J & Hoppe R. (2000) Cultural bias and framing wicked problems. A comparative analysis of structuring the car mobility problem by three institutes for parliamentarian technology assessment. In: Wagenaar H. Government institutions: Effects, changes and normative foundations. Kluwer academic publishers.

Grin J. (2004) Health technology assessment between our health care system and our health. Exploring the potential of reflexive TA. Poiesis & Praxis 2: 157-174.

Grol R, Dalhuijsen J, Thomas S, in 't Veld C, Rutten G, Mokkink H. (1998) Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ 317: 858-861.

Grol R & Grimshaw J. (2003) From best evidence to best pracice: effective implementation of change in patients' care. Lancet 362:1225-1230.

Grol R & Wensing M. (2004) What drives change? Barriers to and incentives for achieving evidence based practice. MJA 180:S57-S60.

Guba EG & Lincoln YS. (1989) Fourth Generation Evaluation. Newbury Park, Calif., [etc.]: Sage Publications.

Haines A & Donald A. (1998) Getting research findings into practice. Making better use of research findings. BMJ 317: 72-75.

Hartman M & van der Wilt GJ. (1999) Kwaliteit en doelmatigheid van preoperatief anesthesiologisch onderzoek in het Canisius Wilhelmina Ziekenhuis.

Hartman M, Rovers MM, Ingels K, Zielhuis GA, Severens JL, van der Wilt GJ. (2001) Economic evaluation of ventilation tubes in otitis media with effusion. Arch Otolaryngol Head Neck Surg 127(12):1471-1476.

Hawkey CJ. (2005) Cox-2 chronology. Gut 54:1509-1514.

Haynes B & Haines A. (1998) Getting research findings into practice. Barriers and bridges to evidence based clinical practice. BMJ 317: 273-276.

Hippisley-Cox J & Coupland C. (2005) Risk of myocardial infarction in patients taking cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs: population based nested case-controle analysis. BMJ 330: 1366.

Hisschemöller M. (1993) De democratie van problemen. De relatie tussen de inhoud van beleidsproblemen en methoden van politieke besluitvorming. Amsterdam: VU Uitgeverij.

Hivon M, Lehoux P, Denis JL, Tailliez S. (2005) Use of health technology assessment in decision-making: coresponsibility of users and producers? International Journal of technology assessment in health care 21(2): 268-275.

Hollander JE, McCracken G, Johnson S, Valentine SM, Shih RD. (1999) Emergency department observation of poisoned patients: how long is necessary?Academic emergency medicine 6(9): 887-894.

Hoppe R & Peterse A. (1998) Bouwstenen voor een argumentatieve beleidsanalyse VUGA, Den Haag.

Hoppe R & Grin J. (2000) Traffic goes through the TA machine: A culturalist comparison, chapter 10 (p. 273-324) in Norman J. Vig & Herbert Passchen (eds.), Parliaments and Technology: the Development of Technology Assessment in Europe. SUNY Press, New York.

Horwitz BJ & Fischer RS. (2001) Current concepts. The irritable bowel syndrome. NEJM 344: 1846-1850.

Hurst JW. (1991) Reforming health care in seven European nations. Health Affairs 10:7-21.

IFN MS Group. (1993) Interferon beta-1b is effective in relapsing remitting multiple sclerosis. Clinical results of a multicenter, randomized double-blind, placebo-controlled trial. Neurology 43: 655-661.

Jabaaij L , Stokx LJ, de Bakker DH. (2001) Indicaties voor het gebruik van rofecoxib. Meer 'off-label' dan 'on-label' gebruikt. Pharmaceutisch Weekblad 136:1538-1541.

Jacobson N. (2007) Social epistemology. Theory for the 'fourth wave' of knowledge transfer and exchange research. Science communication 29(1):116-127.

Janssen HAM, Muris JWM, Knottnerus JA. (1998) The Clinical Course and Prognostic Determinants of the Irritable Bowel Syndrome: A Literature Review. Scandinavian Journal of Gastroenterology 33: 561-567.

Jennings B. (1987) Interpretation and the practice of policy analysis, p. 128-152 in: Frank Fischer & John D. Forester (eds.). Confronting Values in Policy Analysis. The politics of criteria. Newbury Park etc.: SAGE.

Jones AL & Volans G. (1999) Management of self poisoning. British Medical Journal 319: 1414-1417

Kamke K. (1998)The German health care system and health care reform. Health Policy 43(2):171-194.

Kaplan PW. (2004) Reproductive health effects and teratogenicity of antiepileptic drugs. Neurology 63:S13-S23

Kapur N, Cooper J, Hiroeh U, May C, Appleby L, House A. (2004) Emergency department management and outcome for self-poisoning: a cohort-study. General Hospital Psychiatry 26: 36-41

Kellermann AL, Fihn SD, LoGerfo JP, Copass MK. (1987) Impact of drug screening in suspected overdose. Annals of emergency medicine 16(11): 1206-1216.

Kellermann AL, Fihn SD, LoGerfo JP, Copass MK. (1988) Utilization and yield of drug screening in the emergency department. The American journal of emergency medicine 6(1): 14-20.

Kemp R & Weehuizen R. (2005) Policy Learning: What does it mean and how can we study it? Publin Report No. D15. NIFU Step, Oslo.

Kerkhof A. (1985) Suïcide en de geestelijke gezondheidszorg, Lisse: Swets en Zeitlingen,

Klein KB. (1988). Controlled treatment trials in the irritable bowel syndrome: a critique. Gastroenterology 95: 232-241.

Koch P, Hauknes J and Roste R. (2003) Rationalities and innovation policy learning. In: Per Koch and Juha Oksanen (eds.). Goodnip - Good Practices in Nordic Innovation Policies. Part 2: Innovation Policy Trends and Rationalities. STEP, Centre for Innovation Research Oslo.

Langemeijer JJM, Hemrika MH, de Lange JJ. (1996) Doelmatig preoperatief onderzoek in de polikliniek Anesthesiologie. Ned Tijdschr Geneeskd. 140; 1723-1726.

Lehoux P & Blume S. (2000) Technology assessment and the sociopolitics of health technologies. J Health Polit Policy Law 25(6):1083-1120.

Lehoux P, Williams-Jones B. (2007) Mapping the intergration of social and ethical issues in health technology assessment. International Journal of Technology Assessment in Health Care 23(1):9-16.

Lévesque LE, Brophy JM, Zhang B. (2005) The risk for myocardial infarction with cyclooxygenase-2 inhibitors: A population study of elderly adults. Ann Intern Med 142: 481-489.

Levin L, Goeree R, Sikich N, Jorgerson B, Brouwers MC, Easty T, Zahn C. (2007) Establishing a comprehensive continuum from an evidentiary base to policy development for health technologies: The Ontario experience. International Journal of Technology Assessment in Health Care 23(3): 299-309.

Lie RK. (2004) Research ethics and evidence based medicine. J Med Ethics 30: 122-125

Liebelt EL, Ulrich A, Francis PD, Woolf A (1997) Serial electrocardiogram changes in acute tricyclic antidepressant overdoses, Critical care medicine 25(10): 1721-1726

Lindblom C. (1992) Inquiry and change: The troubled attempt to understand and change society. Yale University Press, New Haven, CT, USA.

Longino HE. (1990) Science as social knowledge: values and objectivity in scientific inquiry. Princeton, N.J. : Princeton University Press.

Majone G. (1989) Evidence, argument and persuasion in the policy process. Yale University Press, New Haven, CT. USA.

Malcolm LA. (1989) Decentralisation trends in the management of New Zealand's health services. Health Policy 12, 285-299.

Martinez E, Reilly BM, Evans AT, Roberts RR. (2001) The observation unit: a new interface between inpatient and outpatient care, The American journal of medicine 110: 271-277

Mazmanian DA & Sabatier PA. (1983) Implementation and Public Policy. Boston & London: University Press of America.

McGregor M & Brophy JM. (2005) End-user involvement in health technology assessment (HTA) development: A way to increase impact. International Journal of Technology Assessment in Health Care 21(2): 263-267.

Merck (2004) Merck announces voluntary worldwide withdrawal of VIOXX. News release. Whitehouse Station, NJ: Merck. www.vioxx.com/vioxx/documents/english/_vioxx_press_release.pdf

Meulendijks CFM, Berg van der EJ, Droogleever Fortuyn H, Verkes RJ, Wilt van der GJ, Kramers C (2003) Predicting the need for hospital admission in patients with intentional drug overdose. The Netherlands journal of medicine 61(5): 164-167.

Miles MB & Huberman M. (1994) Early Steps in Analysis. In: Qualitative data analysis: an expanded sourcebook. 2nd ed. Thousand Oaks, CA [etc.] : Sage

Milbank. (2001) Informing Jugdment: case studies of health policy and research in six countries. Milbank Memorial Fund.

Moret M, Reuzel RBP, van der Wilt GJ, Grin J. (2007) Validity and Reliability of Qualitative Data-Analysis: Interobserver Agreement in Reconstructing Interpretative Frames. Field Methods 19(1):24-39

Moret-Hartman M, van der Wilt GJ, Grin J. (2007) Health Technology Assessment and ill structured problems: a case study concerning the drug mebeverine. International Journal of Technology in Health Care 23(3):316-323.

Moynihan R. (2004) Using health research in policy and practice. Case studies from nine countries. Millbank Memorial Fund.

Mkoka S, Vaughan J, Wylie T, Yelland H, Jelsma J. (2003) The pitfalls of translation--a case study based on the translation of the EQ-5D into Xhosa. South African Medical Journal 93, 265-266.

Neeleman J & Wessely S. (1997) Drugs taken in fatal and non-fatal self-poisoning: a study in South London. Acta psychiatrica Scandinavica 95: 283-287.

Netherlands Court of Audit. (1992) Cost containing medicines. 22920, no 2 (in Dutch)

Netherlands Court of Audit. (2003) Between policy and implementation: lessons from a recent inquiry of the Court of Audit. The Hague, SDU. (In Dutch: Dutch Parliament, Dossier 28 831, nrs. 1-2)

Niessen LW, Grijseels E, Koopmanschap M, Rutten F. (2007) Economic analysis for clinical practice-the case of 31 national consensus guidelines in the Netherlands. J Eval Clin Pract.13(1):68-78.

Okma KGH. (2001) Health care, health policies and health care reforms in the Netherlands. Den Haag, the Netherlands: Ministerie van Volksgezondheid, Welzijn en Sport.

Oliver A, Mossialos E, Robinson R. (2004) Health technology assessment and its inluence on health-care priority setting. International Journal of technology assessment in health care 20(1): 1-10.

Outhwaite W. (1996) The Habermas Reader. Cambridge [etc.]: Polity Press

Pawson R & Tilly N. (1997) Realistic Evaluation. London [etc.]: Sage Publications.

Peacock AJ. (2003) Treatment of pulmonary hypertension. BMJ 326(7394):835-836.

Perrone J, De Roos F, Jayaraman S, Hollander JE. (2001) Drug screening versus history in detection of substance use in ED psychiatric patients, The American journal of emergency medicine 19: 49-51

Pieters T. (1998) Marketing medicines through randomised controlled trials: the case of interferon. BMJ 317(7167), 1231-1233.

Polman CH. (1999)[Interferon beta-1b nu ook voor relapsing progressieve vorm van multiple sclerosis? Nederlands Tijdschrift voor Geneeskunde 143, 815-817.

Poynard T, Naveau S, Mory B, Chaput JC. (1994) Meta-analysis of smooth muscle relaxants in the treatment of irritable bowel syndrome. Alimentary Pharmacology and Therapeutics 8, 499-510.

Pressman JL & Wildavsky A. (1973) Implementation. Berkeley etc.: The University of California Press.

PRIMS. (1998) Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing / remitting multiple sclerosis. Lancet 352, 1498-1504.

Raad voor Gezondheidsonderzoek (1998) Advies Medical Technology Assessment. Den Haag: Raad voor Gezondheidsonderzoek; publicatie 15.

Raad voor Gezondheidsonderzoek (2004) Advies HTA-onderzoek; Organisatie van het HTA-onderzoek. Den Haag: Raad voor Gezondheidsonderzoek; publicatie nr 45.

Ragnarsson G & Bodemar G. (1999) Division of the irritable bowel syndrome into subgroups on the basis of daily recorded symptoms in two outpatient samples. Scandinavian Journal Gastroenterology 10, 993-1000

Rein M & Schön D. (1993) Reframing Policy Discourse. In Fischer F & Forester J. The argumentative turn in policy analysis and planning.Durham, N.C.,: Duke University Press

Reuzel RPB & van der Wilt GJ. (2000) Health technology assessment and evaluation. Evaluation 6:4;383-398.

Reuzel RPB. (2001) Health technology assessment and interactive evaluation: different perspectives. Enschede: PrintPartners Ipskamp.

Reuzel RPB. (2004) Interactive technology assessment of paediatric cochlear implantation. Poiesis & Praxis 2: 119-137.

Reuzel RPB, Grin J, Akkerman T (2007) The role of the evaluator in an interactive evaluation of cochlear implantation: shaping power, trust, and deliberation. International Journal of Foresight and Innovation Policy 3(1): 76-94.

Rice GPA, Incorvaia B, Munari L, Ebers G, Polman C, D'Amico R, & Filippini G. (2002) Interferon-beta in relapsing-remitting multiple sclerosis (Cochrane review). In: The cochrane library, issue 1. Oxford: update software.

Rich RF. (1991) Knowledge creation, diffusion, and utilization. Perspectives of the founding editor of knowledge. Knowledge: creation, diffusion, utilization 12(3):319-337

Rich RF & Oh CH (2000) Rationality and Use of Information in Policy Decisions. A Search for Alternatives. Science Communication 22 (2) 173-211.

Rodriguez E, Gallo de Puelles P, Jovell A. (1999)The Spanish health care system: lessons for newly industrialized countries. Health Policy and Planning 14:164 -173.

Rosen R & Gabbay J. (1999) Linking health technology assessment to practice. BMJ 319, 1292-1294

Rossi PH, Freeman HE, Lipsey MW.(1999) Evaluation: a systematic approach. Thousand Oaks, CA [etc.] : Sage Publ

Rutten CL, Post D, Smelt WL. (1995) Het poliklinische preoperatieve onderzoek door de anesthesioloog. I. Minder verrichtingen en preoperatieve opnamedagen. Ned Tijdschr Geneeskd. 139(20):1028-1032.

Ryan GW. (1999) Measuring the typicality of text: using multiple coders for more than just reliability and validity checks. Human organization 58,313-322.

Sackett DL, Haynes RB, Guyatt GH, Tugwell P. (1991) Clinical Epidemiology. A basic science for clinical medicine. p223 - 238. Boston: Little, Brown and Company.

Schieber GJ. (1995) Preconditions for health reform: experiences from the OECD countries. Health Policy 32, 279-293.

Schön DA. (1983) The reflective practitioner. How professionals think in action. Basic Books, New York, USA.

Schön DA & Rein M. (1994) Frame reflection. Towards the resolution of intractable policy controversies. New York: Basic Books.

Schwartz M & Thompson M. (1990) Divided we stand. Redefining politics, technology and social choice. University of Pennsylvania Press.

Severens JL, Brokx JP, van den Broek P.(1997) Cost analysis of cochlear implants in deaf children in The Netherlands. Am J Otol. 18(6):714-718.

Smits R & Leyten A. (1991) Technology Assessment: watchdog or tracker? Towards a comprehensive technology policy.

Stake RE. (1995) The art of case study research. Thousands Oaks: SAGE.

Streiner DL & Norman GR. (1991) Health measurement scales: a practical guide to their development and use. Oxfort University Press.

Taylor D. (2001) Funding medicines for people with multiple sclerosis. BMJ 323, 1379-1380

Tellings A. (1995) The two-hundred years war in deaf education. Nijmegen: University of Nijmegen.

Topol EJ. (2004) Failing the public health - rofecoxib, Merck, and the FDA. N Engl J Med 351: 1707-1711

Trouiller P, Olliaro P, Torreele E, Orbinski, Laing R, Ford N. (2002) Drug development for neglected diseases: a deficient market and a public-health policy failure. Lancet 359: 2188-2194.

Tuinder S, van der Wilt GJ, Knoester P, Hekster Y, Keyser A, Renier W, Deckers C. (2004) Voorschrijven volgens protocol: beroepsgroep moet zelf het initiatief nemen. Medisch Contact 59(3): 79-82.

Tunis SR, Stryer DB, Clancy CM. (2003) Increasing the value of clinical research for decision making in clinical and health policy. JAMA 290: 1624-1632.

van de Graaf H & Grin J. (1999) Policy Instruments, pratiques refliches et apprentisage. Implications pour la gouvernabilité à long terme et la democratie. Espaces et Societes. no. 97-98, p. 63-90.

van den Heuvel WJA, Wieringh R, van den Heuvel LP. (1997) Utilisation of medial technology assessment in health policy. Health Policy 42, 211-222.

van der Horst HE. (1997) Irritable bowel syndrome in general practice. How effective is patient education and counselling? Wageningen, the Netherlands: Ponsen & Looyen.

van der Wilt GJ. (1995) Alternative ways of framing Parkinson's disease: Implications for Priorities in Health Care and Biomedical Research. Organization & Environment. 9(1): 13-48.

van der Wilt GJ, Rovers M, Straatman H, van der Bij S, van den Broek P, Zielhuis G. (2004) Policy relevance of Bayesian statistics overestimated? Int J Technol Assess Health Care 20(4):488-492.

van Dulmen AM. (1996) Exploring cognitions in irritable bowel syndrome. Implications for the role of the doctor. Nijmegen, the Netherlands: Quickprint.

Vermeulen M & de Haan RJ. (1999) Nieuwe therapieën in de neurologie, maar wie wordt er beter van? Ned Tijdschr Geneeskd. 143(35), 1764-1766

Verwey B, Eling P, Wientjes H, Zitman FG (2000) Memory impairment in those who attempted suicide by benzodiazepine overdose. The Journal of clinical psychiatry 61(6): 456-459.

Verwey B, Muntendam A, Ensing K, Essink G, Pasker-de Jong PC, et al. (2005) Clinically relevant anterograde amnesia and its relationship with blood levels of benzodiazepines in suicide attempters who took an overdose. Progress in neuro-psychopharmacology & biological psychiatry 29(1): 47-53.

Wahnschaffe U, Ullrich R, Riecken EO, Schulzke JD. (2001) Celiac disease-like abnormalities in a subgroup of patients with irritable bowel syndrome. Gastroenterology 121, 1329-1338.

Waller, PC, Evans, SJ, Beard, K. (2005) Drug safety and regulation. New powers and resources are needed. BMJ 331: 4-5.

Weiss CH. (1980) Knowledge creep and decision accretion. Knowledge, creation, diffusion, utilization 1(3): 381-404.

Weiss CH & Bucuvalas MJ. (1980) Social science research and decision-making. New York ; Guildford : Columbia University Press.

Weiss CH. (1993) Where politics and evaluation meet. Evaluation practice 14(1);93-106.

Williams I & Bryan S. (2007) Understanding the limited impact of economic evaluation in health care resource allocation: a conceptual framework. Health Policy 80:135-143.

Yanow D. (1996) How does a policy mean? Interpreting policy and organizational action. Georgetown: Georgetown University Press.

Yanow D. (2000) Conducting interpretative policy analysis. Thousand Oaks, Calif., [etc.]: Sage.

Ziekenfondsraad. (1995) Advies toetsing van het geneesmiddelenpakket. Amstelveen: Ziekenfondsraad.

Ziekenfondsraad. (1996) Vervolg-rapport aanspraak interferon-beta-1b. Amstelveen: Ziekenfondsraad.

Ziekenfondsraad. (1998) Rapport Jaarplan Beleidsonderzoek Geneesmiddelen 1999. Amstelveen: Ziekenfondsraad.

Abbreviations

CI	Cochlear Implant
COXIBs	Cox-2 selective inhibitors
ER	Emergency room
HCIB	Health Care Insurance Board (Dutch: College voor Zorgverzekeringen, CVZ)
HTA	Health Technology Assessment
IBS	Irritable bowel syndrome
IFNß	Interferon-beta
MS	Multiple Sclerosis
NSAIDs	Non-steroidal anti-inflammatory drugs
PH	Pulmonary hypertension
PAM	(department of) Policy Analysis of Medicines (Dutch: Beleidsonderzoek geneesmiddelen, BOG)

Summary

Improving the usefulness of Health Technology Assessment

This thesis addresses the problem of the limited impact of Health Technology Assessment (HTA) research on health policy and clinical practice and investigates the ways by which we can improve the usability of research for health policy and clinical practices. HTA is a type of policy research that aims to provide information concerning medical technologies in order to support health care decision making. Although, in many cases, HTAs appear to affect health care decisions, their actual contribution has been frequently debated. Currently, two solutions that seek to increase the applicability of research are being instituted, namely the standardisation of research methods and the active implementation of research findings in clinical practice. An alternative explanation for the limited impact of HTAs is that they do not sufficiently answer the questions deemed important by their potential users, namely policy makers and health care professionals. This alternative explanation is based on insights from the theory of argumentative policy analysis. This theory provides a possible explanation for the failed implementation of policy measures. The development and implementation of a policy is shaped by the interactions between all actors involved. The basic idea of the argumentative approach is that actors' behaviours can be explained by different views on a problem and the argumentation behind these views. The way in which a problem is defined depends on the assumptions the actors have about the situation and their beliefs regarding what is considered good practice (normative values). The implementation of policy measures may only be effective if implementers and target populations consider the proposed policy measure to be meaningful. The latter implies that the proposed policy 1) must make sense in light of their perception of the problem and 2) does not violate their normative values.

Evaluation is usually restricted to assessing whether a specific programme or intervention has, or has not, fulfilled the programme objectives. According to Frank Fischer, the main problem is that such evaluations fail to take into account the underlying assumptions and normative values which influence the evaluation of specific interventions. The outcome of an evaluation depends heavily on the criteria used to evaluate it. Which outcome measures are relevant depends on how the problem is defined and the purpose of the improvement. Fischer presented a framework for including empirical and normative concerns in evaluation. A full evaluation should not only include an assessment of the effectiveness of an

intervention but also an assessment of the more general assumptions and normative values.

HTA might not adequately take into account the possible difficulties that result from divergent problem definitions resulting from differing assumptions and values. Current HTA methods are adequate for solving well-structured problems, when actors agree on what information is required and which norms are at stake. When actors disagree on the nature of the problem, what information is required, and which normative values are at stake, the problem is ill-structured. If the problem is ill-structured, conventional research methods might lead to the over-simplification of the problem.

In this thesis, I will explore the hypothesis that it is imperative to consider the stakeholders' perspectives, even when dealing with technologies that initially appear to be less controversial. The main questions to be answered are:

1. In cases where HTA research or the implementation of subsequent policy measures failed, were the research findings or policy measures congruent with the views of all actors involved?
2. How does involving the policy's target population's influence the design of policy research? Does their participation result in research findings or subsequent policy measures that are congruent with the views of all actors involved?
3. What are the outcomes of an interactive process of problem structuring in policy research, whereby the problem definitions, underlying assumptions, and norms are analyzed and discussed? How does this impact the interventions themselves, the criteria for success, and the research questions considered relevant?

Methodology: reconstructing interpretive frames

A central method in this approach is the reconstruction of someone's interpretive frame: someone's judgement about an intervention, problem definition, underlying background theories, and normative values. This model aims to explain people's acting and assess the likelihood of people's cooperation prior to the implementation of interventions. In chapter 2, the concepts of validity and reliability in relation to this method for analysis are discussed. We presented data concerning the inter-observer agreement of this method for analysis. Reconstructing interpretive frames from verbatim transcribed interviews seems reliable to assess the cooperation of

respondents on proposed solutions. Thus, the method is useful for identifying widely endorsed policy interventions. The method is less reliable when it comes to identifying relevant issues raised by the respondents and the most important problem definitions, at least with analysts not involved in data collection.

Retrospective analysis of research projects

In the next three chapters, case studies are presented in which policy research or subsequent policy measures had failed. The case studies consisted of research projects that were commissioned by the Department of Policy Analysis of Medicines (PAM) of the Dutch Health Care Insurance Board (HCIB), an advisory board for the Ministry of Health. Based on document analysis and interviews, the interpretive frames of actors involved were reconstructed. Aim was to assess whether research results or policy measures were congruent with the problems as perceived by actors involved and their (normative) background theories.

The first case dealt with a project on the drug mebeverine for patients with the irritable bowel syndrome (Chapter 3). Prior to this project, measures to exclude the drug from the health package had failed. Stakeholders disagreed whether the drug fulfilled the criteria to qualify for reimbursement. Research was expected to provide the lacking knowledge. Due to differences in interpretive frames, however, the problem definition shifted during a preceding project on the design of the requested trial. The proposed trial did not fit the problems of the commissioners of the study. However, even if a shift in problem definition could have been prevented, it is questionable whether the requested trial could have prevented failed policy decision a second time. The proposed research did not correspond with the problems perceived by physicians and patients. The funding and usage problems relating to mebeverine were ill-structured. Actors disagreed on the information needed and the norms at stake. Unfortunately, researchers and policy makers did not acknowledge that the problem was ill-structured. As a result, the problem definition shifted during the research project and the subsequent studies endeavoured to answer the wrong question.

The second case dealt with a guideline on the use of lamotrigine set up by the HCIB (Chapter 4). A survey had shown that only a minority of neurologists were familiar with the guideline, and even fewer applied it in practice. The proposed guideline was based on the expected improper use of lamotrigine and aimed to restrict use to patients with refractory epilepsy. The guideline was based on the

available evidence from clinical trials at that time. The results from this study indicate that the problem definitions of policy maker and practising neurologists differed widely, and that the policy measure was conflicting with certain professional beliefs. The guideline might have seemed a solution for the problems of policy makers: to keep the prescription of lamotrigine low to control costs of anti-epileptic drugs. However, target populations did not cooperate and thereby policy failed. Alternatively to guideline development, the HCIB might consider ways of assisting or encouraging the professional community, e.g. by co-funding naturalistic trials that are unlikely to be funded by manufacturers. Policy makers failed to acknowledge how the target population defined problems or what argumentation gave direction to their prescription practices. This case study supports the idea that it is important to establish how target populations of policy measures experience problems and which solutions appear sensible to them at an early stage of policy development.

Based on the findings in these two case studies, one could draw the lesson that more attention should be paid to perspectives of stakeholders. In a project on interferon-beta for patients with multiple sclerosis (Chapter 5), workshops were held in which stakeholders participated. All participants seemed to agree on the desirability of specific policy measures. However, policy measures had never been implemented. Retrospective analysis of this case study revealed that no problems were perceived for which the proposed policy measures could provide a solution. At the Ministry of Health and the HCIB, the subject of IFNß was no longer a high priority. Policy makers considered the policy measures useful for actors in the field. Neurologists, however, did not perceive major problems for which a solution was needed. This study has shown that the organisation of workshops in which target populations participate does not qualify as an appropriate process of interactive policy development. From the beginning, emphasis had been put on a limited number of interventions aimed to control the expected increase in treatment costs. Although target populations participated in policy development, perceived problems and which interventions could provide a solution had been discussed insufficiently.

Implementing lessons from the case studies

Based on the findings in the case on mebeverine, the HCIB made a few changes in their process of commissioning research (Chapter 6 & 7). Besides providing specific evidence, researchers were asked to analyse the problems as they are perceived by the policy's target populations.

In chapter 6, a case study concerning policy measures for prescription of the selective cox-2 inhibitors (COXIBs)is presented. The problem behind the requested study was its frequent off label prescription and high costs. Off label was considered problematic because the uncertainty on the effect of these medicines at long term. Researchers were asked to analyse the problems as perceived by the policy's target populations. Simultaneously, researchers were asked to produce knowledge that was considered relevant to the PAM staff. Interviews showed that researchers seemed to be reluctant to conduct the requested research. Researchers doubted whether PAM's problem definition would be shared by 'the field'. At the time of finishing the study, however, one of the COXIBs was withdrawn from the market because of serious side effects. The results of this case study emphasize that involving target populations in policy research should not imply that target populations control the problem definition and identification of solutions. What went wrong was the way in which PAM staff defined the problem. The PAM's prestudy defined not only the problem, but also the underlying intentions and the direction that needed to be taken to solve the problem.

In chapter 7, a research project concerning the regulation of drugs for patients with pulmonary hypertension is presented. In this project, changes in the commissioning procedure had been implemented to prevent a shift in problem definition. Assumption was that preventing a shift in problem definition could enhance the utilization of knowledge from policy research. As part of the tendering process, the commissioner and the researchers met in an effort to reach an agreement on the research questions. Nevertheless, the participants only concluded they had different views on the situation. Researchers, who were also physicians with expertise in the subject of investigation, redefined the problem based on their experiences and assumptions. As a result, the PAM staff determined that the proposal did not fulfil their needs and thus decided not the commission this project. In hindsight, we can contend that too much emphasis was placed on preventing a shift in the problem definition. A shift in problem definitions may be necessary for solving problems.

What went wrong in both case studies in chapter 6 & 7, is that neither the commissioner, nor the researchers were open to discuss their viewpoints. Although parallel research was commissioned to identify the problems perceived by target populations, the request for evidence on the effectiveness was strongly affected by a specific solution preferred by PAM. For identifying interventions that are congruent with the perceived problems and normative values of all actors involved, problems

and possible solutions should be explored through dialogue and interaction. Preferably, the target populations of policies should be involved in the process of problem structuring to achieve agreement on the most feasible solutions to the perceived problems and the kind of knowledge that research should generate.

Interactive problem structuring

In the eighth chapter, an interactive methodology was used for problem structuring. This project was set up to identify problems and possible solutions concerning the care of patients with auto-intoxication seen at the emergency department. Despite the fact that a high level of care is often unnecessary, these patients are often admitted to the internal ward or intensive care unit. To solve the efficiency-problem, some physicians proposed a pilot study to evaluate the effectiveness of a six-hour observation unit within the emergency department. Twenty respondents from several disciplines had been interviewed, most of them twice. This case study showed that it is possible to perform an interactive process of problem structuring to assess the degree of congruence in perspectives of actors involved. Based on this information broadly supported solutions and relevant research questions were identified. The results of this process differed from a traditional approach in several ways. We concluded that the inefficiency around hospitalization was only one of the possible problem definitions. According to others, equally important was the insufficient use of existing expertise of attention paid towards underlying problems. Other interventions can be considered relevant in order to improve care. Criteria for assessing the success implementation are identified. Finally, relevant research question did not only include the evaluation of proposed interventions (solutions) but questions related to the underlying background theories supporting or contradicting the proposed interventions appeared equally important. Efficiency research rest on the assumption that both comparators are optimized and stable. Interviews revealed that this was not the case in the case study on care for intoxicated patients. Therefore, efficiency research is as yet premature.

Discussion

In chapter 9, the findings from the abovementioned case studies are discussed. In many case studies, interventions were valued differently due to differences in problem definitions. Actors disagreed on what the problem was and which information was needed. These differences in problem definitions resulted from differences in

underlying background theories and normative values. The problems were ill-structured: there was no agreement on which information was needed and actors held different normative values. It was insufficiently acknowledged that problems were ill-structured, as a result, research provided an answer to the wrong question. Consequences were that results were irrelevant to its commissioners or policy's target populations did not co-operate on proposed policy measures. Thus, another approach is needed in HTA, entailing a more extensively process of problem analysis.

Interactive problem structuring appeared feasible and lead also to different outcomes (in terms of relevant interventions and research questions) as compared to a more traditional approach. However, we also observed some difficulties when applying a problem based approach. Firstly, the adoption of an argumentative approach strongly depends on the willingness to learn from other perspectives. It is important to realise that the choice for a problem based approach is funded on normative view towards adequate policy making and the desired relation between policy research and policy making. Main assumption behind the argumentative approach is policy is seen as a process of co-production; i.e the assumption that for successful policy making the co-operation of policy's target populations is needed. Secondly, it asks for another role of research and the researchers. Aim of HTA should not be to provide just evidence, but to provide evidence that is relevant for solving problems. It is the researchers task to analyse the problem to assess to what degree actors agree, if they disagree an attempt should to structure the problem in such way that it is manageable. Thus, the researcher should adequately analyse the problem to identify which question needs to be answered. In case of insufficient congruence in perspectives, some kind of learning should take place to make perspectives more congruent.

Recommendation: Argumentative problem analysis in HTA

Conventional HTA methods need to be embedded in a more problem based approach, whereby an extensive process of problem analysis must take place prior to the design of research projects. The research process should be split up in a "problem analysis phase", "problem structuring phase" and a "knowledge providing phase". All researchers should assess the degree of structuring in order to improve the relevance of their research: argumentative problem analysis. Firstly, the researcher assess which actors are involved, who need to cooperate on possible interventions or are likely to experience its consequences. Next step is to assess how they define the problem, judge

interventions and the normative values they held. Based on this information an estimate should be made on the degree of congruence in perspectives. In case of well structure problems, conventional HTA methods can be used to provide the requested information. In case of ill-structured problems, the problem should structured in order to reach more congruence to be able to identify broadly supported interventions. Methods for interactive Technology Assessment can be used for this purpose.

Samenvatting

Verbeteren van de bruikbaarheid van Health Technology Assessment

In hoofdstuk 1 wordt een inleiding gegeven op het onderwerp van dit proefschrift: de bruikbaarheid van Health Technology Assessment (HTA) en de noodzaak tot een goede probleemstructurering. Er is discussie over de beperkte impact van HTA op besluitvorming in de klinische praktijk of op nationaal niveau. Diverse artikelen zijn geschreven waarin is ingegaan op de vraag hoe het gebruik van onderzoeksresultaten te verbeteren. In het algemeen kunnen hierbij twee strategieën worden onderscheiden: het standaardiseren van onderzoeksmethoden of het onderzoeken van strategieën voor het verbeteren van de implementatie van resultaten in de praktijk. Aanvullend op deze gangbare strategieën, wil ik een andere oplossing introduceren, gebaseerd op inzichten vanuit de theorie van de argumentatieve beleidsanalyse. Een alternatieve verklaring voor de beperkte impact van HTA is dat de resultaten onvoldoende aansluiten bij de problemen zoals die worden ervaren door beoogde gebruikers. De argumentatieve beleidsanalyse is één van de theorieën die een verklaring wil bieden voor problemen met de implementatie van beleidsmaatregelen. De ontwikkeling en implementatie van beleid wordt gevormd door interacties tussen betrokken actoren. De kern van deze theorie is dat het handelen van actoren verklaard kan worden op basis van de wijze waarop iemand een probleem definieert en de onderliggende argumentatie.

In veel gevallen is een evaluatie beperkt tot het meten van uitkomsten. Volgens Frank Fischer is het probleem dat dergelijke evaluaties de onderliggende veronderstellingen en normatieve waarden onvoldoende in beschouwing nemen. De uitkomsten van een evaluatie zijn sterk afhankelijk van de gehanteerde criteria. Welke uitkomstmaten van belang zijn hangt af van de wijze waarop het probleem gedefinieerd is en wat het doel is van de verandering. Fischer presenteerde een raamwerk voor een volledige evaluatie. Een volledige evaluatie omvat zowel een beoordeling van de effectiviteit van interventies als de meer algemene veronderstellingen en normatieve waarden.

In HTA wordt onvoldoende rekening gehouden met het feit dat problemen verschillend gedefinieerd kunnen worden door betrokkenen. Huidige kwantitatieve onderzoeksmethoden zijn vooral geschikt voor zogenaamde goed gestructureerde problemen: situaties waarbij betrokkenen het eens zijn over welk informatie gewenst

is om een probleem op te kunnen lossen en welke normen daarbij in het geding zijn. Indien betrokken het niet eens zijn over welke informatie er nodig is en er tevens verschillende normatieve waarden in het geding zijn is er sprake van een ongestructureerd probleem. Indien er sprake is van een ongestructureerd probleem, zal de conventionele benadering leiden tot een vereenvoudiging van het probleem, waardoor het probleem slechts gedeeltelijk of het verkeerde probleem wordt opgelost.

In dit proefschrift behandel ik de hypothese dat het wenselijk is in HTA rekening te houden met de verschillende perspectieven van betrokken actoren, ook indien men te maken heeft met onderwerpen die op het eerste gezicht niet zo controversieel zijn. Onderzoeksvragen zijn:

1. In gevallen waarin HTA onderzoek of de implementatie van daarop gebaseerde beleidsmaatregelen is gefaald, in hoeverre waren de onderzoeksresultaten of de beleidsmaatregelen congruent met de perspectieven van betrokken actoren (aanluitend bij de wijze waarop het probleem werd gedefinieerd en niet strijdig met normatieve veronderstellingen)?
2. Wat zijn de resultaten van onderzoek en daaropvolgende beleidsmaatregelen indien beoogde doelgroepen van het beleid participeren in beleidsonderzoek? Leidt participatie tot onderzoeksresultaten en beleidsmaatregelen die congruent zijn met de perspectieven van betrokken actoren?
3. Wat zijn de uitkomsten van een interactief proces van probleemstructurering, waarbij probleemdefinities, onderliggende veronderstellingen en normatieve voorkeuren worden geanalyseerd en bediscussieerd? Is het van invloed op de selectie van interventies om een oplossing te bieden voor de problemen, de criteria om de interventies te evalueren en de onderzoeksvragen die als relevant worden ervaren?

Methodologie: reconstructie van handelingstheorieën

Een belangrijke analysemethode bij deze benadering is de reconstructie van handelingstheorieën: de combinaties van iemands oordeel over een interventie, de wijze waarop het probleem is gedefinieerd, onderliggende (normatieve) achtergrondtheorieën, en iemands diepere voorkeuren. Dit model beoogt het handelen van actoren te verklaren en de medewerking van actoren aan bepaalde interventies vooraf in te schatten. In hoofdstuk 2 worden de concepten van validiteit en betrouwbaarheid in relatie tot deze analysemethode bediscussieerd. De resultaten van een analyse van de inter-beoordelaarsvariatie bij het reconstrueren van

handelingstheorieën zijn gepresenteerd. Het reconstrueren van handelingstheorieën op basis van uitgetypte interviews is betrouwbaar als het gaat om het inschatten van iemands medewerking aan bepaalde interventies. Daarmee is de methode geschikt om interventies te identificeren die waarschijnlijk op brede steun kunnen rekenen. De betrouwbaarheid van het identificeren van de belangrijkste problemen voor een actor was echter matig.

Retrospectieve analyse van case studies m.b.v. argumentatieve beleidsanalyse

In hoofdstuk 3, 4 en 5 worden drie case studies gepresenteerd waarin beleidsonderzoek of de implementatie van de daaropvolgende beleidsmaatregelen was gefaald. De case studies hebben betrekking op onderzoek uitgezet door de afdeling Beleidsonderzoek Geneesmiddelen (BOG) van het College voor Zorgverzekeringen. Aan de hand van documentanalyse en interviews zijn de handelingstheorieën van betrokken actoren gereconstrueerd. Met behulp hiervan is nagegaan in hoeverre de onderzoeksresultaten of voorgestelde beleidsmaatregelen aansloten bij de problemen van betrokken actoren en niet strijdig waren met hun normatieve achtergrondtheorieën.

In hoofdstuk 3 wordt de casus over het geneesmiddel mebeverine voor patiënten met prikkelbaar darmsyndroom gepresenteerd. Voorafgaande aan het onderzoeksproject was door het ministerie geprobeerd het middel uit het verstrekkingenpakket te halen. Deze maatregel mislukte echter. Betrokken actoren waren het niet eens over de vraag of het geneesmiddel aan alle criteria voor vergoeding voldeed. Onderzoek naar de effectiviteit van het middel werd uitgezet en de verwachting was dat de resultaten hiervan bruikbaar waren om alsnog succesvol beleid uit te voeren. Gedurende het onderzoekstraject trad er een verschuiving in de probleemdefiniëring op. Het uiteindelijke voorstel voor een onderzoek sloot onvoldoende aan bij de probleemdefiniëring volgens de opdrachtgevers. Echter, ook indien een verschuiving in de definiëring van het probleem kon worden voorkomen, is het niet waarschijnlijk dat het gevraagde onderzoek de problemen zou hebben opgelost. Het onderzoek sloot onvoldoende aan bij de problemen zoals die werden ervaren door potentiele doelgroepen van het beleid. Het probleem rondom de middelen voor het prikkelbare darmsyndroom was ongestructureerd. Dit werd onvoldoende erkend, waardoor tijdens het project een verschuiving in de definiëring van het probleem optrad en de uiteindelijke studie een antwoord op de verkeerde vraag zou geven.

In hoofdstuk 4 is een casus beschreven over een richtlijn opgesteld door het College voor Zorgverzekeringen over het voorschrijven van het anti-epilepticum lamotrigine. De aanbevelingen in de richtlijn zijn gebaseerd op de beschikbare kennis uit wetenschappelijke studies op dat moment. De kern van de richtlijn was dat het voorschrijven van lamotrigine beperkt moest worden tot patiënten met moeilijk behandelbare epilepsie. De reden voor het opstellen van de richtlijn was de verwachting dat het middel inadequaat zou worden voorgeschreven. Een recente enquête toonde dat de richtlijn in de praktijk nauwelijks gebruikt werd. Interviews lieten zien dat beleidsmakers en neurologen het probleem verschillende definieerden. De beleidsmaatregel leek een oplossing voor de problemen zoals gedefinieerd door beleidsmakers: het beperken van het voorschrijven van lamotrigine om de kosten laag te houden. De doelgroepen van het beleid werkten echter niet meer, waardoor het beleid faalde. Een mogelijk alternatieve maatregel om het adequaat voorschrijven van dit middel te bevorderen, is het ondersteunen van de medische professionals, bijvoorbeeld door het financieren van gewenste studies (in natuurlijke setting) die door fabrikanten nooit zouden worden gefinancierd. Deze studie laat zien dat het wenselijk is om bij beleidsontwikkeling de doelgroepen van het beleid in een vroegtijdig stadium te betrekken.

Op basis van deze twee studies kan de conclusie worden getrokken dat het wenselijk is om de perspectieven van doelgroepen van beleid mee te nemen in beleidsondersteunend onderzoek. In een project over beleidsmaatregelen voor het geneesmiddel interferon-beta voor patiënten met Multipele Sclerose (hoofdstuk 5) zijn deze doelgroepen gevraagd te participeren in workshops. Tijdens de workshops leken alle betrokkenen het eens te zijn over de wenselijkheid van een aantal concrete beleidsmaatregelen. Echter, deze maatregelen zijn uiteindelijk nooit geïmplementeerd in de praktijk. Een mogelijke verklaring is dat er door de betrokkenen geen problemen werden ervaren waarvoor deze maatregelen een oplossing konden bieden. Bij het ministerie en het College voor Zorgverzekeringen had het onderwerp geen hoge prioriteit meer. Zij gaven aan dat de maatregelen vooral voor de beroepsgroep meerwaarde konden hebben. De neurologen echter ervaarden geen problemen waarvoor deze maatregelen een oplossing zouden bieden. Deze studie laat zien dat het organiseren van workshops waaraan doelgroepen van beleid deelnemen niet gelijk is aan interactieve beleidsontwikkeling. Een mogelijke verklaring voor het verloop van het project is dat in het project veel nadruk lag op een beperkt aantal beleidsmaatregelen en niet zozeer de problemen waarvoor oplossingen nodig waren.

Tevens werden onderliggende veronderstellingen en achtergrondtheorieën onvoldoende bediscussieerd.

Implementeren van lessen uit retrospectieve analyse van case studies

Naar aanleiding van de bevindingen in de cases over mebeverine heeft het BOG een aantal wijzigingen doorgevoerd in hun aanbestedingsprocedures. Twee projecten waarin de vernieuwde aanpak is toegepast zijn beschreven in hoofdstuk 6 en 7. In deze projecten werd aan de onderzoekers gevraagd om niet alleen bepaalde kennis aan te leveren maar ook de probleemdefinities volgens betrokken actoren te achterhalen en mee te nemen in het project.

In hoofdstuk 6 wordt een casus beschreven over beleidsmaatregelen rond de selectieve cox-2 remmers. Het probleem was dat deze nieuwe medicijnen op grote schaal off-label werden voorgeschreven. Achtergrond was dat er onvoldoende kennis was over de effecten van deze middelen op langere termijn. Onderzoekers werd gevraagd om na te gaan hoe de doelgroepen van beleid problemen rondom de inzet van deze middelen ervaren. Tegelijkertijd dienden zij bepaalde informatie te leveren die relevant was volgens het College voor Zorgverzekeringen. Uit interviews met de onderzoekers die een voorstel hadden ingediend bleek dat zij kritisch waren ten aanzien het gevraagde onderzoek. Volgens de onderzoekers werd de wijze waarop het probleem was gedefinieerd niet gedeeld door artsen, waardoor er weinig draagvlak zou zijn voor maatregelen. Tijdens de afronding van het onderzoeksproject werd één van de middelen van de markt gehaald in verband met ernstige bijwerkingen. Het betrekken van doelgroepen bij onderzoek en beleidsontwikkeling wil dus niet zeggen dat deze doelgroepen volledige zeggenschap moeten krijgen over de wijze waarop het probleem gedefinieerd is. Echter, in de beleidsanalyse werd niet alleen het probleem op een bepaalde wijze gedefinieerd waar doelgroepen kritiek op hadden, maar hierin werd ook een duidelijke oplossingsrichting aangegeven.

In hoofdstuk 7 wordt een casus beschreven over beleid rond twee nieuwe, dure geneesmiddelen (epoprostenol en bosentan) voor patiënten met pulmonale hypertensie. In dit project werden maatregelen genomen om de kans op een verschuiving in de probleemdefiniëring te verkleinen. De veronderstelling was dat het voorkomen van een verschuiving in de probleemdefiniëring de bruikbaarheid van uiteindelijke resultaten zou verbeteren.. Tijdens de aanbesteding van het onderzoek werd een bijeenkomst georganiseerd waarbij overleg tussen de onderzoekers en opdrachtgever mogelijk was. Desondanks trad er probleemtransformatie op in één

van de deelprojecten. De onderzoekers, medisch specialisten, herdefinieerden het probleem bewust op basis van hun eigen inzichten en ervaringen. BOG medewerkers besloten dit onderdeel niet aan te besteden. Achteraf gezien werd er teveel nadruk gelegd op het voorkomen van probleemtransformatie, terwijl een herdefiniëring van het probleem soms noodzakelijk kan zijn om problemen op te lossen.

Het probleem was dat in beide projecten betrokkenen (opdrachtgever en de onderzoekers) niet echt open stonden om hun perspectieven op de situatie te bediscussiëren. Ondanks het feit dat onderzoek was uitgezet om de probleemdefinities volgens doelgroepen van het beleid te identificeren, werd parallel onderzoek uitgezet dat als doel had specifieke informatie te leveren die relevant was volgens het perspectief van de opdrachtgever. Idealiter, zouden doelgroepen betrokken moeten worden bij de probleemanalyse om vervolgens na te gaan welke interventies een oplossing kunnen bieden en welk onderzoek wenselijk is.

Interactieve probleem structurering

In hoofdstuk 8 zijn de resultaten van een interactieve procedure voor probleemstructurering gepresenteerd. Het project had betrekking op het identificeren van ervaren problemen en gewenste oplossingen rondom de acute opvang van patiënten in een ziekenhuis na intoxicatie met een overdosis (genees)middelen. Aanleiding voor het onderzoek was de ervaring dat veel patiënten werden opgenomen op een intensive care, terwijl - achteraf gezien - dit vaak niet nodig was. Een van de betrokkenen stelde voor de doelmatigheid te onderzoeken van een 6-uurs observatie van deze patiënten op de spoedeisende hulp van het ziekenhuis. Twintig actoren werden geïnterviewd, waarvan het merendeel tweemaal, en gevraagd naar ervaren problemen en onderliggende veronderstellingen en argumenten. Met behulp van deze interactieve methode was het mogelijk de mate van congruentie in perspectieven vast te stellen. Vervolgens konden interventies worden geïdentificeerd die op breed draagvlak zouden kunnen rekenen onder de betrokkenen en vervolgens relevante onderzoeksvragen. De uitkomsten van deze interactieve analyse verschillenden van de traditionele benadering op een aantal punten. Allereerst bleek de ondoelmatigheid rond het opnamebeleid slechts één van de probleemdefinities. Volgens anderen waren problemen rond het onvoldoende gebruik maken van aanwezige expertise in het ziekenhuis even belangrijk. Andere interventies werden genoemd om de zorg voor deze patiëntenpopulatie te verbeteren. Ook konden criteria voor goede zorg worden vastgesteld. Tot slot waren niet alleen onderzoeksvragen met betrekking op het

evalueren van effecten van interventies relevant, maar ook vragen naar onderliggende achtergrondtheorieën die bepaalde interventies ondersteunden of tegenspraken. Een belangrijke veronderstelling achter huidig doelmatigheidsonderzoek is dat de interventie die onderwerp van onderzoek volledig is uit-ontwikkeld en geoptimaliseerd. De interviews toonden aan dat het niet het geval was bij dit onderwerp. Doelmatigheidsonderzoek was daarmee nog voorbarig.

Discussie

In het slothoofdstuk worden de bevindingen uit de case studies bediscussieerd. In de meeste case studies werden interventies verschillend beoordeeld door betrokken actoren door verschillen in de probleemdefinities. Actoren waren het niet eens over wat het probleem was en welke informatie nodig was. De verschillen in probleemdefinities kunnen grotendeels worden toegeschreven aan verschillen in achtergrondtheorieën en diepere voorkeuren. De problemen waren ongestructureerd: er was geen overeenstemming over welke kennis nodig was en de actoren hielden verschillende normatieve voorkeuren. Tijdens de projecten werd er onvoldoende rekening mee gehouden dat de problemen ongestructureerd waren. Onderzoek gaf vaak een antwoord op de verkeerde vraag. Als gevolg daarvan waren de resultaten niet bruikbaar voor de opdrachtgever of was er geen medewerking van doelgroepen aan voorgestelde beleidsmaatregelen. Kortom, meer aandacht is nodig voor uitgebreide probleemanalyse in HTA.

Een interactieve probleemanalyse bleek uitvoerbaar en leidt mogelijk ook tot andere uitkomsten in vergelijking met de traditionele benadering in HTA. Echter, bij de uitvoering van een dergelijke benadering werden ook enkele moeilijkheden ervaren. Allereerst is de adoptie van de argumentatieve benadering sterk afhankelijk van de bereidheid om van andere perspectieven te leren. Het is van belang te realiseren dat een probleemgerichte benadering een normatieve keuze omvat ten aanzien van adequate beleidsontwikkeling en de gewenste relatie tussen beleidsonderzoek en beleidsontwikkeling. Een belangrijk uitgangspunt is dat beleid wordt gezien als een proces van co-productie wat veronderstelt dat de medewering van doelgroepen van het beleid nodig is om het beleid succesvol te laten zijn. Ten tweede vergt deze benadering een andere rol voor HTA onderzoek en onderzoekers. Het doel van HTA is niet zozeer het leveren van kennis, maar het leveren van kennis dat relevant is om bepaalde problemen op te lossen. Het is de taak van de onderzoeker om problemen te analyseren en de mate van structurering van het

probleem vast te stellen. Indien betrokken actoren het niet eens zijn, dient de onderzoeker het probleem te structureren op zodanige wijze dat het hanteerbaar is. Indien er onvoldoende congruentie in perspectieven is moet er een proces van leren optreden om meer congruentie te bewerkstelligen. Dit is nodig om oplossingen te kunnen identificeren die op een breed draagvlak kunnen rekenen bij de implementatie.

Aanbeveling: interactieve probleem analyse in HTA

Huidige methoden voor HTA onderzoek zouden moeten worden ingebed in een meer probleem gerichte benadering, waarbij het onderzoek moet worden voorafgegaan door een uitgebreide probleemanalyse. Bij het HTA onderzoek moet onderscheid worden gemaakt tussen een "fase van probleem analyse", een "fase van probleem structurering" en een "kennis producerende fase". Alle onderzoekers moeten de mate van structurering vaststellen voorafgaande aan het onderzoek om de relevantie van het onderzoek te verhogen door middel van interpretatieve probleem analyse. In de eerste plaats moet de onderzoeker nagaan welke actoren bij het onderwerp van studie betrokken zijn, actoren wiens medewerking nodig is bij het uitvoeren van interventies of actoren die mogelijk consequenties hiervan ervaren. In de tweede plaats moet worden nagegaan hoe deze actoren het probleem definiëren, interventies beoordelen en welke normatieve voorkeuren daaraan ten grondslag liggen. Ten derde, op basis van deze informatie dient te worden nagegaan in welke mate het probleem gestructureerd is. Indien het probleem goed gestructureerd is, kunnen conventionele methoden voor onderzoek (kwantitatief onderzoek; bijvoorbeeld kosten-effectiviteitsstudies) worden toegepast om de gewenste informatie te leveren. Indien het probleem slecht gestructureerd is, is verdere structurering van het probleem nodig om meer congruentie in perspectieven te bewerkstelligen. Interactieve onderzoeksmethoden kunnen hiervoor worden toegepast.

Curriculum Vitae

Margriet Moret-Hartman werd op 27 april 1976 geboren in Opijnen. De middelbare school werd gevolgd op het Koningin Wilhelmina College in Culemborg. In september 1994 begon zij met de studie Biomedische Gezondheidswetenschappen aan de Radboud Universiteit te Nijmegen. Tijdens het laatste jaar van haar studie was zij aangesteld als student-assistent op de afdeling Medical Technology Assessment (MTA) van de Universitair Medisch Centrum St Radboud te Nijmegen. Binnen één van de ontwikkelingsgeneeskunde projecten over het optimaliseren van behandeling van patiënten met reumatoïde artritis met metrotrexaat combineerde zij het uitvoeren van haar grote stage en het studentassistentschap. De analyses die betrekking hadden op het effect van de behandeling op kwaliteit van leven scores was onderdeel van de stage. De kosten-effectiviteitsanalyse werd als student assistent uitgevoerd. In 1999 studeerde zij *cum laude* af met als afstudeerrichting epidemiologie. Na het afstuderen werd zij als junior onderzoeker aangesteld aan de afdeling MTA. Daarbij heeft zij eerst enkele jaren gewerkt aan de afronding van enkele doelmatigheidsprojecten, onder andere op het gebied van dermatologie, traumatologie en reumatologie. Gaandeweg raakte zij geïnteresseerd in vraagstukken over de bruikbaarheid van de resultaten van de doelmatigheidsstudies. Daarnaast heeft zij onderwijs verzorgd op het gebied van Health Technology Assessment en Evidence Based Medicine binnen de opleidingen Biomedische Wetenschappen en Geneeskunde van de Radboud Universiteit en de opleiding Advanced Nursing Practice van de Hogeschool Arnhem Nijmegen. Van mei 2006 tot en met april 2008 was zij werkzaam als adviseur richtlijnontwikkeling binnen de vakkern Evidence Based Medicine van het Kwaliteitsinstituut voor de Gezondheidszorg CBO. Sinds mei 2008 werkt zij als staffunctionaris patiëntveiligheid in Gelre Ziekenhuizen te Apeldoorn. Zij is getrouwd met Gerben Moret en samen hebben ze een zoon, Bram.

Publicaties

M Moret-Hartman, R Reuzel, J Grin, GJ van der Wilt. Participatory workshops are not enough to prevent policy implementation failures: an example of a policy development process concerning the drug interferon-beta for multiple sclerosis. Health Care Analysis 2008;16(2):161-175

M Moret-Hartman, GJ van der Wilt, J Grin. Health Technology Assessment and ill structured problems: a case study concerning the drug mebeverine. International Journal of Technology Assessment in Health Care 2007; 23(3):316-323

M Moret, RPB Reuzel, GJ van der Wilt, J Grin. Validity and reliability of qualitative data-analysis: reconstructing interpretative frames. Field Methods 2007; 19(1):24-39

M Moret-Hartman, PD Knoester, YA Hekster, GJ van der Wilt. Non-compliance on the part of the professional community with a national guideline: An argumentative policy analysis. Health Policy 2006;78:353-359

M Hartman, AE van Ede, JL Severens, RFJM Laan, LBA van de Putte, GJ van der Wilt. Economic evaluation of folate supplementation during methotrexate treatment in rheumatoid arthritis. The Journal of Rheumatology 2004; 31(5): 902-908.

M Hartman, M Prins, OQ Swinkels, JL Severens, T De Boo, GJ van der Wilt, PC van de Kerkhof, PG van der Valk. Cost-effectiveness analysis of a psoriasis care instruction programme with dithranol compared with UVB phototherapy and inpatient dithranol treatment. British Journal of Dermatology 2002;147(3):538-544

M Hartman, MM Rovers, K Ingels, GA Zielhuis, JL Severens, GJ van der Wilt. Economic evaluation of ventilation tubes in otitis media with effusion. Arch. Otolaryngol. Head Neck Surg. 2001; 127(12): 1471-1476

N Vink-van Kimmenade, AJM Keyser, J Groenewoud, M Moret-Hartman, YA Hekster. Gestructureerd patiëntendossier Proces tot implementatie als kwaliteitsinstrument en eerste resultaten in de patiëntenzorg. Pharmaceutisch Weekblad 2006; 141(20), 696-699.

PM Welsing, JL Severens, M Hartman, AM van Gestel, PL van Riel, RF Laan. The initial validation of a Markov model for the economic evaluation of (new) treatments for rheumatoid arthritis. Pharmacoeconomics. 2006;24(10):1011-1020.

J Grin, M Moret-Hartman, GJ van der Wilt. Beleidsonderzoek en versterking van het primaat van de politiek in de uitvoering - lessen uit de gezondheidszorg. Beleidswetenschap 2004 /4

PM Welsing, JL Severens, M Hartman, PLCM van Riel, and RFJM Laan. Modeling the 5-Year Cost Effectiveness of Treatment Strategies Including Tumor Necrosis Factor-Blocking Agents and Leflunomide for Treating Rheumatoid Arthritis in the Netherlands. Arthritis & Rheumatism, Arthritis Care & Research 2004; 51(6): 964-973.

AE van Ede, Laan RFJM, Rood MJ, Huizinga TWJ, van de Laar MAFJ, van Denderen, JC, Westgeest AAA, Romme TC, de Rooij DJRAM, Jacobs MJM, de Boo ThM, van der Wilt GJ, Severens JL, Hartman M, Krabbe PFM, Dijkmans BAC, Breedveld FC en van de Putte LBA. Het effect van comedicatie met foliumzuur of folinezuur op toxiciteit en effectiviteit van methotrexaat bij de behandeling van reumatoïde artritis; een gerandomiseerde, dubbelblinde, placebogecontroleerde studie gedurende 48 weken. Ned Tijdschr Geneeskd. 2002; 146; 1322-1327

AE van Ede, Laan RF, Rood MJ, Huizinga TW, van de Laar MA, van Denderen CJ, Westgeest TA, Romme TC, de Rooij DJ, Jacobs MJ, de Boo TM, van der Wilt GJ, Severens JL, Hartman M, Krabbe PF, Dijkmans BA, Breedveld FC, van de Putte LB. Effect of folic or folinic acid supplementation on the toxicity and efficacy of methotrexate in rheumatoid arthritis: a forty-eight week, multicenter, randomized, double-blind, placebo-controlled study. Arthritis Rheum. 2001;44(7):1515-1524.

Dankwoord

Tot slot een woord van dank voor velen zonder wie dit proefschrift er nooit had kunnen liggen. Een aantal mensen wil ik speciaal noemen.

Allereerst wil ik mijn promotoren, Gert Jan van der Wilt en John Grin, danken voor hun kennis en expertise waar ik veel van geleerd heb. Het was fijn om samen met jullie het vraagstuk van de bruikbaarheid van HTA onderzoek uit te kunnen werken. Onze bijeenkomsten waren altijd zeer inspirerend. Gert Jan, je enorme kennis en kritische blik zijn zeer waardevol geweest. Na elk overleg had ik weer veel stof om na te denken. John, jij hebt me geleerd kritisch te zijn ten opzichte van aspecten die binnen de medische wetenschappen zo vanzelfsprekend zijn. Het was zeer leerzaam om begeleid te worden door twee personen vanuit verschillende vakgebieden. Het helpt je een eigen weg te zoeken en ideeën scherp te krijgen.

Rob Reuzel, co-promotor en jarenlang mij kamergenoot. Ik heb met veel plezier met je samengewerkt in Nijmegen. Ik heb veel geleerd van jouw ervaring op het gebied van interactief onderzoek. Behalve stil op de kamer (tenminste, dat viel anderen altijd op), was het vooral ook gezellig. Van gesprekken over serieuze dingen tot af en toe een muzikaal intermezzo (hoog tijd voor een volgend deel?).

Daarnaast wil ik iedereen danken zonder wie ik de projecten niet hadden kunnen uitvoeren. Dank voor mensen van het College voor Zorgverzekeringen voor de mogelijkheid een aantal onderzoeksprojecten kritisch te evalueren. Bert Boer, Yolanda Waterreus, Marcel Eijgelshoven, Nicolien Wieringa, en Floor Rikken, het was leerzaam en goed om één dag per week in Amstelveen en later Diemen mee te lopen. Dankzij een aanvullende financiering van het instituut voor Wetenschap & Samenleving kon ik mijn onderzoek uitbreiden met een interactieve analyse in het UMC St Radboud.

Ik wil alle respondenten danken die bereid waren met mij te spreken over hun werkgebied, problemen en ideeën. Zonder jullie zou ik als onderzoeker niets in handen hebben gehad. Ik hoop dat ik met dit proefschrift een bijdrage kan leveren aan het verbeteren van de relevantie van onderzoek voor de praktijk.

Ook wil ik alle oud-collega's van de afdeling Medical Technology Assessment noemen. Ik heb altijd met veel plezier op de afdeling gewerkt. Hans Severens, jij introduceerde mij tijdens het blok 'kosten-effectiviteitsanalyse' in het vakgebied van MTA. Eddy, Paul, Hans G., Peep, Wija, Wietske en Anke, dank voor de prettige samenwerking binnen onderzoeksprojecten en het onderwijs. Fokkelien en Tilly, de

gezellige uurtjes tijdens lunchpauzes, etentjes en afdelingsuitjes hebben daaraan zeker bijgedragen. Collega-docenten binnen het UMC St Radboud en de HAN, dank voor de prettige samenwerking. Onderwijs naast het onderzoek houdt je scherp. De vragen van studenten dwingen je de lesstof door en door te beheersen. Daarnaast is het vooral ook leuk je kennis over te kunnen dragen aan anderen.

Vrienden en familie, gelukkig was er naast het werk ook tijd voor leuke dingen (al lagen de sociale contacten de afgelopen twee jaar op een iets lager pitje in verband met de afronding van dit proefschrift en de veranderingen thuis). Dank voor de gezellige momenten en jullie interesse. Ik heb een club van goede, inspirerende vrienden overgehouden aan de studietijd in Nijmegen. Joop, onze 'wedstrijd' bleef spannend tot het eind. Binnenkort sta jij hier.

Joris en Steven, ik vind het leuk dat jullie als paranimfen naast mij staan. Jullie steun is zeer waardevol.

Een speciaal woord van dank aan mijn ouders. Van jullie heb ik geleerd om iets aan te pakken en gewoon bij het begin te beginnen. Of het nu gaat om het opknappen van een oud huis of het schrijven van een proefschrift. Dank voor alle hulp en steun. En ma, met het gummetje dat je mij bij mijn afstuderen gaf, wordt de 's' nu weggepoetst.

Tot slot, lieve Gerben, mijn maatje en veel meer dan dat. De afgelopen periode heb ik gemerkt hoe bijzonder onze band is. Na een periode vol spannende ontwikkelingen gaan we nu lekker genieten ... samen met Bram.