Chapter 8:
General discussion
1 Background

The evaluation of work disability has a large individual and societal impact. Every year in the Netherlands tens of thousands of people are being evaluated on their claimed disability and granted or denied a benefit. The way these evaluations are carried out is a result of historical development and has changed considerably over the past two decades [1]. How can one be sure that laws on social insurance are being implemented correctly? As described in the Introduction to this thesis, in the Netherlands the quality of the evaluations that social insurance physicians (SIPs) perform is being debated. Individual SIPs have different ideas on what they are evaluating and how that evaluation can best be performed. Claimants and Institute for Employee Benefit Schemes ('Uitvoering Werknemersverzekeringen' or UWV) are implicated in this debate, too. A common understanding is needed in terms of what quality is and what quality can be expected and who is responsible for what aspect of quality control.

In this thesis the following questions are addressed:
1  What is the object of the evaluation of work disability?
2  What is to be understood by the quality of the evaluation of work disability?
3  How can the quality of evaluation of work disability be controlled?

Six studies have been completed:
1  Organisation of disability evaluation: a survey in fifteen countries
2  Medicolegal reasoning: a focus group and questionnaire study in four countries
3  Guidelines for evaluating work disability: an international survey
4  Evidence-based guidelines for evaluating disability: a comparison in two countries
5  Disability interview protocols in the Netherlands: a comparison of prescriptions and principles
6  Interviews for the evaluation of long-term incapacity for work: a study on adherence to protocols and principles

In this General Discussion attention is paid to the results of the different studies (par 2) and several methodological remarks made (par 3). Answers to the questions are formulated (par 4) and recommendations given (par 5).
Chapter 2: Organisation and quality control of evaluation of disability in different countries.

In thirteen European countries, the United States, and the Russian Federation the organisation of disability evaluations was studied. Similarities and differences were found among legal criteria and in the way these criteria are used. In all countries studied, legal criteria for disability refer to 1) the claimant’s ability (or inability) to perform work that can reasonably be expected from a worker in that profession, 2) health conditions that account for these abilities or inabilities, and 3) opportunities and obligations to receive treatment and return to work activities. These legal criteria are congruent with the criteria of the handicapped role. Legal criteria differ also; for instance, ability and inability are phrased as labour capacity or earning capacity and are rarely operationalised. The Dutch earning capacity is operationalised at the participation level of ICF while the British labour capacity is operationalised at ICF’s activities’ level. Among countries, differences exist in levels of disability, varying from one to seven. Finally, the waiting time before disability can be recognised is either fixed or variable and, in the first case, varies between six months and five years.

The organisation of the process of evaluation differs considerably among countries as well. The precise steps in the process and the connections with health care and labour market are organised quite differently. Legal criteria are operationalised in each country at the Institute of Social Insurance (ISI), as different types of output requirements for the social insurance physicians (SIPs): medical, functional, and rehabilitational. Most often, however, a country combines two or even three types.

The organisation of the evaluations can be seen as an interplay between actors that can be described using the extended script model. Actors in this organisation can be the SIP, the claimant, the ISI, an external supervisory body, the tribunals, the medical professional group, and the lawgiver. Explicit definitions of evaluation of quality were not found. An implicit definition at the legal level would be compliance with legal requirements; at the organisational level it would be compliance with organisational prescriptions, and at the professional level, compliance with professional standards.

Quality control is both indirect and direct. Indirect is, for example, the requirement by ISI that SIPs be qualified medical doctors. Indirect is the time and instruments doctors have for the evaluation and the specification of the required output. Direct quality control is seen in the inspection of case reports by staff doctors at the ISI. Two approaches emerge in quality control: individual SIPs who are monitored by the ISI and SIPs participating in medical committees who steer themselves.
Chapter 3 Medico-legal reasoning: a focus group and questionnaire study

In four countries with different types of output requirements for their SIPs, the medico-legal reasoning of practising SIPs was studied in a case of an elderly construction worker with lower back pain. SIPs were asked to express and agree about the grounds they thought valid in their argumentation. This was first done in focus groups and secondly with a questionnaire, using grounds for argumentation as indicators of reasoning. SIPs in all countries studied proved to interpret disability in a way that meets legal criteria and the handicapped role. Added to that is a requirement of a fair trial. The handicapped role overlaps with the ICF: the situation of participation problems, activity limitations, and impairments, influenced by personal and environmental factors can be seen as an operationalisation of the health condition in the handicapped role. The grounds on the claimant’s health condition and the grounds of medical evidence mostly are grounds for working capacity as well. Fair trial refers to plausibility, consistency (which can relate to ICF), exclusion of non-health related reasons for not working, and inclusion of the personal experience of the claimant.

The differences in output that were found in the study on organisation were not replicated in this study on medico-legal reasoning. The medico-legal reasoning does not simply follow the output requirements of the ISI but the more general handicapped role.

The grounds guide the translation of information into arguments about work disability. It is possible to make these grounds explicit.

Chapter 4 Guidelines for evaluation of work disability

From fourteen European countries the operationalisation of the evaluation of work disability and the use of guidelines were reported by central medical advisers and their staff. Five countries evaluate work disability in terms of all aspects of the handicapped role. In nine countries some aspects are not mentioned. Several countries report correspondence with ICF but nowhere the correspondence is made explicit. Official guidelines in social insurance medicine for evaluating work disability are found in Germany, Ireland, the Netherlands and Switzerland. Guidelines can be characterised as medical or procedural. Common topics of medical guidelines are the medical condition itself (including origin, risk factors, course, diagnostic procedures, treatment and ICD classification); return to work activities; and evaluation of work disability. This is an operationalisation of the handicapped role.

Common topics of the procedural guidelines are descriptions of the concepts of disease and incapacity to work as evaluated in social insurance, independence of SIP between claimant and ISI, qualification of the SIP, ways to determine the existence of disease, ways to determine functional capacity, and quality criteria of evaluations. These can all be seen as operationalisations of the concept of a fair trial.
Chapter 5 Evidence base of medical guidelines in evaluating disability
Two countries use medical guidelines: Germany uses six and the Netherlands sixteen. The quality of development of four pairs of guidelines on similar pathologies was studied using the AGREE instrument. All guidelines showed similar AGREE scores with only minor differences. Existing guidelines all meet the AGREE criteria of ‘scope and purpose’ and of ‘clarity and presentation’. The procedures of looking for and incorporating evidence in the guidelines do not meet the AGREE criteria. The evidence is reported to be lacking for precise recommendations. The recommendations with regard to incapacity for work are expressed in non-specific, general terms. AGREE expects the guidelines to be drafted with involvement of all stakeholders and editorially independent. These requirements are only partly met. Client involvement is restricted and controversial. In Germany the guidelines are developed within the German Institute of Social Insurance (DRV). This reduces their independence.

Chapter 6 Disability evaluation interview protocols: comparison
The instrument that is used most often in daily practice is the interview with the claimant, which is susceptible to bias. One way of dealing with this bias is to use protocols. In the Netherlands three protocols have been developed to conduct the interviews in disability evaluation. These protocols were compared according to their similarities and differences through interviews with the authors of the protocols followed by a group discussion and comparison with existing scientific literature. The protocols all prescribe a semi-structured interview that varies in strictness. The topics that are prescribed vary in detail but all match the handicapped role and ICF. The procedural prescriptions aim at establishing a fair trial. The protocols are practice-based and have not been validated. The principles applied correspond with existing scientific findings but are not evidence-based.

Chapter 7 Disability evaluation interview protocols: application
The adherence of Dutch SIPs to interview protocols and their underlying principles was studied using a questionnaire among experienced Dutch SIPs. The results show a professional consensus about several basic assumptions. The principle of argumentative evaluation of disability and the principle of conducting a semi-structured interview are supported by over ninety percent of the respondents. Twelve interview topics are basically always addressed by over eighty percent of the respondents. All respondents used some form of protocol, either published or of their own making. Interestingly, no relationship was found between the SIPs’ training and the use of a particular protocol. This consensus provides firm ground to develop further into principles of disability evaluation interviewing.
3 Methodological remarks

This thesis studies the evaluation of work disability, and this section discusses the strengths and weaknesses of this research. Different articles focus on evaluating work disability and identifying its quality aspects, including medical, legal, and organisational aspects. Such a comprehensive approach is new. The Council of Europe [2] and Mabbet et.al. [3] report on the legal phrasing of criteria for disability but not on the way the criteria are applied. The OECD [4] has studied social insurance in many countries, but has focused on policies of granting benefits and promoting return to work, not on the evaluations. Four studies make comparisons between countries; this is a new approach as well. By using the EUMASS network and other countries, a good variation in practices is reached. Comparing countries prevents a bias from one’s national perspective but may limit the applicability of results in one particular country.

The research presented is mainly qualitative, establishing reliability through triangulation of data. For the questions at hand, and the state of science in the field, this seems appropriate. The findings are, on the whole, fairly consistent and can be related to existing literature. Many findings lend themselves to future quantitative research.

The question of what is being evaluated and how quality is defined and controlled in different countries is studied in legal criteria, in output requirements of the ISI and on the practice level of the SIPs. The legal criteria were found to be comparable and this seems reliable as reference was made to published texts. Earlier research [2,3] yielded more general but similar results. Organisational aspects like the definition of output and quality control were different among countries and this seems reliable, as reference was made to established policies and administrative prescriptions. Organisational aspects can probably be more differentiated as claims may be different. This thesis considers only straightforward first evaluation of work disability. In practice, organisational aspects may, within a country, vary from one region to another, a variation that was not studied. To have a valid picture of the full range of evaluations of work disability, these differences need to be described. Other studies with regard to the organisation of evaluation of work disability were not found and so cannot be compared. Practices were studied through questionnaires and focus groups but not examined in vivo; consequently, the results do not have a proven validity for the day to day work in disability evaluation. The terms that are used are not completely consistent among countries and thus the information may, at times, be flawed. This may be a matter of language but also of culture. This seems likely, for example, with regard to the question of what is evaluated in practice: some respondents distinguished between the health condition of a claimant and his functional capacity and scored both. Others probably did not make this distinction and scored only one of the two. Flaws may also have occurred when looking for procedural guidelines.
Some respondents did not perceive these as guidelines and did not report their existence. It seems unlikely that no prescriptions exist in these countries. A study on medico-legal reasoning (chapter 3) is interesting to do in more countries and based on more, and different, cases. All in all, the precise definition of the object of the evaluations in practice in different countries needs to be confirmed in further research. The impact of treatment and coaching of sick leave for the evaluation of work disability has not received much attention in this thesis. Differences do exist and sick leave history is included in the evaluations (chapter 2) but this aspect needs more specific attention. The different studies were performed over the past six years against a background of developing practice, society and science [5,6,7]. In the Netherlands disability evaluation now is done in a different legal scheme and under a different approach [8]. To the research questions of this thesis the changes probably do not influence the answers. Explicit definitions of quality of work disability evaluation were not found in any country. This surprising result seems a reliable finding as it was explicitly addressed during all visits. Implicit definitions or common understandings probably do exist and need to be explicit in further research. The policies to control quality were studied from the perspective of ISIs. One visible aspect of quality control is the use of guidelines. In this study several guidelines were identified and studied. It is likely that in practice more guidelines exist, albeit informally. It is also possible that local professionals have more mechanisms to ensure quality. This was not studied but it is worth the trouble to do so. The guidelines that were found were analysed from a medical point of view (handicapped role) and a procedural point of view (fair trial). The validity of these viewpoints for guidelines needs to be further confirmed as the material found was limited.
4 Answers to the questions:

This paragraph answers the research questions for the Netherlands and when appropriate from an international perspective.

Question 1: What is to be understood by evaluation of work disability?
At the legal level, as presented in chapter 2, the handicapped role [9] seems most fit to describe what is being evaluated. This is the case in all fifteen countries studied. At the legal level a fair trial is not mentioned but can be taken to be self-evident.

At the organisational level, as presented in chapter 2, this is less clear. Different countries apply different output requirements that focus on aspects of the handicapped role: the medical condition, functional capacity, and the rehabilitation perspective. Clearly, different ISIs put different emphasis on aspects of the handicapped role. The handicapped role remains the core concept on the organisational level, however.

At the professional level, as presented in chapters 3, 5, 6, and 7, the handicapped role is most fit to describe what is being evaluated. In answering the questions in chapter 3, the respondents sometimes seem to equal the health condition with the (in)capacity for work. A requirement of fair trial is apparent in all four the countries studied [chapter 2].

In the situations studied in this thesis, evaluation of work disability is performed in a public arrangement. On legal, organisational, and professional levels, disability is defined differently. These findings are consistent with Lipsky [10] and Veen [11] who found that various levels in bureaucracies had different working definitions of their task. It is interesting that these differences exist and they must be an obstacle to a common policy of quality control.

Is a common answer possible? If we look at the results of the studies that are presented in chapters 2, 3, 4, 6, and 7, one answer is that an evaluation of work disability is to conclude if the claimant meets the criteria of the handicapped role, based on an examination according to the practices that are considered appropriate in the arrangement at hand.

This definition calls for operationalisation of two aspects: the handicapped role and practices that are considered appropriate. The handicapped role refers to:
1. the claimant’s restricted capacity to function in work
2. his damaged state of health as explanation of 1
3. his behaviour to recover and to take up work
4. possibilities to improve his health and functional capacity

Using the handicapped role agrees with existing opinions as expressed in Waddell & Aylward [12] and with the advice of the Health Council of the Netherlands [13]. The operationalisation of the object of disability evaluation by the Health Council [13] is similar to the handicapped role. The concept of role suggests more coherence in the evaluation than the seemingly independent tasks of evaluation as formulated by the Health Council of the Netherlands. The different aspects are not evaluated separately but can
be seen as conclusions that can be drawn from an evaluation of all four aspects. The Health Council of the Netherlands defines disability evaluation as the evaluation of 1) socio-medical history, 2) actual functional capacities, 3) the prognosis, and 4) actual treatment and socio-medical coaching. This is not universally recognized in the Netherlands: current practice is to look at the evaluations as consisting of a main part, functional and earning capacity evaluation, and several more or less arbitrary parts [14]. In other studies, the evaluation of disability is presented as consisting only of evaluating functional capacity [15,16]. The Health Council of the Netherlands identifies evaluating actual functional capacities as the central task of the SIP. Conceptualising functional capacity as a separate characteristic of a claimant suggests an objective evaluation of functional capacity, which approach turns out to be difficult to accomplish [15,16,17]. Using ICF [18] as the main concept of disability in social insurance tends to support this approach. ICF was developed as a classification of consequences for chronic ill health, not as an instrument for disability evaluation in social insurance. In the scientific literature the ICF model is proposed as representing essential elements of disability [19]. In the studies presented here, ICF was not in official use in any legal scheme. ICF fits with what is evaluated but represents only part of it: the time perspective and the fair trial are not addressed. The practices that are considered to be fit in the arrangement at hand refer to a fair trial. This aspect of the evaluation is less developed and more implicit than the handicapped role. Yet this consequence of the legal context in which the evaluations are performed is presumably universally present. The fair trial aspect is addressed in procedural guidelines that were found in four countries and possibly exist to some extent in all countries. The operationalisation of disability evaluation by the Health Council of the Netherlands also indicates aspects of a fair trial formulated as quality conditions. In this study the fair trial seems, at the professional level, to have a more central role: without application of a fair trial an evaluation of work disability is not complete.

From the study on medico-legal reasoning (chapter 3) and in the guidelines from Germany and the Netherlands (chapter 4) and the disability interview protocols in the Netherlands (chapters 6 and 7), we can conclude that a fair trial refers at least to:
1. Independence of the SIP towards claimant and ISI;
2. Ways to determine the existence of a plausible and consistent picture of impairments, disabilities, and handicaps;
3. Ways to determine functional capacity (parting from the claimant’s opinion and verifying and completing this information).

Ad 1): The first requirement seems self-evident but is hardly operationalised. The script model of disability evaluation [20,21] illustrates that ways have to be found to establish a balance of interests. A way to do this on the individual level, as is demonstrated in chapter 6, is to use the introduction
part of the interview. On the group level of SIPs independence is shown by using guidelines for evaluation practice.

Ad 2): The second requirement fits with existing regulations but is well known to be difficult in practice. It rules out the requirement that disease exists only when a medical diagnosis has been proved, which can be particularly problematic in cases of unspecific diseases like chronic pain and fatigue. It is also challenged in the discussion between the doctors on the grounds that this is not a real medical, diagnostic, approach [22].

Ad 3): The third requirement fits existing regulations and opinions [8,23,24]. It shows that an evaluation of work disability is not a potentially objective measurement but rather a conclusion of a legal structure, based on conjecture and refutation. In Dutch literature this is called the argumentative evaluation of disability. This point of departure dictates a certain balance of power in the individual evaluation, the claimant having a substantial position.

To the extent that this answer of the first question is applicable in other countries needs more study. The present findings suggest that it is applicable. And no countries were found in which the answer does not apply. Different countries are in a different phase with regard to explicit definition of evaluation of work disability. Using the answer given here may help to speed up this development.

Question 2: What is quality of evaluation of work disability?

At the legal level, specific quality requirements were not found. Application of the law and proceeding with a fair trial are implicit quality requirements. At the organisational level, quality is organised rather than defined: specification of the output, evaluation by qualified personnel, and file inspection by colleagues are common measures. At the professional level, a definition of quality of evaluation was not found. In several countries guidelines have been produced. The medical and procedural guidelines represent professional agreement about the professional standards in the situations described in the guidelines. Implicitly, the professional definition of quality is therefore “performance according to professional standards.” Direct determination and control of quality requires an operationalisation of quality. How can determination of quality of disability evaluation be realised?

Looking closer at the handicapped role, partly answers this question.

1. The capacity to function in work is a relational matter between a person’s capacities and work demands. Measuring capacities in different situations yields different results [25]. A person’s capacity to function in specific work depends partly on demands and circumstances of that particular work. It is difficult to imagine a measurement that includes all possible capacities under all possible circumstances. So called Functional Capacity Evaluations do give valuable information but cannot pretend to answer all questions [16]. Besides this technical problem, there is also a moral aspect: how much pain, sacrifice, and suffering

157
are reasonable to ask from a disabled person to endure to be (partly) economically independent?

2 The aspect of the damaged health condition requires this condition to be explanatory of insufficient capacity to work to meet the criteria for a benefit. Recent endeavours to support this aspect of the evaluations with medical guidelines invariably lead to the conclusion that a specific disease is likely to bring about certain restrictions in functioning, but mono-causal relations are the exception, not the rule. Many claimants are found to suffer from a combination of illnesses that have a combined influence on the claimant’s capacity to function [13]. One Dutch procedural guideline prescribes that not the diagnosis but the consistent constellation of plausible manifestations of disease make up the requirement of a damaged health. This suggests much room for individual interpretation by SIPs, which is indeed found [15,26,27,28].

3 The aspect of behaviour towards recovery and resumption of work is partly empirical (what actions and interventions lead to recovery and resumption of work? What actions have been undertaken by this claimant?). Few cures or return to work programs are proven effective for the population that is evaluated for long-term incapacity for work [29]. Therefore, the question is: has he or she done and is he or she doing what can be reasonably expected, considering his or her situation and knowledge about effectiveness of certain interventions? So this evaluation is normative too.

4 The aspect about future possibilities is partly speculation about the answers of the earlier questions. The degree of certainty depends partly on these earlier questions supplemented by uncertainty about the future and by sound epidemiological evidence, if present.

It seems clear from these four aspects that the quality of the individual evaluations cannot simply be determined based on the quality of their output on the basis of measurable criteria that are independent of client and SIP. All four aspects are known to harbour dilemmas and no clear-cut answers are possible [14]. This explains why quality control in social insurance appears to be mainly indirect, as found by Meershoek et. al. [17]. This may seem unsatisfactory, but in the area of professional judgments is not at all exceptional [20]. Professional discretion is demanded in all situations where relevant grounds for the decision are determined by the situation itself, more than by rules or knowledge. Judges call this, in their verdicts, “the facts and circumstances of this particular case.” This discretion can be handled in a sufficient fashion by ensuring the competence of the experts that perform the evaluations and by ensuring that they agree. This is partly common practice in social insurance already: we found that specific education of the SIPs is usual in many countries as is the inspection of doctors’ reports by other doctors. A more direct way of enhancing SIP’s agreement is to do evaluations in committees. This was found in Belgium and in former Eastern-European countries. This is an application of the
mechanism of Spearman and Brown that explains why a group opinion of experts is more likely to be correct than the experts’ individual opinions [20].

Following this line of thought, quality of evaluations is found in the agreement of experts [30]. This is an intersubjective criterion instead of an external objective criterion. Quality of evaluation of work disability is thus to be defined as experts’ agreement on (aspects of) these evaluations. How can this quality be operationalised? And how free are the experts to follow their own opinions?

Operationalisation of the quality of evaluation of work disability can be found in how the evaluations are operationalised: to conclude if the claimant meets the criteria of the handicapped role based on an examination according to the practices that are considered fit in the arrangement at hand. Moving from the above answers, the operationalisation of quality includes:

1. SIPs agreement on the conclusions in terms of the handicapped role
2. SIPs agreement on the information gathering according to principles of fair trial

The study on Medico-legal Reasoning supports the idea that both principles guide SIPs in different countries. This conclusion is supported by the Guideline study and the studies on Disability Interview Protocols in chapters 4 and 6 respectively. As indicated earlier, the handicapped role seems to be broadly, although not universally, accepted. This is yet to be established of the fair trial.

One question is how far expert agreement is possible and necessary. SIPs can agree on several levels:

− About principles and available knowledge in general. This agreement is found in education and by drafting evidence based guidelines. In this thesis, education was not studied. The procedural guidelines that describe the principles in the Netherlands are found to be relatively old and less coherent than the German and Swiss guidelines. The relevant knowledge as presented in the medical guidelines is limited and of general value.

− About application of principles and knowledge in individual cases. This agreement is found in sharing evaluations and case reports, and is done in training situations and peer tutorial sessions as used in the Netherlands. It is not known how effective these are in establishing agreement among SIPs. Their effect will be restricted to small groups and leave unanswered the question of the collective agreement.

To what extent this answer of the second question is applicable internationally also needs more study. The available evidence suggests that it is applicable. Countries like Germany and Switzerland have well developed procedural guidelines and the Netherlands and Germany have well developed medical guidelines. Other countries might fruitfully start from these guidelines to develop their own.
Question 3: How can quality of work disability evaluation be controlled?

Quality control consists of feedback loops from output or outcome of the evaluations of work disability to input or throughput [31]. At the legal level this means adapting the criteria according to outcome such as the disability statistics. This is done infrequently, although it does occur and then with great impact [6,7]. Quality control is usual at the organisational and the professional level. Feedback at the organisational level exists in many countries, notably in using the file inspection for feedback. The results are not published in any country studied. At the professional level, qualification is controlled mainly by continuous education.

From the previous reasoning it follows that an individual SIP can make an evaluation of good quality if he or she uses the expert consensus in every aspect of it.

Quality is primarily a matter of the professional group of SIPs. This description gives, however, not the complete image of quality control. The profession of Social Insurance Medicine is not a free market offer but is performed between parties described in the script model of disability evaluation. Consequently, SIPs need to decide about the quality of their work in a manner that convinces their environment: the contractor Institute for Employee Benefit Schemes (UWV), the claimant, and society as a whole. Claimant, UWV, and SIP are, according to the script of the evaluations, the first parties to contribute to quality control, each in their own fashion, individually and as a group. What do these parties do in the Netherlands?

The influence of the individual client is limited to the individual evaluation, which is not insignificant. The client is the object of the evaluation but also the subject. The better the claimant is instructed about the rules of the evaluation and the more aware he is of his situation, the better he will be able to claim a proper evaluation and influence the evaluation of his own disability.

Clients’ organisations in the Netherlands are becoming more active in supporting their members and in participating in guidelines for SIPs. In the study about EBM guidelines [chapter 5] clients’ organisations’ involvement in medical guidelines was found to be in development but still restricted and controversial.

The individual SIP can realise quality by gathering information according to professional rules of evaluation and by linking this information to professionally accepted grounds into arguments that conclude about the client’s handicapped role. Like many professionals, the SIP stands between managerial instructions that often are considered mechanic [32,33] and full autonomy to decide, which is considered to lend itself to the use of strictly personal convictions [17, 30]. A balanced position seems to be to rely on professional standards that satisfy the administration and that leave room for tailor-made evaluations [34,35]. This requires that the SIP participates in permanent professional education and that he or she has the attitude and the situation to work according to professional standards.
For this, the SIP needs to know the existing professional consensus. The guidelines describe the consensus to some extent. Medical guidelines define criteria and grounds with regard to specific diseases. However, the medical guidelines are not strongly evidence-based and give only general answers to the individual questions that the SIPs face. The Dutch procedural guidelines provide some answers with regard to rules, but these answers are relatively old and lack the coherence of the German and Swiss guidelines. About the social norms, no explicit sources are available. SIPs can be more explicit by identifying the grounds they base their arguments on. These grounds can be characterised in different ways; one is according to their source: knowledge (preferably scientifically proven), rules (preferably legally based), and norms (preferably socially shared). See Figure 1:

![Diagram of argumentation and sources of grounds](image)

**Figure 1:** Argumentation and sources of grounds.

At the group level, it is new that the Dutch SIPs define their professional standard in guidelines. This traditionally is the domain of the UWV as is the case in other countries [chapter 2]. In the study on organisation of disability evaluation, it was found that much room existed for active control of quality, especially in terms of professional standards. Since 2003, sixteen medical guidelines have been drafted by the Health Council of the Netherlands and by the Dutch society of Social Insurance Physicians (NVVG). These guidelines support the evaluations but any accounting of how these guidelines are applied in daily practice is missing, except for the case report inspection by UWV staff. The results of these inspections are not published. Even more recent is the advice of the Health Council of the Netherlands that the professional group develops medical case law ("mediprudentie"), a collection of well documented and professionally commented case descriptions. These case descriptions are to demonstrate the professional consensus on a case by case base. Recently this approach was tested [36] and it is being implemented now. UWV, the contractor of the evaluations, can facilitate and check the application of professional standards of the SIPs, which the SIPs see as
Chapter 8

...problematic [14, 15, 37]. The study on organisation of disability evaluation [chapter 2] dates from 2003 and is partly outdated for Dutch practice in 2009. Since 2003 guidelines have been implemented at UWV and peer tutorial sessions have been extended to draft illustrative cases (precursors of *mediprudentie*, see above). A consistent policy of monitoring and controlling quality has, however, not been officially published. Internationally this answer again seems applicable. However, the balance between ISI and the professional group appears to differ among countries. Applying this approach calls for tailor made solutions.

The parties most concerned with quality control of evaluation of work disability are ISI, SIPs and their organisations, and claimants and their organisations. Their mutual influencing can be summarised as in Figure 2:

![Diagram](image)

**Figure 2:** Parties in quality control of evaluation of work disability.
5 Recommendations:

The conclusions suggest certain actions that can be taken to improve quality in evaluating work disability in the Netherlands. Quality needs to be realised at the practice level, so the recommendations will focus on this level first. Next, we will consider quality control at management and other levels. When applicable, recommendations for other countries will be provided.

Practice level
The main challenge for quality improvement in the Netherlands is for practicing SIPs. They are unsure about the quality of their work and have to resolve these uncertainties by using practices and grounds that represent expert consensus and scientific evidence. Their job is to perform the disability evaluations according to professional guidelines and standards that are constantly evolving.

Making the grounds for disability evaluation explicit in the reports of individual evaluations will enhance the quality of these medical reports, as the grounds used can be evaluated. Developing consensus about these grounds would be a crucial step forward, as consensus will improve professional practice and contribute to the transparency and legitimacy of the disability pension schemes that are open to constant criticism about their capacity to select the right people for a disability benefit.

In terms of the practice of evaluating incapacity for work, both medical and procedural guidelines are important instruments to improve and control quality. It can be expected that professionals’ guidelines will endorse the legal security of claimants as well. For these reasons, the development, testing, and implementation of guidelines deserve support.

Consensus about the interview routine can be used to further develop professional consensus on SIPs’ attitude and on the structure of the interviews. Without striving for a detailed and universally applicable protocol, relevant topics can be selected that are well circumscribed and based on consensus and evidence about what constitutes disability.

If such measures are developed, implemented, and controlled, the quality of the evaluations will improve in terms of transparency and reproducibility. The evaluations will become more comparable and more accessible to scientific research on behaviour of both the SIP and the claimant. The proposed innovations will enable claimants to better prepare themselves for the evaluations, which will make their position more equal to that of the SIP, in favour of the satisfaction of the claimants.

The professional group
It is their professional society (NVVG) that faces the challenge of uniting the SIPs and of developing professional practice guidelines and the medical case law. An essential first step is that the professional group agrees that the evaluation of work disability is not limited to the evaluation of
functional capacity. Evaluation of work disability needs to be based on operationalisations of the handicapped role and on the principles of a fair trial. Another essential step is to define when an evaluation can be said to meet quality standards, including input, process, and output aspects of the evaluation. The development of medical case law will result in many examples.

Over the past years many initiatives in professional development have been implemented. Further action is necessary to consolidate the scientific tradition in social insurance medicine as a stable infrastructure is not established yet. Three developments are particularly promising for the near future: harmonising the procedural guidelines, revising medical guidelines according to AGREE criteria, and developing medical case law.

Procedural guidelines have been developed in the past but in a piecemeal fashion. A guideline on the evaluation of clients’ behaviour towards recovery and resumption of work is underway. Procedural guidelines would profit from more cohesion to improve their applicability and transparency. An option is to produce one comprehensive new procedural guideline. The guidelines need to be updated anyway.

Medical guidelines will need updating in the near future as well. Applying AGREE criteria more rigorously will help in their development and will indicate the lack of evidence and promote targeted research. The basic conditions for optimal guideline development can be met already.

Medical case law (mediprudentie) may be one of the core instruments in the coming years as it is a promising way to demonstrate the professional standards on a case by case basis. Medical case law indicates how principles and knowledge can be applied in individual cases.

Consensus about the essentials of a disability interview will substantially contribute to valid evaluations and good documentation. This consensus can provide a starting point to develop a guideline that can be implemented in practice and evaluated. For implementation, more than a single training will be needed. Some form of supervision and support will be necessary to maintain and improve the application. The everyday practice of a growing diversity of claimants calls for research and development in communication for SIPs. Companies, medical specialists, and occupational physicians will need to be trained for their new roles as well.

Internationally, it seems possible and worthwhile to harmonise the efforts of the professionals. As seen in the study on medico-legal reasoning, the handicapped role and the fair trial turns out to be pivotal for the evaluation of work disability per se. Possibly this applies further than work disability, to participation problems in general. Studies on medico-legal reasoning, but also the making of case descriptions, should be tried internationally as well, perhaps in this order so as to find any incompatibilities in professional grounds. The EUMASS network can facilitate this for European countries.


Institute of Social Insurance (ISI)
All actions where professionals take initiatives, as mentioned above, need to be performed in collaboration with the ISI and with claimants’ representatives. The Dutch ISI, UWV, is responsible for both the administrative and medical quality of the deployment of the law, notably the ‘Work and Income, according to Labour Capacity Act’ (WIA). Overlap in responsibilities of ISI and the professional group easily leads to tensions between UWV, SIPs, and professional organisations. Professional quality control will benefit from a supportive and challenging attitude of UWV toward the professionals. One innovation can be that the SIPs report the functional capacity of claimants and other aspects of the handicapped role. The application of a fair trial needs to be elaborated and implemented. Considerations of efficiency are relevant for both ISI and professionals and need to be weighed against professional standards. A transparent trade-off will help make tensions clear and manageable.

Claimants
Claimants as individuals are not easily in a position to demand and develop quality evaluations as they are object in the evaluations. But claimants are in a position to identify problems in daily practice, and their representatives can contribute to the development of professional guidelines. Recent experiences in the Netherlands with guideline development suggest that their input can be very valuable. This deserves support as the representatives of the clients can be the SIPs’ allies in the quest for quality. In other countries too, claimants can be involved in developing guidelines, if it is possible to find clients’ representatives who can share the basic tenets of disability evaluation.

Supervisor
The Ministry of Social Affairs and Employment monitors the performance of ISI in applying the legislation on work disability. Consequently, they are in a position to demand, facilitate, and verify that SIPs, ISI, and claimant organisations cooperate to promote the quality of evaluations. The model of the script of evaluation suggests that a countervailing power is needed. This prevents a one-sided demand on efficiency by the ISI opposing professionals’ guidelines that would be detrimental to the quality and effectiveness of the disability evaluation. The Ministry, representing the public interest, is in a position to provide checks and balances, and has the opportunity to involve the Health Department to warrant cooperation between the worlds of health and labour [38].

Future Research
All recommendations provided will need to be supported by scientific evidence and proper scientific evaluation. Studies are needed to verify if guidelines are helpful to establish a good quality evaluation. Various questions need to be answered: How can medical guidelines be developed that are more
firmly evidence-based? How can implementation be stimulated in practice? What is the effect of procedural guidelines in the effort to improve and control quality? Does medical case law enhance consensus between SIPs about the correct grounds to use and about proper information gathering and use of evidence? Answering these questions is a matter of independent research. Practical instruments to establish agreement of experts need to be developed, which can serve as gold standards to test the effectiveness of the suggestions made.

The guidelines would profit from considering the handicapped role as the object of evaluation and from more specific recommendations. For this, and to support individual evaluations, epidemiological research is needed on the functional capacity that people with disabilities have and how they deal with these capacities in their daily lives. With regard to less common diseases, the evidence on the prevalence and incidence of impairments will always be limited. With the introduction of new therapies too, much will be unknown about the impact on people’s work capacity. An alternate, more generic approach for chronic conditions is to use existing evidence about the generic effects of chronic medical conditions and diseases on working capacity and participation opportunities. Epidemiological studies departing from this point of view can be more effective than starting new studies for every disease separately.

The difficulties between professionals and their employers are by no means specific for SIPs. Both employers and professionals seek an optimum way of achieving excellence within boundaries of efficiency, which is an interesting field of research that can help solve these problems. SIPs provide good case examples as they form large populations that are relatively homogeneous in their tasks and are, on that task, comparable among countries.

So far, in discussing the quality of the evaluations, emphasis has been on structure, process, and output aspects such as education of SIPs, quality of guidelines, good application of the guidelines, and good communication during the evaluation following a well developed interview protocol. In the end, evaluations result in people with or without a benefit and with or without work or accommodations. It would be interesting to monitor the fate of claimants on the labour market after their evaluation. This would provide important feedback to the SIPs about the opportunities people with disabilities in reality have on the labour market.

Most of these recommendations for research are relevant in other countries too. It is interesting to note that in disability evaluation so many similarities are found among countries. That opens fascinating perspectives for the international development of the social insurance medicine discipline, in education and in research. The adoption of guidelines and case law elsewhere would show the effect of different conditions in different countries. The same goes for epidemiologic study of populations of claimants and people who receive benefits. Another interesting field is the question of supposed incompatibility between evaluating clients and trying to help the evaluated clients. This is done differently in different countries.
and international comparison would help to find opportunities that are not found yet. Finally, professionals working in administrative organisations are found in all countries studied in this thesis. The aspect of quality control in such circumstances also deserves international comparison.

**Future quality**

Evaluation of work disability is a most relevant activity from both an individual and a social perspective. It is underrated if its societal impact is compared to the efforts to study and improve daily practice. Many causes can be identified for this lack of scientific interest and the lack of profiling and professional pride of the social insurance physicians are among them. In the Netherlands over the past few years, quite a few initiatives have been taken to scientifically endorse the evaluation of work disability. The findings in this thesis show the need to continue. If the recommendations are brought into practice and are effective, professionals can enhance the quality of their work. The legitimacy of the system and the sense of an equitable treatment with the claimants will be supported by this. That is an opportunity not to lose.
Chapter 8
References

15. Bont AA, Berendsen L, Boonk MPA, Brink JC van den. At the doctor’s office. [in Dutch] Zoetermeer CTSV 2000
Chapter 8

23 Law on work and income according to working capacity (WIA) [in Dutch] http://lexi.us.nl/wet-werk-en-inkomen-naar-arbeidsvermogen visited 4-9-2009
31 Imaki, M. Kaizen, the key to Japan’s competitive success New York Random House 1993
32 Brink, G van den, Jansen T, Pessers D (eds) Professional pain. Why the Netherlands don’t work properly. [in Dutch] Amsterdam Uitgeverij Boom 2005
33 Weggeman, M. Managing professionals: just don’t! [in Dutch] Schieded Scriptum 2007
34 Freidson, E. Professionalism, the third logic. Cambridge Cambridge Polity Press 2001
36 Committee for preparing medical case law. Medical case law at the start. [in Dutch] Utrecht NVVG 2009
37 Choy, YW. Report about the consultation of social insurance physicians and labour experts by the department of social- medical affairs. [in Dutch, not officially published] Amsterdam UWV 2005
38 Blue print of Participation in Medical Guidelines [in Dutch] Utrecht VGI 2007