# The Shifting Sands of Evidence

# A socio-legal enquiry into the development of medical guidelines

**Friso Johannes Jansen**

**St. Anne's College**

**DPhil in Socio-Legal Studies**

Trinity Term 2017

# ACKNOWLEDGEMENTS

Firstly, I would like to thank all the participants who willingly took the time to answer all my various questions. Without your cooperation this thesis could never have been completed. My supervisors Dr. Marina Kurkchiyan and Dr. Christopher Decker were always generous with their time and constructive in their comments. Their help was instrumental in refining the approach of the study and the theoretical approaches adopted. Their patient reading and re-reading of the manuscript made a vital difference in strengthening the final work. At the Centre for Socio-Legal studies many were generous with their advice, and many more with their friendship. I would like to particularly thank Naomi whose warmth and friendship provided a constant source of support, and laughter, lots of laughter. I would like to thank Petra, Rogier, Inginio, Paulo, and Christina for their encouragement and friendship. Fellow students that made my DPhil a collective endeavour that I would like to show my immense gratitude to are, in no particular order: Po-Hsiang, Paul, Owain, Matilde, Christina, and Felix, as well as all the others that I was fortunate to meet in Oxford. I would like to thank Doede Joustra for his help with accessing some of the Dutch sources in this thesis. Sarah, my vrouw, has been instrumental in pushing me to pursue a DPhil and I would be lost without her. Her critical mind and probing questions have improved this thesis time and time again. I hope to return the favour and give some of the love back she has given me over the years. This thesis has been supported by financial contributions from St. Anne's College Oxford, the Centre for Socio-Legal Studies Oxford, het Vreedefonds, de Fundatie van de Vrijvrouwe van Renswoude and Stichting dr. Hendrik Muller's Vaderlandsch Fonds.

# ABSTRACT

## The Shifting Sands of Evidence

## A socio-legal enquiry into the development of medical guidelines

Friso Johannes Jansen, St. Anne's College

DPhil in Socio-Legal Studies, Trinity Term 2017

Medical guidelines on the same medical condition differ between England and the Netherlands. These guidelines are referred to as evidence-based because they are supposedly based on a systematic searching for and appraisal of medical studies to drive recommendations for appropriate care for specific clinical circumstances. This comparative study interrogates what causes these differences and similarities between guidelines and tries to uncover the mechanisms behind the development of medical practice guidelines. Four case studies, on lower back pain and on type 2 diabetes in both countries, are used to provide a detailed empirical account of the development of medical guidelines. Interviews with guideline developers are combined with a detailed analysis of available guideline documents. The overarching finding of this thesis is that medical evidence plays a more limited and nuanced role in guideline construction than might be expected and that guidelines are manifestations of professional (self-) regulation. Importantly, the research also finds that institutions shape guidelines in a multitude of ways. This study has endeavoured to add to a more nuanced understanding of evidence within the literature: conceptualising evidence as part of a process of a social and institutional construction. This construction is used within a collaborative and communicative process aimed at creating 'objective facts'. Contrary to existing scholarship, this thesis argues that evidence merely informs the understanding of members of guideline groups while a range of economic, cultural, institutional, and political factors, that together form cognitive frames, provide the driving force behind the development of guidelines. Institutional factors have shown to be essential elements in guideline development, influencing all aspects of development through institutional cultures of practice. This study concludes that calling guidelines evidence-based is an important rhetorical instrument, which helps to conceal and legitimize some of the normative choices that are inherent in guideline making.

Word count: 86,800

# Table of Contents

ACKNOWLEDGEMENTS ..... iii
ABSTRACT ..... iv
Table of Abbreviations ..... viii
1 Introduction to Evidence-based Medical Guidelines ..... 1
1.1 Introduction ..... 1
1.2 Research Question and Problem Statement ..... 5
1.3 Selection of Country Case-Studies ..... 6
1.4 Selection of Medical Cases ..... 8
1.5 The Methodological Choices Explained ..... 11
1.6 Structure of the Thesis ..... 18
2 Understanding Guidelines in their Academic Context ..... 22
2.1 Introduction ..... 22
2.2 Evidence ..... 24
2.3 Regulation ..... 39
2.4 Institutional Literature ..... 59
2.5 Factors Shaping Group Decision-Making ..... 72
2.6 Conclusion: Relevance of Literature to this Thesis ..... 90
3 Guidelines in the Netherlands and England ..... 96
3.1 Historical Perspectives: From Consensus to Evidence-Based ..... 96
3.2 Professional Associations and Guideline Actors ..... 104
3.3 The Regulatory Environment of Guidelines ..... 116
3.4 Conclusion ..... 128
4 Lower Back Pain: Guidelines in England and the Netherlands ..... 132
4.1 The English Guideline on Lower Back Pain ..... 132
4.2 Evidence-base for English Lower Back Pain Treatments ..... 139
4.3 Achieving Consensus: Internal Workings of Guideline Groups ..... 150
4.4 External Responses: Rejection of the Guideline ..... 156
4.5 Conflicts of Interest: Limited Relevance ..... 162
4.6 Guideline Drafting in Institutions: Success Within Constraints ..... 164

4.7 Costs: A Complicating Factor for the English Guidelines ..... 170
4.8 The Dutch Guideline on Lower Back Pain ..... 172
4.9 Evidence-base for Dutch Lower Back Pain Treatments ..... 175
4.10 Internal Consensus Building: How Results were Reached ..... 184
4.11 External Responses to the Guideline: A Deafening Silence ..... 188
4.12 Conflicts of Interest: A Limited Role ..... 190
4.13 Institutional Factors: Pressures to Conform ..... 192
4.14 Comparison of Guidelines: Difference or Similarity? ..... 200
5 Diabetes Type II: Guidelines in England and the Netherlands ..... 212
5.1 The English Guideline on Diabetes Type II ..... 214
5.2 Evidence-base for English Treatments for Diabetes Type II ..... 214
5.3 Achieving Consensus: Internal Workings of the English Guideline Group ..... 234
5.4 External Responses: Confrontations over Clinical Choices ..... 243
5.5 Conflicts of Interest: Potential Influences ..... 245
5.6 Institutional Factors impacting on Guideline Drafting ..... 247
5.7 Costs: Cost-Effectiveness Modelling and Clinical Choices ..... 249
5.8 The Dutch Guideline on Diabetes Type II ..... 254
5.9 Evidence-base for Dutch Treatments for Diabetes Type II ..... 254
5.10 Achieving Consensus: Internal Workings of the Dutch Guideline Group ..... 265
5.11 External Responses: Acceptance and Resistance ..... 271
5.12 Conflicts of Interest: An Anti-Industry Bias? ..... 273
5.13 Institutional Factors: Too Much Independence? ..... 279
5.14 Costs: Inclusion by Omission ..... 281
5.15 Comparative Analysis: Differences Dominate ..... 283
6 Guidelines in a Comparative Sense ..... 294
6.1 Evidence: Social and Institutional Construction of Objectivity ..... 294
6.2 Consensus: Different Mechanisms throughout the Process ..... 302
6.3 Institutions: Cultures of Institutional Practice ..... 315
6.4 External Factors: Limited Influence ..... 325
6.5 Conflicts of Interest: The Difficulties of Attribution ..... 331
6.6 Cost-effectiveness: Considerable Tensions between Ideals and Practice ..... 334

7 Conclusion ..... 347
7.1 Evidence ..... 347
7.2 Institutions ..... 355
7.3 Group Decision-making ..... 358
7.4 Costs ..... 362
7.5 External Factors ..... 364
7.6 Regulation ..... 365
7.7 Concluding Remarks ..... 368
Bibliography ..... 372

## Table of Abbreviations

| | |
|---|---|
| ACCORD | Action to Control Cardiovascular Risk in Diabetes |
| ACM | Authority for Consumers and Markets |
| ADVANCE | Action in Diabetes and Vascular disease |
| AGREE (II) | Advancing guideline development, reporting and evaluation in healthcare |
| BEAM | United Kingdom Back Pain Exercise and Manipulation Randomised Trial |
| CBO | National Association for Quality Assurance in Hospitals |
| CCG | Clinical Commission Group |
| EBM | Evidence-Based Medicine |
| EBRO | Evidence-based guideline development |
| EU | European Union |
| GDG | Guideline Development Group |
| G-I-N | Guidelines-International-Network |
| GP | General Practitioner |
| GRADE | Grading of Recommendations Assessment**,** Development and Evaluation |
| HSCA 2012 | Health and Social Care Act 2012 |
| HTA | Health Technology Assessment |

| | |
|---|---|
| ICER | Incremental Cost-effectiveness Ratio |
| IGZ | Dutch Health Inspectorate |
| ISO | International Organization for Standardization |
| ISDA | International Swaps and Derivatives Association |
| KNMG | Royal Dutch Federation of Medicine |
| LHV | The National Association of General Practitioners |
| NCC | National Collaborating Centre |
| NHG | The Dutch Royal College of General Practitioners |
| NHG-Standaarden | The Dutch Royal College of General Practitioners Guidelines |
| NICE | National Institute for Care and Health Excellence |
| NZA | Dutch Healthcare Authority |
| OMS | Order of Medical Specialists |
| RCT | Randomised Control Trail |
| QALY | Quality-Adjusted Life-Year |
| UKPDS | United Kingdom Prospective Diabetes Study |
| VADT | Veteran Affairs Diabetes Trial |
| Zorginstituut | The National Health Care Institute |

## Table of Statutes

Health and Social Care Act 2012 ..... 102, 116, 122-125

Wet op de geneeskundige behandelovereenkomst 1994 ..... 44,117,194

## Table of Figures and Tables

Figure 1. The Evidence Hierarchy based on Knaapen ..... 28

Table 1: Characteristics of the medical fields of the chosen guidelines ..... 9

Table 2: Costs and QALYs of Diabetes Drugs ..... 250

# 1 INTRODUCTION TO EVIDENCE-BASED MEDICAL GUIDELINES

## 1.1 Introduction

This thesis provides a socio-legal enquiry of the development of medical practice guidelines. It develops a comparative qualitative analysis of guideline documents as examples of professional self-regulation in healthcare. In doing so, this thesis aims to provide insight into the 'black box' of guideline making, and aims to elucidate the role of medical experts in using scientific knowledge to draft professional self-regulation.

We would all like to receive the best medical care possible. If we go and see our general practitioner (GP) we expect to receive the same standard of care regardless of which part of the country we live in and which practice/clinic we visit. This, however, is not always the case. Widespread practice variation exists: treatment for the same disease might differ from one doctor to the next. To address this issue, and to increase the accountability of medical professionals, consensus-based medical practice guidelines were developed in the latter part of the seventies in the USA, and quickly spread to the Netherlands and England.[1] The aim of these guidelines was to reduce practice variation by achieving consensus among a large group of medical professionals on the recommended treatment for medical

[1] JJE van Everdingen, *Consensusontwikkeling in de geneeskunde* (Bohn, Scheltema en Holkema 1988).

conditions. As a result, if guidelines would be widely followed, treatments would not differ needlessly.

However, in the 1990s the fact that these guidelines were developed on the basis of consensus among experts was questioned and there was a call to base guidelines on a systematic appraisal of all available scientific evidence: one that would provide universal knowledge to base medical decisions on. This "evidence-based medicine"[2] movement[3] had a profound effect on the medical world, and is listed by some as one of the most important developments in modern medicine.[4] As a result of this movement, all medical practice guidelines are now described as evidence-based: it is these guidelines that this thesis will consider. In order to develop a socio-legal consideration of evidence and guidelines the Dutch and English healthcare systems have been chosen as case studies as they provide clear comparative examples of this topic.

---

2 An approach to medicine that promotes the use of clinical epidemiology and stresses the importance of Randomised Control Trials to inform treatment decisions in medicine. Among many: David L Sackett and others, 'Evidence Based Medicine: What It Is and What It Isn't' (1996) 312 BMJ 71; Frank Davidoff and others, 'Evidence Based Medicine' (1995) 310 BMJ 1085; David Armstrong, 'Clinical Autonomy, Individual and Collective: The Problem of Changing Doctors' Behaviour' (2002) 55 Social Science & Medicine 1771; GH Guyatt, 'Evidence-Based Medicine' [1991] ACP J Club A; Evidence-Based Medicine Working Group, 'Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine' (1992) 268 JAMA 2420; Alvan R Feinstein MD and Ralph I Horwitz MD, 'Problems in the "Evidence" of "Evidence-Based Medicine"' (1997) 103 The American Journal of Medicine 529.

3 Catherine Pope, 'Resisting Evidence: The Study of Evidence-Based Medicine as a Contemporary Social Movement' (2003) 7 Health: 267.

4 Eric Mykhalovskiy and Lorna Weir, 'The Problem of Evidence-Based Medicine: Directions for Social Science' (2004) 59 Social Science & Medicine 1059.

The motivation for this study lies in the observation that when comparing guidelines across countries, the same diseases are often treated in different ways without an obvious explanation.[5] This is something that requires academic study as there is no reason to suspect that the population of England and the Netherlands is so different that this could explain these differences. As evidence-based medicine endorses the systematic use of the best available evidence, and both countries have access to the same (English-language) medical studies, it might be expected that the guidelines would reflect these similarities. It follows that the question is, how despite access to the same evidence do differences occur? And conversely how is it that some elements of guidelines are similar despite guidelines referring to different evidence?

While examining similarities and differences in guidelines is an appropriate starting point, this study moves beyond this to uncover the mechanisms behind the development of guidelines. In this way the study uses medical guidelines as an example of a process of policy and regulatory development and analyses them using a socio-legal approach. While guidelines are extensively debated within the medical literature, a study that approaches them from a socio-legal perspective has, to my knowledge, not been undertaken before. Understanding guideline development in a socio-legal sense can add further to ongoing

---

5 Bart W Koes and others, 'Clinical Guidelines for the Management of Low Back Pain in Primary Care: An International Comparison' (2001) 26 Spine 2504; Jako S Burgers and others, 'Inside Guidelines Comparative Analysis of Recommendations and Evidence in Diabetes Guidelines from 13 Countries' (2002) 25 Diabetes Care 1933.

debates about the role of law in society, and the role of experts in promulgating law and other normative orderings. It can also add to debates in medicine about the value of evidence-based medicine as a method. Uncovering the process of the development of guidelines is not only relevant for the field of medicine and socio-legal studies, but links to wider debates about self-regulation and policy. Opening up the 'black-box' of guideline development may help uncover broader lessons about the use and construction of scientific knowledge by experts.

This thesis will build on earlier work which studied the process of guideline formation using ethnographic methods, as well as empirical studies of the group process of guideline production using interviews, quasi-experiments, or surveys.[6] Although these earlier studies provide an important understanding of some of the elements of guideline formation and negotiation, research on the topic has so far been relatively piecemeal, with various aspects considered in isolation. This study draws together case study findings and

[6] Tim Hopthrow, Gene Feder and Susan Michie, 'The Role of Group Decision Making Processes in the Creation of Clinical Guidelines' (2011) 23 International Review of Psychiatry 358; Claudia Pagliari and Jeremy Grimshaw, 'Impact of Group Structure and Process on Multidisciplinary Evidence-Based Guideline Development: An Observational Study' (2002) 8 Journal of Evaluation in Clinical Practice 145; Rosalind Raine and others, 'An Experimental Study of Determinants of Group Judgments in Clinical Guideline Development' 364 The Lancet 429; Harry Hemingway and others, 'Appropriateness Criteria for Coronary Angiography in Angina: Reliability and Validity' (2008) 149 Annals of Internal Medicine 221; Lou Atkins and others, 'The Process of Developing Evidence-Based Guidance in Medicine and Public Health: A Qualitative Study of Views from the Inside' (2013) 8 Implementation Science 101.

tries to explore all relevant factors in as much detail as possible to provide transferable lessons for guideline development.

## 1.2 Research Question and Problem Statement

To understand the mechanisms that drive guideline development this study provides a detailed qualitative analysis of the drafting process of two sets of guidelines in the Netherlands and England. Driven by socio-legal inquiries about how people make and experience the law, this study is particularly interested in identifying the various contextual mechanisms that operate within these processes and how these mechanisms might explain any differences as well as similarities between medical practice guidelines. To this end a qualitative multi-country study will be conducted that provides two case studies on lower back pain guidelines in the Netherlands and England, as well as two case studies on diabetes type II guidelines in the same countries.

The overarching research question of this study is: What factors and mechanisms explain the differences between the development of the 'evidence-based' guidelines? It uses the lower back pain and diabetes type II guidelines in the Netherlands and England as case studies.

In exploring this question, the study addresses a number of other sub-questions; these include: What role does evidence play in creating differences and similarities in guidelines, and what types of 'evidence' is used? How do institutional framings impact on

guideline development? What is the role of group decision-making in the development of guidelines? What is the role of evidence on cost and cost-effectiveness in the formation of guidelines? How do external factors determine choices made in guideline development? How do these factors and mechanisms impact on guidelines as a form of professional self-regulation? The following sections provide the justification for the choice of case studies and details the methodology used in this study.

## 1.3 Selection of Country Case-Studies

The Netherlands and England have been selected as country case studies as are examples of countries that share similar characteristics but nonetheless have differing guidelines. By comparing two similar countries, we can establish a nuanced understanding of guidelines in the context of a specific healthcare system and mandated by a specific guideline development organisation. This study combines a detailed textual analysis of the relevant guideline documents, and where possible also compares earlier versions or later updates of the guideline, with a detailed look at the negotiation process. This design also allows us to assess the effects of the institutional arrangements, as well as the impact of the structure of the healthcare system on guidelines.

Both countries subscribe to the same ideas of evidence-based healthcare and refer to the same sources for evidence. This can be seen in the use of the same medical databases for searching medical evidence and the restriction to English (or Dutch) language

publications and not French, German or Spanish publications. Policy makers participate in international collaborations on healthcare innovations and regulation: examples of this are the G-I-N network that collaborates on medical guidelines and medical collaborations between hospitals in the Netherlands and England.[7] Academic hospitals in the Netherlands, especially those attached to universities, have explicit relationships with hospitals in England and collaborate on several research projects. In addition, Dutch medical personnel often visit England (admittedly much less so the other way around) and practice as doctors there; when they return they bring knowledge of the English healthcare system back with them. Added to these specific factors, the two countries have similar general characteristics: both are considered among the best healthcare systems in the world, depending on which publication you read.[8] Both are in Western-European, rich, and well-developed countries with long traditions in trying to improve the quality of healthcare. All these factors make it interesting to study whether similarities or differences can be found in medical practice guidelines.

---

7 Among many: https://www.elixir-europe.org/about-us/collaborations, eyehospitals.eu.

8 Sarah Thomson and others, 'International Profiles of Health Care Systems, 2013' (Commonwealth Fund) <http://www.commonwealthfund.org/~/media/Files/Publications/Fund%20Report/2013/Nov/1717_Thomson_intl_profiles_hlt_care_sys_2013_v2.pdf> accessed 14 May 2014; Health Consumer Powerhouse, 'European Health Consumer Index' <http://www.healthpowerhouse.com/files/EHCI_2015/EHCI_2015_report.pdf> accessed 16 April 2016; 'Mirror, Mirror on the Wall, 2014 Update: How the U.S. Health Care System Compares Internationally - The Commonwealth Fund' <http://www.commonwealthfund.org/publications/fund-reports/2014/jun/mirror-mirror> accessed 17 June 2014.

## 1.4 Selection of Medical Cases

The two areas of medical practice guidelines that have been chosen are lower back pain and diabetes type II. These areas of medicine were selected following an initial literature review of suitable comparative topics. This initial analysis focused on areas which had a history of guidelines production, a treatment regime that is comparable, and covered conditions that are prevalent in both countries. Back pain and Diabetes type II both fit this bill. Guidelines on both conditions have a long history going back to the early nineties, both conditions affect a large part of the population, and the treatment regimens are comparable. There are several reasons why these two areas of medicine provide for an interesting comparison. The following table highlights the key characteristics of the two fields of medicine that will be the focus of the thesis. The two fields were assessed on the four characteristics that follow from the literature reviewed in chapter two: evidence, institutions, conflicts of interest and costs.

*Table 1: Characteristics of the medical fields of the chosen guidelines*

| Angle of analysis | The Lower Back Pain Field | The Diabetes Field |
|---|---|---|
| *Evidence* | This is a field of medicine that is moving slowly, and presents no radical changes in the last years. | This is a field of medicine that is developing fast. Every year new medication comes on the market. |
| *Institutions* | In the Netherlands several groups of medical practitioners make their own guidelines. In England NICE has established a multidisciplinary group to handle this process. | In the Netherlands several groups of medical practitioners make their own guidelines. In England NICE has established a multidisciplinary group to handle this process. |
| *Conflicts of interest* | The risk of conflicts of interest is assessed as low, given the low profitability of lower back pain treatment. | Diabetes care is a huge and growing market, the massive marketing budgets spend by the world's largest pharmaceutical companies make the chance of conflicts of interest high. |
| *Costs* | The cost of treatment for lower back pain is relatively low (although the societal costs can be large) costing no more than £8,000 per year. | The cost of treating diabetes is high, as diabetes is a chronic life-long condition for which there is no cure, the cumulative cost of dealing with diabetes is substantial. |

As is clear from the table the two medical fields also provide a contrast to each other, and this allows us to compare and contrast the various relevant aspects of the formation of guidelines to help us answer the research questions. This comparative exercise can take place across two different axes of enquiry: across countries and across medical

fields. The differences in institutional national contexts allow us to explore whether this causes differences between guidelines.

Having successfully identified the field of medicine to focus on, the next step is to select the guideline producing organisations to further demarcate our enquiry. In England the choice is relatively straightforward: no-one can escape the dominance of NICE (the National Institute of Health and Care Excellence) as a producer of guidelines. This organisation and the guidelines on lower back pain and diabetes authored under NICE will be one half of our case study. In the Netherlands, there are several organisations of medical professionals that produce guidelines. Among them are the KNMG (Royal Dutch Federation of Medicine), NHG (Royal College of General Practitioners), and several organisations of medical consultants across the medical specialists: for example, surgeons, anaesthesiologists, and paediatric doctors. Of these organisations, GPs have had a long history of producing guidelines: since the eighties they have produced both a guideline on back pain as well as on diabetes.[9] Choosing to focus on the NHG in the Netherlands facilitates a meaningful comparison because we will only have to compare guidelines drafted by two different institutions. The only potential downside in choosing GPs is that unlike NICE their guidelines are not produced in a multidisciplinary group of professionals. This, however, could also allow this study to test, as part of the research design, whether

[9] So called NHG-standaarden, more detail is provided in chapter 2.

changes in the professional background of the medical professional drafting the guideline has a defining influence on the guidelines.

## 1.5 The Methodological Choices Explained

The conceptual/analytical framework used to structure this thesis comprises six different elements of analysis, which when taken together aim to provide a comprehensive overview of the guideline formation process in the Netherlands and England. These elements are: evidence base, achieving consensus, external responses, conflicts of interest, institutions and costs. The following section shows how these elements of the framework will be analysed and the methods underpinning the study.

This study compares England and the Netherlands using a case-study approach. There are good methodological reasons for choosing a comparative case study approach and using interviews and document analysis. Choosing this approach allows an in-depth study of the process of constructing guidelines, while at the same time enabling a comparative approach. This study has chosen a 'most-similar case study' selection process for the countries concerned, combined with a 'most-different case study selection for the guidelines.[10] This is appropriate as the initial literature survey observed a variety in outcomes in particular guidelines between the two countries despite the similar background

[10] Robert K Yin, *Case Study Research: Design and Methods* (Sage 2014).

in guideline production in the two countries. To assess what (combination of) factors caused these differences, it is helpful to try to minimize external factors that could complicate the comparison. Of course, this does not go so far as to say that we should select two identical cases, the cases do differ slightly, but an attempt has been made to minimise the differences.

However, parallel to this the guidelines on diabetes and lower back pain have purposely been chosen as coming from very disparate medical fields so as to maximise the difference between them. The combination of these case studies enables analysis of the full spread of cases within the wider set of 'medical practice guidelines' and helps to draw out broader conclusions on how guidelines are developed.

*Interviews as a Methodological Tool*

To gain a good understanding of the drafting of medical guidelines, interviews with the medical professionals involved in developing these guidelines is an essential part of the research. As Marshall and Rossman note, interviews are particularly suited for asking people about their lived experiences.[11] Using interviews helps us understand not only what the participants did but also how they reflect on what they did. The gap between the written text of a guideline and the recollection of the experience of drafting the guideline provides

[11] Catherine Marshall, *Designing Qualitative Research* (4th ed, SAGE 2006) <http://www.loc.gov/catdir/toc/ecip0519/2005026958.html> accessed 23 April 2016.

us with insights that can be fruitfully analysed. Contrasting different views of the same event from different standpoints using a multitude of interviews can help build a rich understanding.

The medical professionals involved in drafting the guidelines were identifiable through the published guideline documents. As many as possible were approached and asked to participate in a face-to-face interview, a total of 16 were interviewed 8 in both countries. In addition to interviewing medical professionals, two pilot interviews were held in the Netherlands with a specialist in medical guidelines and a compliance officer in a rural hospital to test the feasibility and appropriateness of the topic list and to help narrow the study. Furthermore, pilot interviews were held with a senior official at NICE and two officials at the Zorginstituut Netherlands to discuss broader aspects of the work of guideline groups.

The interviews were conducted in a semi-structured fashion using a topic list developed based on the literature discussed in chapter 2. They were conducted face-to-face as well as over the phone, and lasted between 45 minutes and three hours. The interviews in the Netherlands where conducted in August 2014 as well as June 2015, and the interviews in England were conducted in February-May 2015. The material emerging from the interviews was then transcribed and analysed linking relevant elements to the different

parts of the literature. This was done using a simple coding system: coding parts of the interviews in relation to their relevance for various parts of the framework.

Quotes from the interviews are presented in such a way that they cannot be attributed to a single participant. To this end sometimes the gender of the speaker has been changed (in groups with only one female group member), or quotes are ascribed to the group without disclosing the role of the person in the group (chairman, technical lead, member, surgeon, GP etc.), to as far as possible prevent easy identification. Sometimes this restricts the studies ability to explain statements in light of the position the person has in the guideline group. However, it was felt that allowing participants the opportunity to speak without feeling restricted outweighed this difficulty. This also allows the study to talk in detail about the specific area (back pain, diabetes) the group worked on. This is a vital advantage compared to other studies which used interviews or observations and did not disclose the disease under discussion. It is of greater importance to know in detail the field the group worked in, as this potentially helps explain more of their behaviour than their specific professional background.

*Analysis of Guideline Documents as a Methodological Tool*

The interviews with the members of the guideline group provide valuable information about the process of drafting a guideline, nonetheless, it is vital to scrutinize the wording of the final guideline document and review the draft versions if they were available.

In England NICE publishes a comprehensive list of documents that outline the guideline process: this includes scope documents, meeting notes, appendices with background studies and draft versions of the guideline, as well as comments received from registered stakeholders. In the Netherlands only the final guideline document is available. All these documents are filed online and have been studied to be able to reconstruct as closely as possible the development of the guideline. The responses in medical journals to the guideline, as well as any responses in the press, also add to the evidence base of material available.

These documents were then analysed in a systematic fashion starting with the final guideline document. The English and Dutch guideline documents were compared section by section. Three aspects were explored in detail for each guideline: (i) any references to evidence given in the document; (ii) the reasoning behind the adopted position on a treatment/medication/advice; and (iii) any reference to costs.

References to evidence in the guideline section were analysed in detail. This analysis was driven by the question: what evidence did the guideline group rely on and

how did this differ from the evidence relied on by the other guideline group? The documents referenced within a guideline were also examined (for example a controlled case study, randomised control trial, systematic review etc.), if particular reference was made to the document or when it was important to better understand the reasoning employed. An overview of all the material referenced is then made to be able to compare the guidelines across the whole process and to assess to what extent the evidence-base underlying the guideline is comparable. This systematic process also highlights where decisions have been taken with reference to expert opinion (of members of the guideline group or otherwise), and so highlights what type of evidence is used and where evidence seemed to be lacking.

This information then fed into the next stage of analysis, which examined the reasoning given in the guideline document. Stated reasons were compared across countries and provided opportunities for questions during the interviews. The reasoning was further analysed to see to what extent reasons given were based on a reference to evidence, expert opinion, societal cost or other considerations of the guideline group. Contrasting this analysis with the answers given in interviews allowed for a better understanding of the foundation for decisions in the guideline.

The third element analysed is any reference made to cost. Costs are taken here to be a separate category, although there are clear links with both the evidence-base and the reasoning in the guideline document. Nonetheless, it is useful for analytical purposes to

look at costs separately. Costs are often presented in a separate document, the calculations and reasoning employed uses a different logic, one more closely aligned with economic methods, and the actors involved sometimes differ from those involved in other parts of the guideline process.

By pulling these various elements together the guideline documents provide an integral part of the thesis. A critical reading of these documents helps provide the impetus to describe the gap between the 'law in the books' and the 'law in action'. The text in these documents is not approached as definitive, and no attempt is made to discover their absolute meaning in the sense that some legal research purports to when analysing text in a system of law. Instead, the focus of this research is on the text as the embodiment of the compromises struck in the process of negotiation. In addition to only looking at what is said in the guideline, it is also important to understand what is not said. What issues are not dealt with, or recommendations not made, might say a lot about the drafting process. When these omissions are spotted (through comparing various guideline documents), this fed in to the interview process.

*Limits of the Methodology*

In this research the focus is placed on the process of negotiation of a guideline, and as a result, less attention is paid to the implementation of the guidelines, which is only briefly explored in this thesis. In this sense, this study is more interested in aspects of the *process of development* than in exploring the impacts. To establish the details of this process of

negotiating guidelines, this research has chosen to examine guidelines which have already been negotiated, and thus looks back on a past process. This study does not include an analysis of guidelines as they are being negotiated, and it chose not to adopt the methods of participant observation. This research, however, does have the possibility to contrast the final guideline documents with the interviews with participants, and in this sense, is able to go deeper into the underlying compromises that have to be struck than if only a guideline in progress could be studied. Furthermore, this approach also allows for the inclusion of more comparative case studies over a more focused approach on one guideline using ethnographic methods.

In using interviews there are naturally also downsides and limits. There is often no objective truth about the events described, and all participants will offer their subjective interpretation of events informed by their own role in the process. As a researcher one needs to be aware of this type of bias and try to use a process of triangulation to minimise the effect on the results. Despite these limitations, the chosen research methodology seems to be the best possible fit for understanding the mechanisms that drive differences and similarities in guideline development.

## 1.6 Structure of the Thesis

Chapter 2 provides an introduction to the academic context of medical guidelines. It presents an overview of relevant literature on several aspects of guideline development:

evidence, regulation, institutions and factors shaping group decision-making. These strands of literature will then be used to analyse the research question, exploring to what extent the existing literature captures the development of medical guidelines.

Chapter 3 introduces the field, presents the various relevant actors operating in the Dutch and English healthcare system and provides a historical backdrop to the development of medical practice guidelines. This chapter will explain the development from consensus-based guidelines to evidence-based guidelines and its effect on guideline development. It discusses the main English and Dutch healthcare organisations that affect guideline development.

Chapters 4 presents a case study of one Dutch and one English guideline on the treatment of a-specific lower back pain. It presents a Dutch and English guideline on back pain providing an in-depth discussion of the development process of these guidelines; using the similarities and differences between these guidelines to explore the mechanisms for development of guidelines in these cases, and drawing insights from comparing them.

Chapter 5 presents a case study of one Dutch and two English diabetes type II guidelines. It presents these guidelines contrasting the development of the Dutch with the English guidelines. This case study consists of systematically analysing the differences and similarities to highlight mechanisms that operate in the process of developing guidelines.

Chapters 6 draws on the case studies to present a comparative analysis of all four cases. In this chapter the findings from both chapters 4 and 5 are brought together to establish whether connections exist between the various case-studies and establish which mechanisms and factors play a role in guideline development in the Netherlands and England and to what extent.

Finally, chapter 7 provides a conclusion drawing upon the literature analysed in chapter 2 and the analysis conducted in chapter 6. Conclusions will be drawn on to what extent this study fills existing gaps in the literature and presents new findings on the development of evidence-based medical guidelines.

# 2 UNDERSTANDING GUIDELINES IN THEIR ACADEMIC CONTEXT

## 2.1 Introduction

In this chapter, the existing scholarship that engages with medical guidelines is presented and the issues the literature presents are discussed with the aim of positioning this thesis within the field of study. It is important to connect the research in this study with literature from various different disciplines to allow for a broad analysis of the development of medical practice guidelines in a comparative perspective. The literature that is instrumental in understanding the process of development of medical guidelines can be positioned under four broad categories: evidence, regulation, institutions, and group-decision making.

An important element of this thesis is an engagement with the notion of evidence. Evidence is used in the medical field in a specific technical sense. It is understood to refer to a body of *knowledge*, and not just any (medical) knowledge but knowledge that conforms to standards in the methods that were used to create it. It is important to engage with literature on medical evidence, and evidence-based medicine, of which it is a part. Engaging with this literature is important for understanding the development of evidence-based medical guidelines because guideline drafters will be influenced by their conception of evidence in making a guideline. This influence comes in at least three ways: a conception of what evidence is informs what drafters see as informing a guideline, it shapes their understanding of the process of making a guideline and it influences their ideas about what

constitutes a good guideline. The debates in the literature about what counts as evidence, the process of using evidence, and measures of success in making guidelines will be explored.

Regulatory literature is a second source of information that helps to understand the drivers and the processes that influence the development of medical guidelines. It will be argued in this thesis that guidelines are not mere forms of advice, that can be easily ignored, but rather are regulatory tools used by professionals for two distinct purposes: to ward off the influence of the state on the profession, and to influence other professionals and steer the quality of their practice of medicine. The regulatory literature provides interesting theories and debates on the nature of self-regulation by professionals, and this study will be positioned as a contribution to a better understanding of these aspects of the regulatory state.

The work of a guideline group furthermore purports to be, at least in theory, independent. This thesis argues an essential aspect of the development of guidelines can be explained by specific features of its institutional embedding within a guideline making organisation. The institutional literature informs the approach of this study by examining the interplay and tensions between institutions and the guideline groups that work for them.

Finally, guideline making groups will also be studied at the micro level. The interactions of guideline makers and in particular their processes of decision-making form

the heart of this work. Accordingly, several different theories and approaches that analyse group-decision making will be explored to allow this work to add to existing important scholarship in this field.

## 2.2 Evidence

Guidelines are commonly defined as: 'systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances.'[1] In practice guidelines can have different names: protocols, standards, practice statements or guidelines. In this study I will use medical practice guidelines, or just guidelines, when describing all of these.

Medical practice statements are described as 'systematically developed.' This description points to two types of guidelines: consensus-based or evidence-based guidelines. The consensus-based development of guidelines refers to those guidelines adopted in the 1980s where large conferences where held where healthcare professionals decided, by consensus, to adopt recommendations and to draft guidelines.[2] This approach has been replaced, slowly but steadily, by a focus on an evidence-based development of

---

1 Marilyn J Field and others, *Clinical Practice Guidelines Directions for a New Program* (National Academy Press 1990) 38.

2 Nick Black and others, 'Consensus Development Methods: A Review of Best Practice in Creating Clinical Guidelines' (1999) 4 Journal of Health Services Research & Policy 236; JJE van Everdingen, *Consensusontwikkeling in de geneeskunde* (Bohn, Scheltema en Holkema 1988).

guidelines. This historical development, and its description in the literature is directly relevant to this study as it colours the conception guideline developers have of what making a guideline entails. Medical practice guidelines did not move to being described as evidence-based without the impetus of a broader movement that has had (and still has) an effect on medicine: evidence-based medicine. To understand how guideline group members might approach the task of developing a medical guideline an analysis of the evidence-based medicine movement and its central philosophy is in order.

The term evidence-based medicine was coined in 1991 by Guyatt[3] and describes an approach to clinical decision-making that builds on findings from clinical epidemiology and the promotion of the Randomised Control Trial (RCT) by Cochrane.[4] This approach set out to challenge doctors who would take decisions based on clinical expertise or knowledge gained from other professionals, and started promoting the systematic use of medical studies in clinical decision making. Sackett describes this approach in the following terms: 'The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic

---

3 GH Guyatt, 'Evidence-Based Medicine' [1991] ACP J Club A.

4 Iain Chalmers, 'The Cochrane Collaboration: Preparing, Maintaining, and Disseminating Systematic Reviews of the Effects of Health Care' (1993) 703 Annals of the New York Academy of Sciences 156; AL Cochrane, J Glyn Cox and T Francis Jarman, 'Pulmonary Tuberculosis in the Rhondda Fach' (1952) 2 British Medical Journal 843; AL Cochrane, *Effectiveness and Efficiency: Random Reflections on Health Services* (Nuffield Provincial Hospitals Trust 1972).

research.'[5] The take up of this call has been dramatic: journals[6], research institutes[7] and medical training[8] have all been reformed based on these ideas. Evidence-based medicine has not been interpreted in a uniform way. While Sackett stresses the need to select the appropriate evidence for the clinical question at hand, one of the defining features of evidence-based medicine is the introduction of an evidence-hierarchy in medicine.

The background to the term evidence-based medicine lies in the medical discipline of epidemiology, a discipline that studies diseases on the level of populations. RCTs are a type of scientific experiment to determine whether a treatment is effective through the use of randomly assigned participants into either a control group (who receive no treatment) or an intervention group (who receive treatment), in order to establish causal links. RCTs are seen as scientifically more rigorous and thus are labelled as a more valuable way of assessing treatments and diagnosis. When RCTs are coupled with an emphasis on the need to consult and appraise the scientific literature in making clinical decisions they can improve the evidence-base of those decisions.[9] Other forms of scientific evidence are also

---

5 David L Sackett and others, 'Evidence Based Medicine: What It Is and What It Isn't' (1996) 312 BMJ 71, 71.

6 Frank Davidoff and others, 'Evidence Based Medicine' (1995) 310 BMJ 1085.

7 The Oxford Centre for Evidence-Based medicine is but one of many examples.

8 JH Shin, RB Haynes and ME Johnston, 'Effect of Problem-Based, Self-Directed Undergraduate Education on Life-Long Learning.' (1993) 148 CMAJ: Canadian Medical Association Journal 969.

9 Evidence-Based Medicine Working Group, 'Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine' (1992) 268 JAMA 2420; Roger L Sur and Philipp Dahm, 'History of Evidence-Based Medicine' (2011) 27 Indian Journal of Urology : IJU : Journal of the Urological Society of India 487.

included in producing medical guidelines, such as expert opinion, cohort studies and case studies. Cohort studies are a form of longitudinal prospective study used to establish a correlation between risk factors and a certain disease. While case-studies test for a predefined causal attribute, such as smoking, in a group of people with a disease, say lung cancer, compared to a group of people without lung cancer. If a significantly greater amount of people smokes in the group with lung cancer there is a strong correlation but no causation between smoking and lung cancer. Critically, as these types of studies are observational they can only provide correlations while randomised control trials, if conducted properly, can establish causal links.

In evidence-based medicine, forms of evidence are ranked by type and placed in a hierarchy of quality of evidence: with RCTs given most weight, cohort studies and case studies given a medium ranking, and expert opinion ranked as of lowest value. As such evidence-based medicine proposes a hierarchy of evidence as displayed in this diagram.

*Figure 1. The Evidence Hierarchy*[10]

Evidence-based guidelines use all of forms of evidence, and the relative assessment of their value, to produce recommendations for treatment. From this we can deduce that there is a preference in evidence-based medicine to favour knowledge that is the result of statistical inferences for population level research on diseases. It is precisely this hierarchy that is innovative in evidence-based medicine; nobody goes so far as to say that earlier

---

[10] © Some Rights Reserved CC-BY-NC-SA 3.0 licence (https://creativecommons.org/licenses/by-nc-sa/3.0/). Levels of Evidence, Evidence-Based Practice in the Health Sciences Tutorials, Information Services Department of the Library of the Health Sciences-Chicago, University of Illinois at Chicago. https://uic.blackboard.com/courses/1/cont.library_tutorials.jamied.1/content/_4166951_1/dir_NursingResearchDesign.zip/index.html.

decisions in medicine were not based on any evidence, only that an explicit hierarchy of evidence is new. This epidemiological approach to medical knowledge is true to its foundations in epidemiological research, but therefore also imports a specific epidemiological view of medicine into mainstream thinking.

To further understand the precise nature of evidence in specific evidence-based medical guidelines it is important to outline the broader canvas of scientific knowledge and use of evidence by scientific communities. The use of knowledge, falsifiability of scientific claims, paradigms, and the Duhem–Quine thesis, will be discussed briefly to understand the foundations of evidence-based medicine, and certain critiques formulated against it, and its relevance to this thesis.

The use of evidence in science has been discussed extensively by Thomas Kuhn and Karl Popper: in general Popper is interested in the logic of science, while Kuhn focused on the history of science.[11] Kuhn contends that it is a scientist's work within a paradigm that structures how they see their data, this means that our observations are always theory-laden and can never reveal the 'natural order' of things. These theorists are relevant for the

[11] Thomas S Kuhn, *The Structure of Scientific Revolutions* ([2d ed, enl, University of Chicago Press 1970).

formulation of guidelines as their work examines the nature of science in society and open up 'scientific procedures' for critical scrutiny by social scientists.

Popper's claim that scientific knowledge depends on the falsifiability of statements established a specific way of looking at the logic of knowledge, and by introducing this criterion intended to provide a way of demarcating science from non-science.[12] Progress in science can therefore be achieved through the falsification of theories of knowledge, and their consequent replacement by theories that have greater explanatory power to explain the phenomena the theory relates to. Many have critiqued Popper and his distrust of induction as a method of science, but a full assessment of this critique is beyond the scope of this thesis.[13] Popper's work is important for our discussion as randomised controlled trials (RCTs), an important element of evidence-based medicine, are based on a logic that looks for falsifiable statements. It will be explored in this thesis how members of guideline groups evaluate evidence based on an understanding of science informed by these philosophical critiques.

Lakatos has extended both Popper's and Kuhn's work by identifying research programs, comparable to Kuhn's paradigms, which consist of a 'hard core' together with

---

12 Karl Raimund Popper, *The Logic of Scientific Discovery* (Psychology Press 2002).

13 For example Paul Feyerabend, *Against Method* (Verso 2010); David Miller, 'Some Hard Questions for Critical Rationalism' (2014) 15 Discusiones Filosóficas 15.

auxiliary hypothesis.[14] Lakatos suggests that auxiliary hypotheses can be abandoned without threatening the theoretical core of a theory. Only when this hard core is refuted a paradigm (or research programme) shift occurs, this is in line with Kuhn's historical view of science. Lakatos' auxiliary hypotheses were inspired by the work of Duhem. Duhem[15] and Quine[16] extended Lakatos' theory, by stating that it is impossible to test a scientific hypothesis in isolation because this requires background assumptions about other hypotheses that need to be correct. So, if one falsifies a single hypothesis it is not known whether this was because the hypothesis is indeed false or because one of the background assumptions or auxiliary hypothesis are false. Again, this is relevant for the analysis of evidence-based medicine because RCTs aim precisely to do what Duhem and Quine suggest is impossible: to test hypothesis in isolation. Duhem and Quine suggest that it is possible to falsify a theory, or a bundle of hypothesis as a whole, so in this sense their work is congruent with that of Kuhn and Lakatos. Equally, it could be argued that guideline group members may not want to evaluate evidence in isolation, but examine groups of medical studies to help them decide whether a change of clinical practice is warranted. As a result, these philosophical concerns may feed directly into the development process of a

---

14 Imre Lakatos, *The Methodology of Scientific Research Programmes* (Cambridge University Press 1999).

15 Pierre Maurice Marie Duhem, *The Aim and Structure of Physical Theory* (Princeton University Press 1991).

16 WV Quine, 'Main Trends in Recent Philosophy: Two Dogmas of Empiricism' (1951) 60 The Philosophical Review 20.

medical guideline. Understanding academic thinking on philosophy of science is directly relevant for this thesis.

*Critiques of the Evidence Hierarchy*

An epidemiological approach has not been accepted uncritically within medicine: there are several critiques that are important in considering the role of evidence. To understand the role evidence plays for guideline makers we need to understand how doctors and policy-makers understand and use knowledge.

Evidence-based medicine has been challenged within the literature on the grounds of its selectiveness and conservative employment of knowledge.[17] For example, Chin-Yee[18] is insightful in his critique of the limited nature of knowledge in evidence-based medicine. Chin-Yee addresses 'underdetermination.' Underdetermination is the idea that 'empirical evidence alone provides insufficient grounds for acceptance or rejection of

---

17 Among many: Abhaya V Kulkarni, 'The Challenges of Evidence-Based Medicine: A Philosophical Perspective' (2005) 8 Medicine, Health Care and Philosophy 255; Benjamin H Chin-Yee, 'Underdetermination in Evidence-Based Medicine' (2014) 20 Journal of Evaluation in Clinical Practice 921; Aaron Michael Cohen, P Zoë Stavri and William R Hersh, 'A Categorization and Analysis of the Criticisms of Evidence-Based Medicine' (2004) 73 International Journal of Medical Informatics 35; Maya J Goldenberg, 'On Evidence and Evidence-Based Medicine: Lessons from the Philosophy of Science' (2006) 62 Social Science & Medicine 2621; John Worrall, 'What Evidence in Evidence-Based Medicine?' (2002) 69 Philosophy of Science S316; Edwin Harari, 'Whose Evidence? Lessons from the Philosophy of Science and the Epistemology of Medicine' (2001) 35 Australian and New Zealand Journal of Psychiatry 724.

18 Chin-Yee (n 17).

scientific theories, and therefore underdetermines our beliefs about the world'[19]. The relevance of this argument for our study is his contention that we need several different forms of knowledge to overcome the problem of underdetermination. Chin-Yee rightly draws attention to the fact that 'evidence-based medicine' is quite a specific way of conceptualising medicine, and that there are other complementary ways of generating knowledge, in his words:

> Such an epistemic attitude imposes limitations on medical research: by failing to probe core beliefs regarding the mechanisms of disease and therapeutics, EBM fails to generate a theoretical framework that is necessary for hypothesis generation and integrating knowledge to develop new treatments. Although clinical studies are an important component of medical research, clinical trials are not the sole epistemic authority in medical knowledge nor are they effective as a singular research heuristic.[20]

While these arguments are mainly directed towards clinical science, and the epistemological underpinnings of research, these ideas can be neatly transposed into the domain of guideline construction. This literature reminds us that there are competing knowledge claims at play and that the evidence hierarchy as envisaged by evidence-based medicine is not the only way of conceptualising evidence in medicine. In the formation of guidelines, there are two complementary critiques possible of 'evidence' generated by randomised clinical trials: one draws attention to the fact that there might be

---

19 ibid 922.

20 ibid 921.

methodological limitations in conducting an RCT, and the second that highlights that some knowledge claims cannot be made by using clinical trials. Both elements are relevant if we want to conceptualise the nature of evidence use in guideline creation.

If we take the first point about the methodological limitations of RCTs, we can see a real challenge for the use of evidence in medicine. We can challenge whether the results of an RCT are valid in the 'real' world, and whether these results should lead to the conclusion that a causal relationship exists between the intervention and the alleviation of symptoms. This is a pertinent point in many discussions in guideline groups existing empirical research suggests.[21] Critiques on the methodological limits of RCTs reveal the amount of assumptions necessary before we can infer a causal relationship between Treatment X helping to cure Disease Y. It also helps understand that appraisal of evidence is of vital importance in deciding to what extent a certain study can dislodge an entrenched idea about a given disease or condition. A parallel with Kuhn's paradigm is appropriate here; in that one study that disproves a 'theory' about how a disease should be treated does not immediately necessitate a change, but if a body of knowledge exists that points in a different direction this might dislodge the orthodox position and lead to a new paradigm.

---

[21] Tiago Moreira, 'Diversity in Clinical Guidelines: The Role of Repertoires of Evaluation' (2005) 60 Social Science & Medicine 1975.

The second critique by Chin draws attention to the fact some knowledge cannot be established by RCTs. It is also necessary to understand disease pathology and rely on basic science to help establish the effectiveness of treatments because they reveal the underlying mechanisms of disease.[22] The evidence-based medicine approach thus fails to ask the question *why* treatments work, and focusses only on the fact that they do appear to work. Clinical knowledge about disease pathology is necessary, however, to distinguish between classes or groups of patients for which a certain study is relevant and others for which that is not the case. This helps us understand the postulated objective nature of evidence-based medicine. The evidence hierarchy suggests that higher quality evidence is more objective and thus better, however, we now know normative assumptions about the nature of knowledge underpin this hierarchy. In this context knowledge is neither 'natural' nor complete, but rather based on several normative assumptions about the nature of knowledge and thus is in essence politically and socially constructed.

It is this political nature of knowledge that deserves further scrutiny. This thesis will address these arguments about knowledge and engage with these philosophical concerns in an empirical way. This will provide an important addition to this strand of the evidence-based literature because so far the question of whether these philosophical doubts about the theoretical underpinnings have any effect on how clinicians think of evidence-

---

[22] Chin-Yee (n 17) 926.

based medicine has not been addressed. By combining the literature that paints a normative picture of what evidence-based medicine should look like, with knowledge on how this is brought into practice, a more detailed understanding of the role of evidence-based medicine in influencing the development of medical guidelines can be reached. In this way, this thesis can be seen as filling a gap in the literature because these arguments have not been brought to bear in empirical study of guidelines.

*Objectivity and Evidence*

Evidence-based medicine is not only analysed from a medical viewpoint, the literature in the tradition of science and technology studies understands guidelines in a broader sense as an attempt to create objectivity. A different strand in the science and technology literature stresses the normativity of the use of instruments like guidelines and it is helpful to contrast these two different approaches. The (postulated) objectivity of evidence has been problematized in this literature as socially bound, context dependent, and historically shifting. The literature behind this claim sits largely in the field of science and technology studies, which build on Latour's actor-network theory.[23] For example, Cambrosio et al have shown how evidence-based medicine sits within a broader tradition of shifting ways of creating objectivity in medical science which is based on the collective production of

[23] Bruno Latour, *Reassembling the Social: An Introduction to Actor-Network-Theory* (OUP Oxford 2005).

evidence.[24] The authors chart an historical transformation of what objectivity means in science, culminating in what they call 'regulatory objectivity'. They posit that different forms of objectivity are in play concurrently:

> As previously noted, recent historical studies have transformed scientific and medical objectivity into a subject of historical inquiry. They have shown how different historical periods have produced different types of objectivity that have subsequently persisted either as autonomous forms or in combination with other types of objectivity. Thus, whereas the objectivity of a statement was once guaranteed by the knowledge and experience of the author of the statement, later periods have tended to privilege mechanical or instrumental objectivity that replaces experts' subjectivity with mechanically produced inscriptions.[25]

This shift from expert to mechanical objectivity can be taken one step further; as Cambrosio et al explains current arrangements are a combination of human expertise and objects:

> There has therefore been a reversal of objectivity: whereas mechanical objectivity displaces the bearer of objectivity from the human expert to the object, regulatory objectivity turns the focus away from objects towards collective forms of expertise combining people (clinicians, researchers, administrators, patients, etc.) and objects (entities, instruments, tools, techniques, etc.) connected by specific coordination regimens.[26]

---

24 Alberto Cambrosio and others, 'Regulatory Objectivity and the Generation and Management of Evidence in Medicine' (2006) 63 Social Science & Medicine 189.

25 ibid 193.

26 ibid 194.

These specific ‘coordination regimens’ that make the collective production of regulatory objectivity possible include medical guidelines. This approach does not stress the intricate detail of what is created, but rather focuses on the external effect of that creation: and so the impression of objectivity is created for people viewing the guidelines. One thing we can take from this discussion of regulatory objectivity is the collective nature of considering and creating evidence. Noting that the evidence that goes into an evidence-based guideline is transformed through a collective process to produce ‘objective scientific facts’ at the end of the exercise. This approach categorically highlights the social nature of the process of guideline construction and the social choices that are made within this process. While Cambrosio et al claim a clear shift to regulatory objectivity, and suggest coordination regimens play a vital role, they leave unanswered how exactly these coordination regimens would function, and how it is that people would see the result of this process as objective. It is these questions that this thesis addresses. This strand of argument is important because it provides insight into a possible way to conceptualise how members of guideline groups might work with medical knowledge and how they can create guidelines that are accepted by medical professionals.

The literature on evidence-based medicine and the use of evidence in guidelines presents a varied picture. On the one hand evidence-based medicine is heralded as a fundamental change in doing medicine while others show the importance of ‘traditional’ sources of knowledge. While much of the medical literature, in its nature, presents

evidence-based medicine as a fact and devotes itself to understanding the best way of implementing it, the philosophical and science and technology studies literature provide a more critical engagement with this movement and the conceptual language it has introduced. In this study evidence will play a major role because it describes a large part of the language that is used in guideline documents and the discourses surrounding guideline documents. This study will address *evidence* in a socio-legal enquiry and interrogates what evidence means for people, how they use it, and what the effects are of this particular language. It seems plausible that the way people negotiate the meaning of a technical term like evidence will expose a gap between the text-book definition of evidence and the reality of its socially constructed use.

## 2.3 Regulation

Guidelines are, in contrast to their names, not mere forms of voluntary guidance that can be ignored without consequence but rather a specific form a professional self-regulation. Before developing this argument further, it is helpful to set out an analysis of the regulatory literature and the position self-regulation has within it. This analysis starts with a look at the development of the regulatory state and presents an analysis of the changing nature and conception of the role regulation plays in society. Before going on to establish arguments which view guidelines as regulatory tools that play a role in the regulatory state. This literature helps us explore the argument that guidelines as a regulatory tool can been seen

as a method to ward off (undue) influence of the state on a given profession. The second body of literature relevant here deals with issues of professionalization and professionalism, and from this literature we can develop ideas about how guidelines are used as a regulatory tool by medical professionals, and are aimed at achieving behavioural change in clinical practice. The contribution this thesis aims to make to debates in these two fields is to show how they are connected, and bring these disparate literatures together, in such a way that it illuminates how this particular form of medical regulation functions in a western welfare state. A separate section is devoted to the issue of transparency, and the insight the regulation literature provides on the importance of the transparency of regulatory initiatives.

Regulation as a discipline can be approached through an exploration of the literature that is concerned with the changing nature of western states. The regulation literature developed as a response to the (perceived) changing nature of the state's power and the way the state regulates the behaviour of its citizens. One influential definition is Black's 'decentred' definition of regulation:

> Regulation is the sustained and focused attempt to alter the behaviour of others according to defined standards or purposes with the intention of producing a broadly identified outcome or outcomes, which may involve mechanisms of standard-setting, information-gathering and behaviour-modification.[27]

---

[27] Julia Black, 'Critical Reflections on Regulation' (January 2002) 20 <http://www.lse.ac.uk/CARR> accessed 14 April 2016.

The classical model of regulation of citizens behaviour would consist of law-making in the form of concrete laws that ban certain conduct and contain detailed rules on how to achieve certain societal goals that legislation would promote, or centralised planning measures as was common in the former communist East-European States. The contribution of the regulation literature is to focus on how in reality the state governs its citizens, its governance arrangements, and to discuss how this has shifted from a command-and-control type of regulation to a regulatory state now achieving its goals through regulatory action. This shift is first explored and argued in relation to the EU by Majone.[28] He contends that the EU should be thought of as a regulatory state because instead of centralised administration of policies it now attempts to change behaviour through independent regulatory agencies.

This article meant a watershed in the thinking about the role of the state and Majone's thesis has been confirmed and modified in subsequent work: The regulatory

---

28 Giandomenico Majone, 'The Rise of the Regulatory State in Europe' (1994) 17 West European Politics 77.

state[29], The Audit Society[30], Regulation Inside Government[31], The New Regulatory State[32], The British Regulatory State[33], and Regulatory Capitalism.[34] All these approaches, despite differences in approaches and language, refer to the central phenomenon of the regulatory state, which is analysed by Wincott as the shift 'typically interpreted through conceptual oppositions that have taken us from government to governance, rowing to steering, providing to facilitating, national sovereignty to post-nationalism after the sovereign state and the welfare state to the regulatory state.'[35] Recently as a result of the economic crisis and its aftermath, there have been voices in the literature that question the validity of Majone's thesis for the EU today- but that analysis falls outside the scope of this work.[36]

---

29 ibid.

30 Michael Power, *The Audit Society* (Oxford University Press 1997) <https://global.oup.com/academic/product/the-audit-society-9780198289470> accessed 27 May 2016.

31 Christopher Hood and others, *Regulation inside Government: Waste Watchers, Quality Police, and Sleaze-Busters* (Oxford University Press 1999).

32 John Braithwaite, 'The New Regulatory State and the Transformation of Criminology' (2000) 40 British Journal of Criminology 222.

33 Michael Moran, *The British Regulatory State: High Modernism and Hyper-Innovation: High Modernism and Hyper-Innovation* (Oxford University Press 2003).

34 D Levi-Faur, 'The Rise of Regulatory Capitalism: The Global Diffusion of a New Order' (2005) 598 The ANNALS of the American Academy of Political and Social Science 200.

35 Daniel Wincott, 'European Political Development, Regulatory Governance, and the European Social Model: The Challenge of Substantive Legitimacy' (2006) 12 European Law Journal 743, 743.

36 James A Caporaso and others, 'Still a Regulatory State? The European Union and the Financial Crisis' (2015) 22 Journal of European Public Policy 889.

The model of the regulatory state also brought with it changing ideas on how to best govern professions that enjoyed a great deal of autonomy. The work by Power on the audit society deals with this component particularly well. In his work he provides an account of how the regulation of banking used to be a form of 'club government'[37] to a far more formalised relationship where the banking profession is more clearly regulated and conduct prescribed through detailed rules. The banking profession is now required to regulate its own behaviour, by self-regulation[38], and formulate rules, within a wider web of detailed governmental rules. In return for this (relative) freedom the government prescribes detailed policy outcomes that this regulation by the profession was meant to satisfy. Governments police this new arrangement through establishing independent regulators, who set standards and monitor the compliance of the profession. A result of these developments are central to this thesis: the increase of *mandated self-regulation*. The relationship between profession and state can be described in terms of that of mandated self-regulation. This means a situation where a collective group, such as medical professionals, is designated to formulate rules within a broadly defined framework set by the government.[39] Both the profession in the UK and the Netherlands, are subject to this

---

37 Michael Moran, *The British Regulatory State: High Modernism and Hyper-Innovation* (Oxford University Press 2003).

38 Described by Black as the situation of a group of persons or bodies, acting together, performing a regulatory function in respect of themselves and others who accept their authority.

39 Anthony Ogus, 'Rethinking Self-Regulation' (1995) 15 Oxford Journal of Legal Studies 97; Julia Black, 'Constitutionalising Self-Regulation' (1996) 59 The Modern Law Review 24.

mandated regulation. In England directly through NICE, in the Netherlands as a result of the Act on the Medical Treatment Contract.[40] Guidelines can be developed independently but have to fit within this broader framework. These developments are directly relevant to this project as the medical profession is confronted by the broader regulatory changes: medical guidelines play an essential role in a system of mandated self-regulation and warrant further investigation.

*The Impact of Healthcare Regulation in England and the Netherlands*

Regulatory developments are not confined to the financial sphere but have affected healthcare regulation as well. A good account of the changing landscape of British medical regulation is provided by Wright. He describes the regulation of the NHS as a result of the Labour reforms between 1997 and 2007 with the themes: 'the division of labour between state and society, the division of state, the codification of relationships, the development of meta-regulatory techniques for enforced self-regulation.'[41]

The emergence of NICE in 1999 as a special health authority fits well with the trend described in the regulatory literature and the themes described by White. The National Institute for Health and Care Excellence changed the division of labour between state and

---

40 Wet op de geneeskundige behandelovereenkomst 1994 (Stb 1994, 837).

41 John SF Wright, 'The Regulatory State and the UK Labour Government's Re-Regulation of Provision in the English National Health Service' (2009) 3 Regulation & Governance 334, 337.

society as it now fell within the realm of the state to support the regulation of the medical profession through medical guidelines. This is covered in the literature on self-regulation, which notes how professionalization was the result of a further division within the state, moving from direct ministerial control to a model with several bodies at arms-length with various levels of independence. For example, NICE formed a codification of the relationship with the medical profession, moving from an informal settlement with the General Medical Council and the Royal Colleges who used to be in charge of guidelines, to a formalised relationship with a government body. NICE is now an essential link in a chain of self-regulation that fits neatly with White's thesis on the development of meta-regulatory techniques for self-regulation. NICE can be described as using meta-regulatory techniques precisely because it facilitates (rather than drafting itself) the production of medical practice guidelines. The guidelines produced in turn form part of self-regulatory regime, one which is mandated through processes of compliance involving auditing and financial incentives. By overseeing guideline production NICE is not merely facilitating but also controls – as a regulator – guideline production. It could be said that it is through these controls, which don't take the form of formal rules but are themselves guidelines in the form of manuals on guideline production that NICE tries to shape the production of guidelines.

In the Netherlands, the NHG is a professional body of GPs and as such produces self-regulation in the form of guidelines.[42] This points to an important institutional and regulatory difference between the two countries, as the NHG is independent from the state, and the production of guidelines is not regulated directly by the state in the Netherlands. Does this mean that the Netherlands has not become a regulatory state, and that we should describe the model in the Netherlands as that of the 'old' welfare state? This would be a misunderstanding of the Dutch healthcare system.

The Dutch healthcare system is as much part of a regulatory state as the British healthcare system. The Dutch system is based on private hospitals and private healthcare insurers who compete to attract patients/insured. It is managed by the *Nederlandse Zorgautoriteit* and *Zorginstituut Nederland* who, similar to the English system, regulate the (quasi) market in healthcare. The details of the Dutch healthcare system are covered in greater detail in chapter 3. The key point for the current discussion is that the NHG and its guideline functions is one area where medical practitioners, it seems, have retained autonomy. This is because there is no central guideline producing institution in the Netherlands, but rather a multitude of guideline producing organisations: of which the NHG is one. Other institutions, such as the Zorginstituut, point towards the conclusion that the Netherlands is much like England and there is mandated self-regulation at play, even if

---

42 Called NHG-standaarden, www.nhg.org.

they initially appear more dispersed in their operation. The complicated precise interrelation between these elements of healthcare systems, institutions, and guidelines, is further explored in chapters 4 and 5.

*Guidelines as Regulatory Tools*

The changes that the rise of the regulatory state has brought to the regulatory systems and institutions of healthcare in both the Netherlands and England is relevant to this thesis because it provides one of the reasons why we see a multiplication of guidelines, it feeds a demand for accountability partly fuelled by different arrangement of healthcare governance as a result of the rise of the regulatory state in (western) Europe. Of course, the position is more nuanced than that, endogenous developments in medicine also played a role.[43] The proliferation of guidelines as such is not of central relevance to this thesis, but importantly this growth has led to a change of the role of guidelines, so much so that I would claim guidelines can now be seen as *regulatory tools.* As far as I am aware this argument has not been advanced in the scholarly literature before, but it follows from the existing regulatory literature and would place medical guidelines in the same category as the International

---

[43] Ruud van Herk, Rita Schepers and Anton F Casparie, 'Huisartsen en zelfregulering. De ontwikkeling van intercollegiale toesting en standaarden voor huisartsen tussen 1970 en 1990' (1994) <https://lirias.kuleuven.be/handle/123456789/161090> accessed 12 June 2014; Cornelia de Vos, 'Kwaliteit, disciplinering en sturing. Een historisch sociologisch onderzoek naar de vormgeving van kwaliteit van zorg in ziekenhuizen = Quality, discipline and governance. A historical sociological study into the design of quality of care in hospitals' (Erasmus Universiteit Rotterdam 2014).

Swaps and Derivatives Association (ISDA) master-agreement, International Organization for Standardization (ISO) standards, and financial standards developed by the Basel Group[44], for example.

To support the argument that guidelines are attempts at standard-setting, information gathering and behaviour modification to achieve certain outcomes, a short history of their development is set out. Guidelines started out, when first developed in the Netherlands and England as measures to ensure that scientific medical knowledge more was widely available and kept up to date. Guidelines were also designed to support internal auditing of medical professionals by other medical professionals, as Everdingen and Klazinga describe in detail.[45] Guidelines quickly developed a broader appeal and they became instruments to improve the quality of care.[46] With this development it also became possible for external parties to use guidelines as a transparency mechanism over what the

---

44 Julia Black, 'Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes' (2008) 2 Regulation & Governance 137; Julia Black and David Rouch, 'The Development of the Global Markets as Rule-Makers: Engagement and Legitimacy' (2008) 2 Law and Financial Markets Review 218.

45 Everdingen (n 2); Nicolaas Sieds Klazinga, 'Quality management of medical specialist care in the Netherlands. An explorative study of its nature and development' (Erasmus Universiteit Rotterdam 1996) <http://www.bmg.eur.nl/fileadmin/ASSETS/bmg/Onderzoek/Promoties/Promoties_1996/Klazinga_1996_web.pdf> accessed 24 September 2014.

46 R van Herk, 'Artsen onder druk. Het kwaliteitsbeleid van de medische beroepen in Groot-Brittannië, Nederland en België tussen 1970 en 1996, als gevolg van de interne en externe druk op de zelfregulering van artsen.' (Erasmus Universiteit Rotterdam 1997) <http://repub.eur.nl/pub/18416/971023_HERK,%20Rudi%20van.pdf> accessed 24 November 2014; Vos (n 43).

profession regards as good quality care.[47] Both in England and in the Netherlands the government use guidelines as a tool to reduce practice variation, and increase quality, while simultaneously saving costs by further rationalising care.[48]

Guidelines have become central objects in the complex interplay between the internal regulation of the profession and external control and oversight of self-regulation by the government. This has made guidelines tools of mandated self-regulation used by professionals to control professionals and thus a form of professional self-regulation, and *at the same time* an instrument of control for governments to co-opt professionals to fulfil government aims: such as, standardisation of care, cost-effectiveness, rationalisation and patient involvement.[49]

A crucial result of these developments is the fact that guidelines always suffer from a *tension* between these two interconnected aspects: a regulatory tool for control over clinicians and regulatory tool to satisfy and *limit* government control. By defining autonomy for the clinician in medical guidelines the profession tries to prevent the

---

47 N Klazinga, 'Compliance with Practice Guidelines: Clinical Autonomy Revisited' (1994) 28 Health policy (Amsterdam, Netherlands) 51.

48 Vos (n 43).

49 In chapter 3 the history of guideline development will be considered in greater depth in relation to the specific case studies under consideration. Transparency and its importance in understanding guidelines will be explored separately below.

government controlling this aspect of their professional role. It is this tension that will form a large part of the questions analysed in this thesis. This study will analyse whether this tension does indeed have an effect on the development of guidelines and how this might be reflected in the text of guidelines.

To explore in greater depth what guidelines as regulatory tools for professionals might mean, and how they are employed, we have to look beyond the regulatory literature to academic work on the sociology of professions.

*Professionalization*

The sociology of professions provides us with a framework to address the importance of guidelines as a project of professional regulation. Evetts provides a useful description of three central concepts in this sociological sub-field: profession, professionalization, professionalism.[50] Where profession refers to 'a distinct and generic category of occupational work'[51], Evetts notes there fails to be a generic definition that demarcates a profession from a mere occupation. Professionalization, on the other hand, is referred to 'as the process to achieve the status of profession and has been interpreted as the process to pursue, develop and maintain the closure of the occupational group in order to maintain

---

50 Julia Evetts, 'Sociological Analysis of Professionalism: Past, Present and Future' (2011) 10 Comparative Sociology 1.

51 ibid 6.

practitioners own occupational self-interests in terms of their salary, status and power as well as the monopoly protection of the occupational jurisdiction'.[52]

Medicine has always been regarded as a profession (notwithstanding the failure to arrive at a generic definition) and its professionalization has been described in terms of the 'medical dominance' thesis advanced by Freidson[53], but also in more pessimistic terms as suffering from 'proletarianization and de-skilling.'[54] The most relevant element of this literature for our discussion of guidelines, however, is the sub-field that engages with professionalism and its link to trust. As Evetts states:

> Professionalism in occupations and professions implies the importance of trust in economic relations in modern societies with an advanced division of labour. In other words, lay people must place their trust in professional workers (electricians and plumbers as well as lawyers and doctors) and some professionals must acquire confidential knowledge.[55]

Guidelines are a part of this process in a couple of interrelated ways. They are part of making trust more explicit in some aspects of the relationship between the doctor and the

---

52 ibid.

53 Eliot Freidson, *Professionalism, the Third Logic* <http://www.press.uchicago.edu/ucp/books/book/chicago/P/bo3621932.html> accessed 2 June 2016.

54 John B McKinlay and Joan Arches, 'Towards the Proletarianization of Physicians' (1985) 15 International Journal of Health Services 161; John B McKinlay and John D Stoeckle, 'Corporatization and the Social Transformation of Doctoring' (1988) 18 International Journal of Health Services 191.

55 Evetts (n 50) 4.

patient: because they mediate the doctor-patient relationship by making explicit what advice the doctor relies on (whether the patient is aware of this is another question), and introducing an external source of knowledge parties rely on. Furthermore, guidelines relate to the 'confidential knowledge' Evetts refers to: as they make explicit previously 'implicit' knowledge understood without codification by medical practitioners. Through the publishing of guidelines, some of the knowledge healthcare professionals' use has become (more) public and are now open to criticism by external parties. The focus for the enquiry in this study is the possibility of control of professionals (and by extension the doctor-patient encounter) through guidelines. Fournier provides a good assessment of this process in describing professionalism as a disciplinary tool 'which inscribes "autonomous'' professional practice within a network of accountability and governs professional conduct at a distance.'[56]

Guidelines used by NICE or the NHG, irrespective of the differences in institutional context we might observe, both strive for 'professionalization' through guidelines. This aim of course is not described in direct terms using the language of disciplinary tools but rather it is presented as part of broader mission of improving the quality of care. One element that is likely to play a role, based on this literature, in the deliberations of guideline makers is what 'image' they have of their audience (healthcare professionals) and the fact

---

[56] Valérie Fournier, 'The Appeal to "Professionalism" as a Disciplinary Mechanism' (1999) 47 Sociological Review 280, 280.

that they will have to respond to what healthcare professionals want out of a guideline. Guidelines in this respect can be more or less ambitious depending on the level of change they demand in the working practices of healthcare professionals. While attention is paid to the micro-politics of guideline production[57] within a guideline group, and research exists on the political consequences of guideline making for other healthcare actors like the government[58], this relationship between professionals in the guideline group and those outside is neglected. There is reason to include this aspect in this study because of the way a guideline group defines their role in the regulation of professionals, through guidelines, directly influences the amount of leeway they find to adopt recommendations that change clinical practice. It is likely, but as yet unclear, that the room for changes in clinical practice correlate with the specific position of the guideline producing organisation in relation to the professional associations of the various medical professionals affected by a guideline. In a more corporatist healthcare system, such as the Dutch system, one could expect less room for clinical change than in a more top-down organised guideline making process in England. So, it is plausible that in a system with more dispersed guideline making, the relationship between the healthcare professionals that have to follow the guidelines and those who draft it are closer than in a system where guidelines are always made by a multidisciplinary group of professionals. In the Netherlands one would thus expect

---

57 Ruth McDonald and Stephen Harrison, 'The Micropolitics of Clinical Guidelines: An Empirical Study' (2004) 32 Policy & Politics 223.

58 Moreira (n 21).

guidelines that are less likely to change medical practice then in England. This study will explore whether these assumptions hold true.

*Transparency as a Structuring Force Influencing Guidelines*

An important element that drives the success of guidelines as regulatory tools is the transparency of the process of developing the guideline. Transparency has several different functions and shapes the development of guidelines in various ways. Understanding how calls for transparency are described in the regulatory literature and the debates surrounding this concept will provide further pointers for an empirical examination of the process of guideline development.

Transparency is closely connected to processes of accountability in modern regulatory states. Evidence-based medicine and tools surrounding its rise, quality assessments, performance indicators, rankings, clinical audits and also guidelines, are all in various ways aimed at creating a greater transparency of professional practice for outside observers.[59] Some of the literature on this topic paints the response of professionals in negative terms and suggests active resistance by professionals to these transparency

[59] Power (n 30); Charlotta Levay and Caroline Waks, 'Professions and the Pursuit of Transparency in Healthcare: Two Cases of Soft Autonomy' (2009) 30 Organization Studies 509.

initiatives.[60] Levay and Waks however shine a light on a more positive engagement of professionals with these transparency initiatives in the Swedish healthcare setting:

> Professional involvement in transparency technologies in the two cases developed despite initial resistance, largely in response to outside pressures. Professionals perceived the need to *legitimize [stress in original]* their work in the face of external demands (…) However, the transparency techniques also seemed to correspond to professionals' own conceptions of how to *develop [stress in original]* quality in modern healthcare organizations. Accreditation and national quality registries could be used both to assure minimal quality levels and standardized procedures and to develop new ways of organizing work.'[61]

Their argument is interesting because they show how by working with these initiatives professionals retain control over the criteria of their own evaluation: 'On the contrary, they still managed to control many of the premises and criteria of evaluation. Professionals were neither independent in evaluating their own work, nor in the hands of external evaluators, but enjoyed a negotiated, "soft" professional autonomy.'[62]

This literature suggests that transparency projects can lead to resistance and a loss of autonomy but at the same time form the basis of strengthening autonomy, within the new boundaries set as a result of the interaction of the state and the profession. The thesis

---

60 Haridimos Tsoukas, 'The Tyranny of Light: The Temptations and the Paradoxes of the Information Society' (1997) 29 Futures 827; Marilyn Strathern, 'The Tyranny of Transparency' (2000) 26 British educational research journal 309.

61 Levay and Waks (n 59) 519.

62 ibid 523.

will explore whether the argument that medical guidelines are used in such a positive way, to establish autonomy is valid, and build on the work of Levay and Waks by exploring transparency in guidelines. Transparency that is specific to medical guidelines has not received attention in the literature so far, and this is a gap which this thesis aims to fill.

How transparent guideline makers are during their drafting may help to provide an explanation for some of the differences in guidelines between countries. For example, in England recent guidelines made by NICE all follow the GRADE-approach to grade evidence.[63] This is a process developed to help guideline makers to consider which studies should be included in the evidence-base of a guideline. This works through a process of developing an assessment of the strength of the certainty that the result of the study is not a coincidental result but represents a casual relation. In other words, this instrument helps make explicit the potential sources of bias in the studies that underpin guidelines. The results of the GRADE analysis are available online and are open to the public: this improves the visibility of the guideline drafting process in the case of England. Some guidelines might not follow this approach which might explain some of the differences we

63 Gordon H Guyatt and others, 'GRADE: An Emerging Consensus on Rating Quality of Evidence and Strength of Recommendations' (2008) 336 BMJ 924; 'Grading Quality of Evidence and Strength of Recommendations' (2004) 328 BMJ 1490.

observe between guidelines. The specific relevance of this literature will be examined in the context of the case studies during the following chapters.

Through the construction of evidence tables, the GRADE-methodology tries to improve the visibility of the relevance and strength of the evidence used to base recommendations on in guidelines. This is an important aspect of the claim that guidelines are evidence-based: it provides for a possibility to check the basis for every recommendation made in a guideline and assess, as a reader, whether you agree with the analysis of the evidence performed by the guideline group. Knaapen places this development in the key of the 'EBM-ness' of guidelines: 'Not (quantitative) Evidence but specific procedural rules are the central feature of 'being EBM', reflecting a new type of objectivity that Cambrosio and colleagues … have called 'regulatory objectivity.'[64] This has been identified as important by the literature because the GRADE approach towards selecting evidence leads to the adoption of evidence that is placed higher in the evidence hierarchy and to the rejection of evidence of lower quality. In this way these guidelines

64 Loes Knaapen, 'Being "Evidence-Based" in the Absence of Evidence: The Management of Non-Evidence in Guideline Development' (2013) 43 Social Studies of Science 681, 701.

might differ from those that do not adopt this approach and use less systematic ways of selecting evidence.

Another example of a transparency tool that has gained prominence in the field of guidelines is the AGREE tool, alongside the updated AGREE II.[65] This is a tool that defines several procedural criteria that guidelines have to meet: making it possible to compare guidelines on different diseases and different countries.[66] The claim that guidelines that score highly on these procedural criteria are also high quality guidelines is questioned within the literature.[67] The AGREE tool promotes transparency of procedures for developing guidelines, as it is online and available for public review.

These attempts at improving the transparency of guidelines could have an indirect effect on guideline makers and explain guideline differences because they help steer a process of evidence selection and define criteria for attaching weight to some evidence (studies with low degrees of bias, and so high confidence in the result). If these criteria are

---

65 'Development and Validation of an International Appraisal Instrument for Assessing the Quality of Clinical Practice Guidelines: The AGREE Project' (2003) 12 Quality and Safety in Health Care 18; Melissa C Brouwers and others, 'AGREE II: Advancing Guideline Development, Reporting and Evaluation in Health Care' (2010) 63 Journal of Clinical Epidemiology 1308.

66 Pablo Alonso-Coello and others, 'The Quality of Clinical Practice Guidelines over the Last Two Decades: A Systematic Review of Guideline Appraisal Studies' (2010) 19 Quality and Safety in Health Care 1.

67 Joseph C Watine and Peter S Bunting, 'Mass Colorectal Cancer Screening: Methodological Quality of Practice Guidelines Is Not Related to Their Content Validity' (2008) 41 Clinical Biochemistry 459.

followed this might lead to different choices in what is defined as ‘evidence’ and as a result lead to different recommendations between guidelines. Plausibly this could also lead to more consistency in selecting evidence in a response to the transparency brought about by these instruments. As a result, this thesis will discuss the role transparency initiatives have on guideline decision-making.

*Making Sense of Guidelines as Regulation*

The regulatory and sociological literature enable the argument that guidelines are a particular form of professional mandated self-regulation that straddle the divide between regulation to control professionals and regulation to control (further) state intervention by controlling the criteria of evaluation. This study will discuss how those two sometimes contradictory rationales play out in practice, and in this way contribute to the literature in both the field of sociology as well as regulation. Furthermore, the role of the transparency initiatives identified in this process will be considered.

## 2.4 Institutional Literature

So far, we have mainly considered guidelines and guideline drafters in isolation, as if guidelines are produced in a vacuum. But we should not neglect the institutional environment in which guideline drafting takes place: we should consider the impact of institutions on guideline drafting. To date the institutional context of guideline drafting has

not been considered explicitly in the main body of medical guideline literature. This represents an important gap in the literature and more research is needed to draw attention to the institutional dimensions of guideline drafting. Institutional understandings offer a more nuanced understanding of the ways in which guidelines differ: allowing us to see the role institutions have themselves in promoting specific guidelines.

This study will focus on the mechanisms structuring institutions and what their relations to actors, within institutions, are as this thesis does not aim to provide a full institutional history of medical regulation. By focussing on the institutional element of guideline drafting it is expected that new insights might be gained about the factors that influence guideline drafting by bringing to the fore elements that have so far been neglected in the discussion of guidelines. The comparative design of this study is particularly suited bring more clarity to this issue by enabling a comparison between relatively similar medical communities, one in the Netherlands and one in England. This section will present three different institutional literatures that are particularly relevant for the study of the development of guidelines: sociological institutionalism, institutional regulation of conflicts of interest, and the institutional role of cost and cost-effectiveness as a factor that influences the development of guidelines.

*Institutional Theory*

The work of guideline groups is central to the development of guidelines. Guideline groups are presented as being independent in their operation, and it is perhaps for that reason that an institutional focus is absent in the literature. While attention is paid to the micro-level of guideline development by Pagliari et al.[68], and to European cooperation between different guideline makers[69], the relationship between guideline making institutions and guideline groups and how they could potentially influence each other often goes unquestioned.

Modern sociological institutionalism[70], the term used by Meyer[71], provides a way of conceptualising organisations and their relation to other actors. In this strand of the organisational literature, actors are both empowered and constricted by the institutional

---

68 C Pagliari, J Grimshaw and M Eccles, 'The Potential Influence of Small Group Processes on Guideline Development' (2001) 7 Journal of evaluation in clinical practice 165.

69 Loes Knaapen, 'European Regulation and Harmonization of Clinical Practice Guidelines', *European Union Public Health Policy: Regional and Global Trends* (Routledge 2013).

70 Paul J DiMaggio and Walter W Powell, 'The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields' (1983) 48 American Sociological Review 147; Walter W Powell and Paul J DiMaggio, *The New Institutionalism in Organizational Analysis* (University of Chicago Press 1991).

71 John W Meyer, 'Reflections on Institutional Theories of Organizations', *The SAGE Handbook of Organizational Institutionalism* (SAGE Publications Ltd 2008).

context of organisations and institutional forces provide rationales for action. Meyer describes this in these terms:

> Here, actors are substantially empowered and controlled by institutional contexts, and these contexts go far beyond a few norms or network structures (…) Institutions, in these conceptions are packages or programs of an expanded sort. 'Regimes' is a term employed in political science for the idea—organizational packages infused with cultural meaning.[72]

Actors in this conception do not act in an unconstrained way out of self-interest[73] but are instead within bounds that constrain certain courses of action and legitimize others. In this realist (or constructivist), perspective actors are not only influenced by their institutional environment but also constructed by it.[74] It is this approach this study will adopt and build on because it allows for a more nuanced understanding of the conflicting forces of actors and institutions.

In exploring the notion of the actor in modern society Meyer and Jepperson provide an interesting account of a specific future of modern actorhood, one that provides a useful

---

72 ibid 791–792.

73 An axiomatic feature of economic models of human agency.

74 This thesis therefore does not seek to engage with the work of Foucault or that of Luhmann, even though the relevance of their thinking for a potential construction of the role of actors in institutions is acknowledged.

description for the purposes of this study.[75] They present the notion of 'agentic actors' that do not only work for, or because of, their immediate interest but also work towards becoming a more effective agent for himself and others.[76] And these actors are highly standardised through a process of isomorphism[77], where organisations become structurally similar and the way actors behave in them conforms to standardised models of individualism. In describing a society of these 'agentic actors' they suggest how collective action and social status is bound up with a pretence for the interests of others, by which they mean the expression of cultural ideals. They describe the role of professionals in these terms:

> The great enhancement of the centrality of "disinterested" professionalized others is a notable development of modern society, one difficult to account for in terms of interests or functional requirements. (Indeed demonstrations of functionality are conspicuous for their absence.) The definers and carriers of social agency—especially those more removed from direct responsibility for mundane self-interested actorhood—are granted substantial authority and prominence. The most canonical authority is that of the professional who serves no actor but rather the high truths of the rationalized natural and moral universe: pure other-hood (…).[78]

---

75 John W Meyer and Ronald L Jepperson, 'The "Actors" of Modern Society: The Cultural Construction of Social Agency' (2000) 18 Sociological Theory 100.

76 ibid 107.

77 DiMaggio and Powell (n 70).

78 Meyer and Jepperson (n 75) 115.

Restating their argument more simply: they describe that for individuals to be accorded a high social status in society, they are required at least putatively to describe their endeavours to be directed towards the achievement of cultural goals. This provides interesting new insights for conceptualising the process of developing medical guidelines. This literature suggests that actors might be self-interested and act to further these interests but will be confronted by the necessity of the institutional scripts that are connected to doing this activity have to present themselves as being disinterested, neutral, enactors of scientific knowledge[79] to further the interests of medicine.

This study will use this sociological institutionalist approach to understand and interpret the experiences of collectively producing medical guidelines, while at the same time accounting for the cultural scripts produced and maintained by the guideline making organisations of which they form a part. In trying to see how the development of guidelines might be a contextualised and socially constructed process of collective organisation this theory will be useful.

*Institutional Regulation of Conflicts of Interest*

While sociological institutionalism provided a general approach to conceptualising the role of institutions in the development of guidelines for the purposes of this study a

[79] The 'evidence' as described in detail earlier.

more detailed analysis of the institutional regulation of conflicts of interest is particularly relevant. The prevailing medical norms (or scripts in the institutional literature) mandate that (medical) organisations have a policy to deal with (actual or perceived) conflicts of interest. Here we can observe a form of institutional isomorphism, as described earlier, in so far as the NHG and NICE adopt very similar institutional procedures and policies in this area.

Potential conflicts of interest have led to an institutional response in medical regulation: both NICE and the NHG recognise that conflicts of interest might occur and require disclosure. The NICE guideline manual[80] and the NHG 'procedurebook'[81] require both financial as well as non-financial conflicts of interest to be reported on disclosure forms before the start of the making of a guideline, and these declarations have to be kept up to date. The NHG procedure book states that if there is doubt about the independence of a guideline group member he/she can be excluded from the guideline group.[82] The NICE policy on conflicts of interest states that if there is a specific financial interest the person cannot attend the meeting the interest relates to, and this might also be the case for specific

---

80 'The Guidelines Manual' <http://publications.nice.org.uk/the-guidelines-manual-pmg6> accessed 26 May 2014.

81 NHG,'Procedureboek' <https://www.nhg.org/sites/default/files/content/nhg_org/uploads/procedureboek_versie_090810.pdf> accessed 17 November 2014.

82 ibid 15.

non-financial interest. All other types of non-specific interest need to be declared but do not preclude participation.[83] In this way both institutions try to find a balance between including the experts who might have conflicts of interest and profiting from their expertise while avoiding a biased guideline. Both institutions therefore provide the option that someone with a conflict of interest can participate, but prevents that person from influencing decisions where there is a direct financial conflict of interest.

The conflicts of interest literature lays the groundwork for this study to be able to analyse whether these policies are adhered only on paper or also in practice, and whether we are able to discern any association between a conflict of interest and the recommendations in a guideline. The underlying assumption of much of this literature is that without such institutional controls on participation, guideline groups will draft guidelines that are biased towards the pharmaceutical industry. This study will endeavour as far as possible to uncover whether this can be said to be the case, or whether this institutional regulation provides an effective bulwark against such undue influence.

---

83 NICE, 'Code of Practice for Declaring and Managing Conflicts of Interest' 7 <https://www.nice.org.uk/Media/Default/About/Who-we-are/Policies-and-procedures/code-of-practice-for-declaring-and-managing-conflicts-of-interest.pdf> accessed 20 June 2016.

The concept of 'regulatory capture' expresses the idea that guideline making bodies could be biased on an institutional level.[84] Regulatory capture literature suggests that NICE and the NHG could be 'captured' by the pharmaceutical companies to produce and promote guidelines that favour their products, in contrast to generic medication made by competitors, to give an example. Regulators would then no longer work in the 'public interest' making evidence-based guidelines to increase the quality of medicine, but instead work to further the private interests and profits of companies.

The fact that NICE and NHG have policies on paper aimed at dealing with individual conflicts of interest of guideline group members, does not preclude this form of bias. Direct funding of NICE and NHG would be one way in which regulatory capture could occur. There are, however, no direct financial links between the industry and the guideline making bodies, as both bodies are funded by the state or through a combination of private fees by professionals and state funding. Another way in which these bodies could be 'captured' is termed 'cognitive capture'[85]: where guideline making institutions start to identify (too) strongly with the interests of industry instead of keeping a professional distance. This might be a factor that influences guidelines and could perhaps explain differences in guidelines between countries. The role of pharmaceutical companies in the

---

84 George J Stigler, 'The Theory of Economic Regulation' (1971) 2 The Bell Journal of Economics and Management Science 3.

85 Daniel P Carpenter and David A Moss (eds), *Preventing Regulatory Capture: Special Interest Influence and How to Limit It* (Cambridge University Press 2014).

construction of guidelines, and their contacts with NICE and NHG, will be analysed in this thesis to assess the strength of this factor.

*The Role of Costs in Medical Guidelines*

It is important to make the role of the costs of medical treatment more explicit and highlight the influence this has on building consensus in a guideline group. In order to understand how this issue can steer the evaluation of evidence and how institutional factors influence the importance of cost considerations.

An extensive literature that evaluates cost in healthcare exists, mainly emerging from the fields of economics and health policy. A strand of this literature is concerned with designing and validating economic models to use in cost-benefit analysis[86], while others

[86] Joan Costa-Font, Christophe Courbage and Alistair McGuire, *The Economics of New Health Technologies* (Oxford University Press 2009) <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199550685.001.0001/acprof-9780199550685> accessed 1 May 2015; Karl Claxton and others, 'Discounting and Cost-Effectiveness in NICE – Stepping Back to Sort out a Confusion' (2006) 15 Health Economics 1; Andrew J Palmer and others, 'The CORE Diabetes Model: Projecting Long-Term Clinical Outcomes, Costs and Cost-Effectiveness of Interventions in Diabetes Mellitus (Types 1 and 2) to Support Clinical and Reimbursement Decision-Making' (2004) 20 Suppl 1 Current Medical Research and Opinion S5.

apply these models in systematic reviews of the cost-effectiveness of specific treatments.[87] The use of these cost-effectiveness and cost-benefit analyses techniques has been most prominent in the growing field of health technology assessment.[88]

To give a sense of the scale of this literature, for one disease (diabetes) in one journal (health technology assessment) there are already 44 health technology assessment reports published.[89] Health technology assessments are used in evaluating whether new drugs or technologies should be covered by the healthcare system, and decision-makers rely partly on these economic models to assist decision-making. In England, the literature has critically engaged with the cost-effectiveness thresholds NICE uses in its decision-

---

87 E Loveman and others, 'The Clinical and Cost-Effectiveness of Patient Education Models for Diabetes: A Systematic Review and Economic Evaluation' (2003) 7 Health Technology Assessment <http://www.journalslibrary.nihr.ac.uk/hta/volume-7/issue-22> accessed 29 April 2016; 'ZonMw - A Randomised Cost-Effectiveness Study on the Surgical versus Conservative Treatment of Sciatica -' <http://www.zonmw.nl/nl/projecten/project-detail/a-randomised-cost-effectiveness-study-on-the-surgical-versus-conservative-treatment-of-sciatica/hoe-nu-verder/> accessed 19 June 2014.

88 HTA Commissioning Board, 'Standing Group on Health Technology' (1998) 2 Health Technology Assessment <https://www.researchgate.net/profile/David_Greatbatch/publication/13370611_Qualitative_Research_Methods_in_Health_Technology_Assessment_A_Review_of_the_Literature/links/0912f50b7847ab963e000000.pdf> accessed 20 May 2016; Pascale Lehoux and Stuart Blume, 'Technology Assessment and the Sociopolitics of Health Technologies' (2000) 25 Journal of health politics, policy and law 1083; Patricia W Stone, Christine R Curran and Suzanne Bakken, 'Economic Evidence for Evidence-Based Practice' (2002) 34 Journal of Nursing Scholarship 277; David Banta, 'The Development of Health Technology Assessment' (2003) 63 Health Policy 121; Carl May, 'Mobilising Modern Facts: Health Technology Assessment and the Politics of Evidence' (2006) 28 Sociology of Health & Illness 513.

89 https://www-journalslibrary-nihr-ac-uk./collections/diabetes/

making.[90] The literature suggests that within NICE cost-effectiveness has a central place in the drafting of guidelines through the use of in health technology assessments. These assessments use an economic threshold above which they will usually not recommend a technology or drug. The expression of this threshold is in ICER's, incremental cost-effectiveness ratio, per quality-adjusted life year (QALY) gained. This threshold is set at £20,000-£30,000 per QALY gained.[91] NICE is careful to say that cost-effectiveness is not the only reason for treatment selection, and that it also funds technologies or drugs with higher costs.

While in the field of health technology assessment cost and cost-effectiveness are analysed and extensively debated this has not been true of the use of these methods within guidelines. This is all the more surprising because cost-effectiveness modelling is such an important rationale, at least in England, that influences not only health technology assessments but also guidelines through the reference to cost-effectiveness in the NICE

---

90 'NICE Threshold - Centre for Health Economics, The University of York' <http://www.york.ac.uk/che/research/teehta/thresholds/> accessed 25 February 2015; Nancy Devlin and David Parkin, 'Does NICE Have a Cost-Effectiveness Threshold and What Other Factors Influence Its Decisions? A Binary Choice Analysis' (2004) 13 Health Economics 437; Karl Claxton and others, *Methods for the Estimation of the NICE Cost Effectiveness Threshold* (University of York, Centre for Health Economics 2013) <http://www.york.ac.uk/media/che/documents/reports/resubmitted_report.pdf> accessed 15 August 2016.

91 'Guide to the Methods of Technology Appraisal 2013 |' para 6 <https://www.nice.org.uk/article/PMG9/> accessed 4 July 2016.

guideline manual.[92] It is likely that the effect on drafting a guideline of cost-effectiveness considerations can be considerable. Cost-effectiveness considerations require guideline drafting groups to consider (complex) economic modelling of costs and benefits of particular technologies and drugs, and a sophisticated weighing exercise of cost and benefits against clinical values, patient preferences and other considerations. This study attempts to partly fill this gap by evaluating how cost-effectiveness models and considerations feature, if at all, in the development of guidelines and how, and to what extent, this is reflected in the guidelines produced. This element of guideline development might also allow for an interesting comparison with the Netherlands because cost-effectiveness considerations are unlikely to play the same role there. While the NHG recognises costs as a factor, and recommends prescribing the cheapest medication if the effectiveness is the same[93], no methodology to integrate cost considerations in guidelines is in use there. This can be explained by the fact that in the Netherlands the health technology assessments are performed by the Zorginstituut Netherlands, an advisory body to the Secretary of State for Health, Social Care and Sport.[94] The Dutch NHG has no direct

---

92 'Developing NICE Guidelines: The Manual' ch 7 <https://www.nice.org.uk/process/pmg20/> accessed 15 August 2016.

93 NHG (n 81).

94 Zorginstituut Nederland, 'Kosteneffectiviteit in de Praktijk' <https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/rapporten-en-standpunten/2015/1506-kosteneffectiviteit-in-de-praktijk/1506-kosteneffectiviteit-in-de-praktijk/Kosteneffectiviteit+in+de+praktijk.pdf> accessed 16 April 2016.

involvement with health technology assessment and unlike the situation in England there is no relation between a recommendation in a guideline and the funding of treatment, as these decisions are taken by two separate and independent bodies. Given these differences the question is whether cost considerations force English guideline makers to make decisions to recommend cheaper medication or technologies purely on the basis of cost, whilst the Dutch guideline makers are free to prescribe more expensive medication or technologies regardless of the cost? This study will seek to empirically analyse these assumptions through a comparative assessment of the role of cost in Dutch and English guidelines.

The way that the issue of cost of treatment of disease in guidelines is handled could this have a real influence on the development of guidelines? The institutional environment of guideline making in England forces guideline groups to consider cost considerations. In the Netherlands guideline groups do not have to confront the issue of cost in this way because these decisions are taken by the Zorginstituut, a separate and independent body.

## 2.5 Factors Shaping Group Decision-Making

To analyse the development of guidelines from a socio-legal perspective, the consideration of a range of factors that shape group-decision making processes in constructing medical guidelines is important. At the outset it is essential to state that this study does not purport to give an ethnographic or sociological account of the exact nature of the discussion in

guideline groups. The research does not strive to establish precise account of the discussion of the guideline group members, rather the focus lies on the results these discussions have had on the text of the guideline documents. To this end this section will analyse several literatures that can provide potential factors influence the decision-making in guideline groups and the subsequent text of the guideline that is constructed as a result of this process.

Guidelines are a product of consensus within a group of professionals. Naturally the process of reaching consensus will influence the guideline. Social psychology literature on group-processes paints a picture of the pitfalls that exist in this process. Status hierarchies that exist outside of the group might be reproduced within the guideline group[95]; minorities and women might be underrepresented in the discussion[96], the group might suffer from the social influence from one particular individual to the detriment of the collective decision-making[97], there might be majority pressure to conform[98], as well as outcomes suffering from 'groupthink'.[99]

---

95 JE Driskell and B Mullen, 'Status, Expectations, and Behavior: A Meta-Analytic Review and Test of the Theory' (1990) 16 Personality and Social Psychology Bulletin 541.

96 Alice H Eagly, 'Gender and Social Influence: A Social Psychological Analysis.' (1983) 38 American Psychologist 971.

97 S Milgram, 'Some Conditions of Obedience and Disobedience to Authority' (1965) 18 Human Relations 57.

98 Solomon E Asch, 'Opinions and Social Pressure' (1955) 193 Scientific American 31.

99 Irving L Janis, *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes*, vol viii (Houghton Mifflin 1972).

The potential influence of group dynamics on guidelines has not gone unnoticed in the literature. Several authors have discussed these factors in some detail. Pagliari and colleagues[100] have observed that status hierarchies are indeed to some extent reproduced in guideline groups, through the fact that there is more discussion among high status individuals in guideline groups. Berg et al have drawn attention to the fact that not all group members contribute equally.[101] While Hopthrow et al have highlighted four areas that influence the functioning of guideline groups and suggested ways to make sure groups are effective, these areas are: '(1) information sharing, (2) systematic processing, (3) group formation, and (4) process loss.'[102] For all these areas, Hopthrow et al provide suggestions on how to improve the process. Examples they give include: a diversity of opinions and expertise, providing group socialisation activities to enhance bonding, promote group norms that provide for critical thinking, sufficient time and matching people to specific (sub-) tasks.[103]

---

100 Pagliari, Grimshaw and Eccles (n 68); Claudia Pagliari and Jeremy Grimshaw, 'Impact of Group Structure and Process on Multidisciplinary Evidence-Based Guideline Development: An Observational Study' (2002) 8 Journal of Evaluation in Clinical Practice 145.

101 Marc Berg, Ruud ter Meulen and Masja Van den Burg, 'Guidelines for Appropriate Care: The Importance of Empirical Normative Analysis' (2001) 9 Health Care Analysis 77.

102 Tim Hopthrow, Gene Feder and Susan Michie, 'The Role of Group Decision Making Processes in the Creation of Clinical Guidelines' (2011) 23 International Review of Psychiatry 358, 359–360.

103 ibid 360–364.

If we consider the guideline working groups in the Netherlands there is a study by Zuiderent-Jerak and others on the variation between guidelines.[104] In this study the authors conduct interviews with different guideline developers that shed a light on some parts of the process of guideline development. Some guideline groups have patient involvement while others do not[105], there are sometimes tensions between evidence-based and consensus-based approaches within a working group, and monetary constraints play a role on deciding the scope of a guideline.[106]

Considering the guideline groups in England, Atkins et al paint a detailed picture in their study of three working groups of NICE.[107] The authors establish that some participants see guideline recommendations as aspirational statements while others prefer directly implementable recommendations.[108] This study lends more substance to the claim the professional expertise and status influence group discussions, where higher status

---

104 Teun Zuiderent-Jerak and others, 'Variatie in richtlijnen. Wat is het probleem?' (*Variation in Guidelines. What is the problem? IBMG*, 2011) <http://www.pgosupport.nl/mailings/FILES/plugin_content/176/Regieraad_rapport_variatie_in_richtlijnen.pdf> accessed 12 November 2013.

105 ibid.

106 ibid.

107 Lou Atkins and others, 'The Process of Developing Evidence-Based Guidance in Medicine and Public Health: A Qualitative Study of Views from the Inside' (2013) 8 Implementation Science 101.

108 ibid 106.

professionals have a greater voice, as well as the importance of vested interests of health professionals (in a certain procedure or treatment) in influencing recommendations.[109]

One message that emerges clearly is that the group process involved in making a guideline is a complex one: a process requiring very specific tasks, which necessitates a large amount of expertise and has to be completed within a limited timeframe.

In this study we can build further on this literature by establishing whether the differences that exist in the group composition, and the way that the groups function, does indeed lead to differences in assessment and differences in decision-making and whether the factors that Hopthrow et al. suggests play a role. It is not immediately clear whether the factors Hopthrow et al. identifies will always be present in the same way in different guideline groups. These factors could be present, but not to such an extent that they explain differences in guidelines. This study will use a comparative approach to gain insight on this question.

While this literature gives a general sense of some of the difficulties that members of guideline groups might face during the group processes of developing a guideline, a

---

109 Atkins and others (n 107); Rosalind Raine and others, 'An Experimental Study of Determinants of Group Judgments in Clinical Guideline Development' 364 The Lancet 429; Andrew Hutchings and Rosalind Raine, 'A Systematic Review of Factors Affecting the Judgments Produced by Formal Consensus Development Methods in Health Care' (2006) 11 Journal of Health Services Research & Policy 172; Pagliari and Grimshaw (n 100).

more specific account given some of the political and normative elements of guideline drafting are also of interest to this study. Understanding potential political and normative considerations in guideline drafting will help uncover potential motivations of guideline group members and help shed light on the process that feeds that leads to a guideline text. These literatures also provide arguments to question the orthodox view of evidence-based guidelines as 'the application of evidence to decision-making' and suggest ways in which that picture is more complicated than some of the medical literature allows for.

First, we will consider the politics inherent in evidence-based medicine itself, following the critique expressed in philosophy of science literature discussed earlier. To be able to apply this knowledge to a study of participants developing guidelines we then turn to the sociology of situated judgement to provide insights in the form of repertoires of action that members of guideline groups might employ.

*Medical Politics and Evidence*

The study of medical politics, and the role of medical guidelines, is predicated upon an understanding of the link between knowledge and power in healthcare. The possession of scientific knowledge is the key to power in healthcare. Belkin describes this link as the

'technocratic wish'.[110] In his treatment of managed care he suggests that there is an intimate relation between the scientific justifications given for actions and shifts in power within healthcare. So the shift in evidence-based medicine to favour generalizable data about populations, in the form of RCTs, over uncontrolled studies or clinical experience, is the result of contingent coalition of interests that shape the success of this claim for truthfulness over other competing claims.[111] The way that managed care, and evidence-based medicine by extension, is successful is in its claim to 'scientific obviousness'[112] and as a result is accepted as right and correct, thereby this hides the essentially cultural nature of the choices made to arrive at these scientific claims. Scientific knowledge claims are therefore socially constructed and contestable.

Rodwin observes how evidence-based medicine neatly transposes questions of ideology, values and standards, into technical questions of evidence. He draws a parallel with the progressive era in 1890's American politics that strove to replace 'governance by

---

110 Gary S Belkin, 'The Technocratic Wish: Making Sense and Finding Power in the "Managed" Medical Marketplace' (1997) 22 Journal of Health Politics, Policy and Law 509.

111 ibid 511.

112 ibid 512.

politics with governance by experts' and describes this as a search for a 'technocratic fix.'[113] Goldenberg following Rodwin formulates this succinctly:

> The configuration of policy considerations and clinical standards into questions of evidence conveniently transform normative questions into technical ones. Political issues are not resolved, however, but merely disguised in technocratic consideration and language. Thus the goals of medicine and other normative considerations lie just below the surface of these evidentiary questions, and evidence becomes an instrument of, rather than a substitute for, politics.[114]

This literature suggests that evidence-based medicine and evidence-based guidelines can be political but still claim to be objective and scientific because the political questions are masked as technical questions that only require medical expertise. Understanding how evidence-based medicine produces objectivity claims will help us understand how guideline group members are likely to understand their own work, and how they relate their work to medicine as a whole.

While regulatory objectivity stresses the guideline as an instrument and looks at the guideline as a whole, work by Berg et al inverts this perspective and looks at the normativity inside a guideline, and choices that are made by the participants.[115]

---

113 Marc A Rodwin, 'Commentary: The Politics of Evidence-Based Medicine' (2001) 26 Journal of Health Politics, Policy and Law 439, 441.

114 Goldenberg (n 17) 2630.

115 Berg, ter Meulen and Van den Burg (n 101).

Normativity occurs where values are assigned that are inherently 'good' or 'bad' (or other social values) to a choice: this choice is then dictated by those values based on what is perceived to be 'normal' or correct. Berg et al provide a classification of types of normativity in guidelines that are useful in this regard: 'I: The normativity implied in the instrument itself, II: The assumptions and translations that have to be made each time the instrument is developed for a specific problem, III: The assumptions and translations that are specific to the medical situation addressed by the instrument.'[116]

The second type of normativity here is particularly relevant for this study of guidelines. Normative choices are often judgements that involve weighing costs against effects – and making choices on which costs count, and which do not.[117] When applying this to evidence, normativity can be seen in the choices made around the following questions: does this study count as evidence, is the result of the study relevant to my population of patients, and should the study lead to a change in recommendation? The use of evidence is thus an activity that is imbued with a series of more and less normative choices, which provides a filter to achieve the link between medical knowledge and

116 ibid 80.

117 ibid 81.

recommendations in a guideline. The choice to include one medical study over another, even if driven by seemingly objective criteria, remains a normative choice.

To be able to link elements of normativity more directly to the activity of members of the guideline group we turn now to the sociology of situated judgement [118]: an approach that stresses the social nature of coordinated action by individuals. The 'sociology of situated judgement' refers to sociological and evaluative studies of how actors reach consensus through discursive methods in a discrete social situation. Moreira has applied this approach especially to show how guideline groups deal with evidence based on his in-depth observation of several guideline groups. His approach is valuable for this study as it draws together wider concerns about normativity of guidelines in an abstract sense, with the practical knowledge of observation of social action. Furthermore, he is one of the few authors who base their findings in detailed ethnographic observations of guideline group meetings, this makes his work particularly valuable. He distinguishes four repertoires of evaluation, drawing upon theories of situated judgement.

Moreira's repertoire of evaluation is 'a recurrent form of judgement deployed by social actors in considering the worth of an object (…) as means of harmonizing their

---

[118] Luc Boltanski and Laurent Thévenot, 'The Reality of Moral Expectations: A Sociology of Situated Judgement' (2000) 3 Philosophical Explorations 208.

actions with others.'[119] As Moreira explains, these types of judgement employed by members of guideline groups serve the specific purpose of creating meaning through the use of (medical) knowledge as a mediation of social action.[120] One of these 'recurrent forms of judgement' is called 'Science' and forms the first 'repertoire of evaluation'. Moreira describes this repertoire as one where guideline group members are evaluating the technical robustness of evidence in light of possible criticisms of health researchers, specialists and other guideline groups.[121]

This accords well with our discussion of evidence-based medicine and its attempts to enhance the use of appropriate medical knowledge, such as clinical trials, to influence clinical medicine. However, what Moreira shows convincingly is that this is only part of what happens in a guideline group when clinicians discuss evidence. Considering research that is methodologically sound and satisfies the requirements of evidence-based medicine is not enough: the next step is to analyse this evidence in light of the argumentative forum it is going to be used in. This might affect whether a change that is based on a technical assessment of scientific studies is translated into a recommendation to change practice.

---

119 Moreira (n 21) 1977.

120 ibid.

121 ibid 1978.

Using this sociological approach has the advantage of not stating what guidelines ought to be but rather analysing how they are made.

The second repertoire Moreira discusses is that of practice. In his words: 'where members are concerned with the practical usability of the knowledge in debate in view of routine patient care.'[122] This repertoire stresses the need for the guideline to be useful in the direct clinical encounter between health professional and patient. Moreira explains that guideline group members have discussions involving reference to experiences of clinical practice, where one tries to square the 'typified health knowledge' embodied in statistical relations that come out of scientific studies in various forms and sizes and the uniqueness of the individual encounter between patient and clinician. This repertoire provides an important addition to the debate as it draws attention to the necessary tension that is inherent in the drafting of a guideline. When drafting a general guideline, one tries to give general guidance which necessarily cannot cover every individual application of that guidance. This sets well with the findings on the implementation of evidence-based guidelines and underscores the contested nature of integrating biomedical science with clinical knowledge. Moreira's presents two further repertoires, politics and process, which will be discussed in the following sections. What we can draw from his repertoires of

122 ibid.

science and practice is that evidence only becomes evidence through the form of a social debate and iterative discourse which includes perspectives outside evidence alone.

Moreira has also commented on the professional aspect of guideline production, styling it as the repertoire of 'politics'. In Moreira's words: 'where members are concerned with the acceptability of their recommendation in relation to public health policies and stakeholder views on the condition in question.'[123] This repertoire highlights that recommendations that come out of guidelines that recommend certain treatments or medication will influence an existing balance of power within a (sub)discipline, and so the existing configuration of responsibilities and arrangement in a given (sub)discipline that the guideline affects. But wider interests of pharmaceutical industry or even interests of government policy might be at stake. As Moreira rightly argues, it is the need to be seen to act as a guideline group and make recommendations, which needs to be balanced with the acceptability of a political position and the need to achieve consensus within the group.[124] This accords well with our wider discussion about guidelines as self-regulation and the role of the sociology of professions. While Moreira seems to stress the political role of guidelines in wider institutional fora, such as Royal Colleges and Medical Associations and Industry Bodies, he only implicitly discusses the professionalization consequences of guidelines where recommendations affect individual health professionals. This distinction

123 ibid 1981.

124 ibid 1982.

is not very material, however, because in practice it is likely that discontent with guidelines will often be expressed through the mediation of a Medical Association or Royal College due to (perceived) political choices.

The final relevant repertoire that Moreira distinguishes is that of process. 'Process' here speaks to the means of constructing the methodological reliability of knowledge claims which emerge from the guideline group discussions and are represented in the text of the guideline. As Moreira explains this involves separating the group from the 'outside' and construct the group as a separate social world on itself. Making sure there are no conflicts of interest is given as an example of how this can be achieved. A further element according to Moreira is to make sure that the evaluation of the evidence is not distorted by the social ties that expert members of the group might have.[125] It shows an interesting ambiguity, in that concerns regarding people's interest outside the group might influence their decision in the group is coupled with the effort to establish the group as an independent institution with its own group identity. This independent identity would preclude external influences from intervening in the debates within the group. The process repertoire suggests a new way of assessing the influences of conflicts of interest depending on the particular group dynamics: by linking the internal group dynamics to the external

---

125 ibid 1983.

interests of the group. This is a potentially fruitful way of discussing conflicts of interest that will be pursued further in the case-studies.

Giving such a lengthy exposition on the work of Moreira is justified as his work provides a multitude of good conceptual tools to help strengthening the empirical approach adopted in this study. This thesis hopes to go further however and take a broader approach to understanding the development of guidelines as a complex process where understanding can be strengthened through a combination of various approaches.

It is clear that the politics of evidence-based medicine are important rationales influencing the conception of guidelines. The evidence-hierarchy has been influential and currently, and at least formally, underpins the evidence gathering and appraisal for making evidence-based guidelines. It is, however, not an exclusive epistemological foundation of medical knowledge, therefore necessarily other types of knowledge claims are used in the production of guidelines. Likewise, knowledge claims in guideline production will always be normative, if only in the limited sense envisioned by Berg.[126] How evidence is used and conceptualised, and to what extent reference is made to the evidence-based hierarchy of evidence, will be the subject of this research. What is not well understood is how exactly the interplay between the formal evidence-based medicine logic, medical politics, and other considerations in a guideline group will play out, or to what extent evidence-based

---

126 Berg, ter Meulen and Van den Burg (n 101).

medicine provides an overriding logic for the consideration of evidence. The sociology of situated judgement might provide the tools to analyse these questions, but in this research further attention will be paid to the concrete ways in which the different 'repertoires' might interact and collide with each other. This literature has also hinted on the importance of considering conflicts of interest as an important factor that could influence the negotiations in guideline groups and thus could shape the text of a guideline, it is this literature that will be explored below.

*Conflicts of Interest and Ties between the Pharmaceutical Industry and Clinicians*

There is a large literature that deals with various conflicts of interest between the medical profession and the pharmaceutical industry.[127] Conflicts of interest are viewed as problematic within medicine because of the fear that commercial interests will dictate the availability of medication and priorities in research: this clashes with the view that healthcare is a public good delivered for the benefit of all. Avoiding undue influence is particularly relevant for medical guidelines as the strength of their advice would be perceived as weaker if this advice was seen as biased, and not driven by scientific

127 Committee on Conflict of Interest in Medical Research, Education and Institute of Medicine (U.S.), *Conflict of Interest in Medical Research, Education, and Practice* (Bernard Lo and Marilyn J Field eds, National Academies Press 2009); Marc A Rodwin, 'Conflicts of Interest, Institutional Corruption, and Pharma: An Agenda for Reform' (2012) 40 The Journal of Law, Medicine & Ethics 511; Robert Steinbrook and Bernard Lo, 'Medical Journals and Conflicts of Interest' (2012) 40 The Journal of Law, Medicine & Ethics 488.

knowledge but by other commercial agendas. Likewise, guidelines that are based on a conflict of interest would be seen to benefit those parties that influenced the guideline and this may provide concessions the pharmaceutical industry.[128] To establish whether conflicts of interest have a role to play in guideline construction we first need to understand the literature on the scope of conflict of interest, and what kind of regulation exists to manage conflicts of interests. Furthermore, this review highlights the possibility of 'regulatory capture'. Instances of such capture would mean that the guideline institution itself, such as NICE and NHG, would work at the behest of outside interests and no would longer be impartial.

Goldacre has provided a comprehensive overview of some of the problems in medicine and presents some of the ways in which a conflict of interest might occur, why it matters, and ways to respond to it.[129] Goldacre identifies that studies by the pharmaceutical companies have a higher likelihood of being published, while negative results are far less likely to be published, leading to publication bias. He goes on to identify several ways in which pharmaceutical companies influence clinicians: drug reps from pharmaceutical

---

128 EAM Gale, 'Conflicts of Interest in Guideline Panel Members' (2011) 343 BMJ d5728; Susan L Norris and others, 'Conflict of Interest in Clinical Practice Guideline Development: A Systematic Review' (2011) 6 PLoS ONE <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3198464/> accessed 23 June 2015.

129 Ben Goldacre, *Bad Pharma: How Drug Companies Mislead Doctors and Harm Patients* (Fourth Estate 2012); Ben Goldacre and Carl Heneghan, 'How Medicine Is Broken, and How We Can Fix It' (2015) 350 The BMJ h3397.

companies visit clinicians to promote their drugs, there are industry funded education courses on new drugs, industry funded congresses, clinicians get paid consultancy positions with pharmaceutical companies, ghost-writers are employed to write publications for academics -and the list goes on.

The question of course is whether these financial ties between practitioners and industry have been shown to affect guidelines. There is some evidence that this has indeed been the case. Several studies have examined guidelines on the same medical condition and found a correlation between more positive recommendations and the number of conflicts of interest present in the guideline group.[130] One obvious solution to this problem, presented by some of the literature, might be to categorically exclude people with a (financial) conflict of interest from taking any part in developing a guideline. Other authors, however, point to the fact that this might lead to guidelines that are not based on sufficient expertise of the condition at hand.[131] This debate has led many to the conclusion that the

---

130 James N George, Sara K Vesely and Steven H Woolf, 'Conflicts of Interest and Clinical Recommendations Comparison of Two Concurrent Clinical Practice Guidelines for Primary Immune Thrombocytopenia Developed by Different Methods' (2014) 29 American Journal of Medical Quality 53; Gisela Schott and others, 'Does the Pharmaceutical Industry Influence Guidelines?: Two Examples from Germany' (2013) 110 Deutsches Ärzteblatt International 575; Norris and others (n 128).

131 Derek J Jones and others, 'Conflicts of Interest Ethics: Silencing Expertise in the Development of International Clinical Practice Guidelines' (2012) 156 Annals of Internal Medicine 809; Caroline Price, 'NICE Risks Making Itself a "Laughing Stock" over Guidance on Metformin Alternatives, Say Experts' (*Pulse Today*) <http://www.pulsetoday.co.uk/clinical/diabetes/nice-risks-making-itself-a-laughing-stock-over-guidance-on-metformin-alternatives-say-experts/20009139.article> accessed 5 May 2015.

best way forward is to ‘manage’ conflicts of interest: by allowing participation but restricting involvement at certain moments during guideline development. Different organisations have developed different rules along these lines: some are stricter than others in how they deal with conflicts of interest.

Through an analysis of the factors shaping the decision-making in guideline groups, conflicts of interest, medical politics and group dynamics, we can complement our analysis of the development of guidelines. By linking these micro-level concerns to the text of the guideline it will be possible to trace linkages between aspects of the development of guidelines and the text of the final guideline product. This study will seek to engage with these literatures in order to use them as a basis for developing a more nuanced comparative understanding of medical guidelines.

## 2.6 Conclusion: Relevance of Literature to this Thesis

In this concluding section of this chapter, the various strands of literature that have been discussed will be pulled together and the main issues have been explored to arrive at a firm position for this research within the academic literature- before analysing the context of this research in more detail in chapter 3. The analysis of the literature has covered four main areas: evidence, regulation, institutions and factors affecting decision-making.

Of central concern in this thesis is the concept of evidence. This is used in the specific technical sense of the use of medical knowledge contained in scientific studies which are ranked according to a hierarchy of evidence. Philosophy of science, especially the work of Chin[132] and Goldenberg[133] has illuminated how evidence-based medicine has both methodological and epistemological limits. These critiques will be taken seriously and this thesis will build further on their findings by providing a robust empirical account of the practical understanding and use of evidence by guideline groups. Science and technology studies have examined evidence and place it at the heart of the attempt (through guidelines) to create objectivity[134]. This literature opens up objectivity as a concept to show how it is purposefully socially constructed. An empirical assessment of the strength of these arguments is currently lacking and this thesis will attempt to address this gap in the literature through a comparative study of guidelines. A sole focus on evidence would, however, risk adopting the viewpoint of the medical literature in this thesis. This might limit the opportunity to criticise the underlying assumptions of evidence-based medicine, while asking searching questions about its role in society.

---

132 Chin-Yee (n 17).

133 Goldenberg (n 17); Maya J Goldenberg, 'Iconoclast or Creed?: Objectivism, Pragmatism, and the Hierarchy of Evidence' (2009) 52 Perspectives in Biology and Medicine 168.

134 Cambrosio and others (n 24).

Analysing medical guidelines as regulatory tools to some extent addresses these concerns and at the same time purports to add to the extensive regulatory literature on the nature and role of (mandated) self-regulation in the current governance of the state.[135] This thesis takes the position that guidelines are a specific form of regulation that both regulates professionals and at the same time tries to ward off the encroaching state by fulfilling a mandate for self-regulation. This thesis will break new ground in putting this duality of purpose of guidelines centre stage.

Literature from the sociology of professions, especially Evetts[136] helped conceptualise the duality of guidelines. A critical engagement with some of their earlier findings on the nature of banking regulation, international private regulatory actors like ISO and the Basel group will enrich this position further. It is thought that medical guidelines might have unique features that have not been addressed in the literature before. The specific and particular role of transparency initiatives in shaping the development of guidelines has the potential to add interesting new elements to this field. Guidelines are also used as regulatory tools of persuasion in relation to medical professionals that as Evetts suggests might in some circumstances be instruments of control. This study will seek the views of guideline drafters and hopes to ascertain how they see the relationship between a

---

135 Ogus (n 39); Black, 'Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes' (n 44).

136 Evetts (n 50).

guideline and the wider medical profession, and in this way, adds to and builds on Evetts work.

These combined perspectives on the role of evidence and guidelines as specific regulatory tools will be further complemented by an understanding of the institutional literature. This area has been neglected so far in the study of guidelines, and this thesis hopes to make a contribution to this field by exploring in depth aspects of the institutional context shaping both guideline drafters and the text of guidelines in different countries. This study adopts a sociological institutionalism approach, championed by Meyer.[137] It is this approach that allows for a nuanced understanding of the role of actors (guideline drafters), who are both constrained and shaped by specific institutional scripts, such as conflicts of interest.

Cost-effectiveness, and its potentially major role in negotiating guidelines, especially in England is also explored. An extensive literature that deals with cost in relation to health technology assessments exists.[138] This study will build on this literature by exploring whether these types of studies apply to medical guidelines – so far research of this nature has been absent. Cost-effectiveness considerations have the potential to

---

137 Meyer and Jepperson (n 75); Meyer (n 71).

138 Lehoux and Blume (n 88); May (n 88).

interact with assessments based on evidence: this study hopes to bring these literatures together and explore potential interactions between them.

Finally, this study draws on literature that has engaged with the micro-level of guideline construction, through observations and ethnographic academic research. The work of Pagliari[139], Hopthrow[140], Berg[141], Rodwin[142] and Moreira[143] all have shown in various ways that discussions in guideline groups have strong normative dimensions and can be said to represent political process, on the micro-level again the issue of conflicts of interest and its potential corrupting influence comes to the fore.[144] This suggests the interesting questions that this study will need to confront: how can such a supposedly political process lead to the production of objectivity, as science and technology scholars claim? In conclusion, this study draws on several different academic disciplines to provide a broad and detailed discussion of the development of medical guidelines in the Netherlands and England.

---

139 Pagliari, Grimshaw and Eccles (n 68).

140 Hopthrow, Feder and Michie (n 102).

141 Berg, ter Meulen and Van den Burg (n 101).

142 Rodwin (n 113).

143 Moreira (n 21).

144 Goldacre (n 129).

# 3 GUIDELINES IN THE NETHERLANDS AND ENGLAND

This chapter provides the background and context to help us to understand the emergence of medical practice guidelines. Section 3.1 describes the emergence of medical practice guidelines in the Netherlands and England. It explores the different trajectories that the guideline development process has followed in the two countries – particularly the shift from consensus-based to evidence-based guidelines – while drawing out the implications that this might have for this study. Section 3.2 introduces the relevant professional associations and other actors involved in drafting and negotiating guidelines in these two countries. Differences in the composition of guideline working groups and practices in negotiating guidelines will be explored. Section 3.3 discusses guidelines in the broader institutional context of healthcare delivery, the role of the institutional and regulatory context on guidelines, and the financing of guidelines. Concluding remarks are presented in section 3.4 in the form of a cross-country comparison.

## 3.1 Historical Perspectives: From Consensus to Evidence-Based

It is important to trace the history of the development from consensus-based to evidence-based guidelines in both countries and assess when, and to what extent, the shift from consensus-based guideline development to evidence-based guideline development occurred, as this will enable us to understand the current developments in both countries.

*The Netherlands*

In the Netherlands, the production of so called 'consensus statements' began in 1982, when the National Association for Quality Assurance in Hospitals (CBO) [1] produced a set of clinical practice guidelines that described a 'state of the art' procedure a health professional should follow when confronted with a situation the guideline describes.[2] These consensus statements were produced in response to increasing government pressure to cut budgets, which led to the fear that cost considerations might endanger the quality of care.[3] The scientific council of the CBO who produced these initial guidelines consisted of representatives of the various scientific organisations of medical specialists. The CBO was aimed at improving the quality of care in hospitals, and the guidelines were meant to reduce practice variation and provide means to audit care.[4] In 1987 the National Association of General Practitioners (NHG) followed suit by producing the NHG-Standaarden: clinical practice guidelines aimed at illnesses prevalent in GP surgeries.[5] Klazinga provides a

---

1 Centraal begeleidingsorgaan intercollegiale toetsing

2 N Klazinga, 'Compliance with Practice Guidelines: Clinical Autonomy Revisited' (1994) 28 Health policy (Amsterdam, Netherlands) 51.

3 Michael I Harrison, *Implementing Change in Health Systems : Market Reforms in the United Kingdom, Sweden, and the Netherlands* (SAGE 2004) 144.

4 Teun Zuiderent-Jerak and others, 'Variatie in richtlijnen. Wat is het probleem?' (*Variation in Guidelines. What is the problem? IBMG*, 2011) <http://www.pgosupport.nl/mailings/FILES/plugin_content/176/Regieraad_rapport_variatie_in_richtlijnen.pdf> accessed 12 November 2013.

5 JS Burgers and JJE van Everdingen, 'Evidence-based richtlijnontwikkeling in Nederland: het EBRO-platform' (2004) 148 Nederlands Tijdschrift voor Geneeskunde 2057.

survey of the first 12 years of the CBO clinical guideline program and establishes three clusters of reasons for their (more widespread) use: professionalization, accountability and efficiency.[6] Klazinga argues that guidelines help doctors deliver state of the art care as well as potentially delivering transparency about treatments to patients, insurers, management and government, and possibly helping to ration care and control costs.[7]

The Cochrane Centre was set up in the Netherlands (following England) to conduct reviews of RCTs.[8] Initially the CBO and the Cochrane Centre started working together, followed by a great many other scientific associations of doctors who later joined the collaboration.[9] The shift from consensus-based to evidence-based guidelines, however, was not a radical 'paradigm shift' as some literature suggests.[10] Rather in the Netherlands it meant that evidence-based methods became part of the consensus-based approach to development of a guideline.[11] It follows that although the term 'evidence based' is often used to describe the guideline development process in the Netherlands, a more accurate description of the process would focus on the role of consensus in establishing what the

---

6 Klazinga (n 2).

7 ibid.

8 'Dutch Cochrane Centre' <http://dcc.cochrane.org/> accessed 26 May 2014.

9 Burgers and Everdingen (n 5).

10 Evidence-Based Medicine Working Group, 'Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine' (1992) 268 JAMA 2420.

11 Zuiderent-Jerak and others (n 4).

power and relevance is of certain scientific evidence.[12] This has not prevented evidence-based guideline development from becoming the preferred term to describe guidelines. A good example of this point is the EBRO-platform[13], which brought together organisations that were involved in evidence-based guideline development. The EBRO-platform was launched in 2001 and was financed by the Order of Medical Specialists (OMS). Since then all clinical practice guidelines strive to be evidence-based, and advances were made in further formalizing the procedure for drafting guidelines.[14] Furthermore, textbooks have been produced that promote evidence-based guideline development as the new standard and suggest a (normatively) correct way of making a guideline.[15]

Guideline production remained in the hands of the professional associations from the 1980s until the present day. However, an interesting recent development in the Netherlands is the stronger involvement of the government in guaranteeing guideline quality; a quality assurance institute has been set up as a part of the *Zorginstituut*, a body

---

12 Loes Knaapen and others, 'Pragmatic Evidence and Textual Arrangements: A Case Study of French Clinical Cancer Guidelines' (2010) 71 Social science & medicine (1982) 685.

13 Evidence-based guideline development

14 An example is the 'procedurebook' used by the National College of General Practitioners: 'procedureboek voor de ontwikkeling van NHG-standaarden, NHG 2010'

15 JJE van Everdingen, *Evidence-based richtlijnontwikkeling: een leidraad voor de praktijk* (Bohn Stafleu Van Loghum 2004); JJE van Everdingen, *Handboek evidence-based richtlijnontwikkeling: een leidraad voor de praktijk* (Bohn Stafleu van Loghum, Springer Media 2014).

functioning at arms-length from government.[16] This quality assurance institute has opened a register in which quality instruments, such as guidelines, can be entered. Before guidelines will be allowed on this registry they are assessed by the Zorginstituut.[17] In the Dutch healthcare system, guidelines are now ubiquitous, and all scientific associations of doctors have their own guideline program, as a result the number of NHG-standards has grown to 89.[18] The NHG, however, is only one of a multitude of organisations involved in guideline production: the associations of medical specialists, and several para-medical professions also produce guidelines.[19] The total number of guidelines is hard to establish but it is clear that the figure is in the region of hundreds: several databases exist that give some indications[20] but a specific number is hard to give.

---

16 Zorginstituut Nederland, 'Kwaliteit' (*Quality*) <http://zorginstituutnederland.nl/kwaliteit> accessed 26 May 2014.

17 This development is recent (2014-2016) and so in discussing the Dutch guidelines in this study these developments do not play a role.

18 On 14-2-2017. 'NHG-Standaarden | NHG' (*Standards of the Dutch Scientifc Association of General Practicioners*) <https://www.nhg.org/nhg-standaarden> accessed 25 May 2014.

19 Artsennet a Dutch website for doctors shows around 75 organisations involved in guideline production. 'Medische Richtlijnen per Organisatie' (*Medical guidelines by organisation*) <http://www.artsennet.nl/Richtlijnen/Per-organisatie.htm> accessed 22 May 2014.

20 'Richtlijnendatabase' <http://kims.orde.nl/het-kennisinstituut-voor-u/richtlijnen/richtlijnendatabase.html> accessed 11 June 2014; 'International Guideline Library — Guidelines International Network' <http://www.g-i-n.net/library/international-guidelines-library> accessed 27 June 2014; 'Richtlijnen // Diliguide.Nl' <http://www.diliguide.nl/> accessed 27 June 2014; 'Oncoline' <http://www.oncoline.nl/> accessed 27 June 2014; 'Pallialine' <http://www.pallialine.nl/> accessed 27 June 2014.

*England*

In England consensus-based guidelines were also developed in the late 1980s, although these efforts mainly produced guidelines aimed at specific health districts where NHS trust worked together.[21] There are several examples here of guidelines using this model: The Royal College of Physicians guidelines on stroke provide an example of a development of a national guideline in the early 1990s.[22] The Clinical Standards Advisory Group set up by the NHS executive board produced early guidelines on back pain[23] to reduce practice variation, while the Centre for Health Services Research of the University of Newcastle also produced guidelines with the aim of improving medical research.[24] The real impetus for guideline production, however, came with the government establishing NICE[25], which was set-up in 1999 as a special health authority part of NHS England with the aim of reducing variation in the availability and quality of NHS treatments and care.[26] With the advent of NICE, guideline production became more formalised. NICE has become a

---

21 D Armstrong and others, 'Development of Clinical Guidelines in a Health District: An Attempt to Find Consensus.' (1992) 1 Quality in Health Care 241.

22 Jako S Burgers and others, 'Towards Evidence-Based Clinical Practice: An International Survey of 18 Clinical Guideline Programs' (2003) 15 International Journal for Quality in Health Care 31.

23 Clinical Standards Advisory Group, *Back Pain.* (HMSO 1994).

24 Burgers and others (n 22).

25 National Institute for Health and Care Excellence, formerly called National Institute for Clinical Excellence.

26 http://www.nice.org.uk/aboutnice/whoweare/who_we_are.jsp; National Institute for Clinical Excellence (Establishment and Constitution) Order 1999, S.I. 1999/220.

statutory body as a result of the HSCA 2012.[27] To date[28] 182 clinical guidelines are available, with 57 in development.[29] The procedure to develop these guidelines is based on the evidence-based guideline development model discussed earlier. NICE has a clear national remit; its guidelines are disseminated to all NHS trusts. NICE has therefore largely replaced the earlier more local initiatives. This has also meant that because NICE started out based on the evidence-based model of developing guidelines it does not have an institutional history of using consensus-based methods, that were used in the 1980s and early 1990s.

The establishment of NICE, an independent non-departmental public body, has caused a shift from professional control over guidelines to a situation of (at least on paper) large government influence over guideline production, albeit at arms-length. Where initially guideline production by professional associations wasn't regulated or controlled by government, the establishment of NICE has meant direct government involvement in the financing off and priority-setting for guideline making.

---

27 Health and Social Care Act 2012, s 232.

28 14-2-2017.

29 'Published Clinical Guidelines' <http://guidance.nice.org.uk/CG/Published> accessed 26 May 2014.

*Concluding Remarks: Similarities and Differences*

The history of medical guidelines in the Netherlands and England show some important commonalities: there is a trend towards more formalised guideline production and international cooperation on guidelines through the so called Guidelines International network (G-I-N).[30] In The Netherlands the consensus-based models used in developing the first guidelines still plays a role, while within NICE there is no institutional history of using consensus-based methods for developing guidelines. Evaluating the precise relationship between consensus-based and evidence-based methods in this study, through interviews and the analysis of guideline documents helps to understand the process of negotiation of a guideline and the nature of the choices that are made.

There are further differences if we consider the ownership of medical guidelines. In the Netherlands the various professional organisations develop their own guidelines, and are therefore more likely to feel ownership of these guidelines and follow their own disciplinary outlook in deciding on what to include in a guideline. In England, because NICE has ownership of guidelines, control is removed from professional associations and

30 Loes Knaapen, 'European Regulation and Harmonization of Clinical Practice Guidelines', *European Union Public Health Policy: Regional and Global Trends* (Routledge 2013).

moved to (representatives) of the state. The practitioners who are asked to implement the guideline are further removed from its conception.

A final point of apparent difference is the fragmentation of guidelines production in the Netherlands as compared to the more centralised approach in England. It remains to be seen whether this fragmentation has any meaningful effects on the development of guidelines. One potential way this fragmentation could affect the content in guidelines is through a need to coordinate advice in guidelines between different medical professionals.

## 3.2 Professional Associations and Guideline Actors

We have seen that both professional associations as well as government bodies play a role in guideline production and that guidelines are developed in guideline working groups. We need to understand who is involved in these associations, and working groups, to help to create a detailed understanding of how these guideline groups work and how they contribute to the development of medical guidelines. The specific institutional environment varies between countries so a detailed understanding of how these environments operate will aid the analysis of the development of medical guidelines.

*The Netherlands*

In the Netherlands healthcare professionals are highly organised in numerous scientific and professional associations. These associations have always played an important role in

shaping the healthcare system. Under the corporatist model of the 1970s and 1980s, doctors (under the umbrella of their associations) negotiated directly with the government on tariffs for their care and displayed a great amount of autonomy in the way they practised their profession. Three associations that are among the most influential will be included in this study, and the way they develop guidelines and their wider influence on healthcare will be discussed: the Royal Dutch Medical Association (KNMG), the National Association of General Practitioners (LHV), and the Dutch Order of Medical Specialists (OMS).

The Royal Dutch Medical Association is a federation of 7 associations of health professionals established in 1999.[31] The KNMG is very influential through various regulatory and non-regulatory activities. It is responsible for the regulation of the fitness to practice licence for physicians and their 5 year revalidation process, subject to agreement of the minister of health.[32] The KNMG also produces guidelines on various topics, the most

---

31 'The federation consists of the Association of Public Health Physicians (KAMG), the National Association of salaried Doctors (LAD), the National Association of General Practitioners (LHV), the Dutch Association for occupational Health (NVAB), the Association for elderly care physicians (Verenso), the Dutch Association of Insurance Medicine (NVVG), the Dutch Order of Medical Specialists (Orde van Medisch Specialisten) and a group of individual KNMG members and students.' http://knmg.artsennet.nl/Over-KNMG/about-knmg.htm.

32 One example is the decision to ratify the decision by a commission of the KNMG to set the training requirements for GPs. 'Goedkeuringsbesluit Kaderbesluit En Drie Specifieke Besluiten van de CHVG Betreffende Huisartsgeneeskunde, Verpleeghuisgeneeskunde En Medische Zorg Voor Verstandelijk Gehandicapten' (*Ratification of the decision by the commission for GP care, social care and care for the mentally disabled on the training requirements*) <http://wetten.overheid.nl/BWBR0017529/geldigheidsdatum_10-02-2014> accessed 26 May 2014.

comprehensive of those is on end-of-life care, including euthanasia.[33] Its process for producing guidelines is to establish a commission that consists of representatives from all the various medical disciplines directly involved in the topic of the guideline, who participate based on their expertise not as representatives of a professional association. After a draft guideline is produced there is a consultation in which stakeholders can comment on the guidelines. In practice, these are often the professional associations most closely involved with the topic of the guideline. In addition to its involvement in guidelines the KNMG carries considerable political clout, as it actively lobbies parliament and the health minister.[34] Elements of the Dutch '*poldermodel*'[35] still dictate that measures are preferably taken when there is support in the field, and the KNMG often acts as the voice of health care professionals and so can provide or withdraw this support.

---

33 'KNMG/KNMP Richtlijn Uitvoering Euthanasie En Hulp Bij Zelfdoding' <http://knmg.artsennet.nl/web/file?uuid=9899cb4d-d598-4f58-974c-10ad38153485&owner=a8a9ce0e-f42b-47a5-960e-be08025b7b04&contentid=119543> accessed 22 October 2013.

34 Examples can be found in its website where very regularly comments on new legislative initiatives appear as well as letters being send to the parliament and reports of meetings with the health ministry are published.

35 The poldermodel describes a tradition of consensual decision-making especially prevalent in the Netherland during the purple coalition years when the Labour Party, the Liberal Democrats, and the Conservatives formed a government, but is also used to describe earlier consensual styles of government all the way back to the time of the Republic of the Seven United Provinces of the Netherlands. The use of the world polder is appropriate as it describes the forming of an area of land reclaimed from the sea and protected by dikes that can only come about through cooperation.

Two other Dutch professional bodies that are relevant for this study are the LHV, the National Association of General Practitioners, and the OMS, the Dutch Order of Medical Specialists. The LHV works through its sister organisation the NHG which promotes the scientific foundation of the work of GPs, by publishing the NHG-standards, medical guidelines for GPs: a selection of which will be included in this study.[36] The NHG guideline process is very similar to that of the KNMG, the obvious difference being that the members of the guideline development group are mostly GPs. The consultation is carried out through sending the concept of the guidance to 50 randomly selected GPs, and references are sought from other professional associations. After this the guideline is presented to the 'authorisation committee', which has the final word on whether the guideline should be published and to what extent it might need modification before authorisation would be granted. This committee is composed of several professors of primary care medicine in the Netherlands and tasks itself with providing a final 'quality-check' of the guideline that is developed. The LHV has a membership rate of about 98% of all GPs in the Netherlands[37] and is a vocal supporter of 'GP-care' (*huisartsgeneeskunde*) and actively lobbies to that effect.

---

36 'NHG-Standaarden | NHG' (n 18).

37 'Boetereapport NMa' (*Rapport of a fine by the Dutch Competition Authority*) note 3 <http://medischcontact.artsennet.nl/web/file?uuid=03be9098-ad0e-4555-a31e-d726f780a925&owner=86bbbdd9-5ff8-4303-ab88-f8888509486c> accessed 25 May 2014.

The OMS provides a very similar role for 'specialists', the Dutch term for consultants and surgeons in hospitals. The OMS sponsored the EBRO-platform that works on evidence-based guideline development, and various different scientific organisations of medical specialists come together on this platform informally to work on guidelines.[38] Currently the OMS supports various professional organisations of specialists through its knowledge centre, which promotes and supports working with the evidence-based guideline development method, and maintains a database of guidelines produced according to the EBRO method.[39] OMS does not produce guidelines itself, rather it acts as an umbrella for various scientific organisations of specialists, and experts from its knowledge institute participate in guideline production.

From this brief overview, we can see that professional bodies in the Netherlands combine various roles: they are regulators through setting standards for education of professionals, guideline developers, and provide advocacy services. Guideline production is not an activity in isolation but is part of a broader drive for quality improvement by the professional bodies. The knowledge institute of OMS explains this well: 'guidelines are the basis of a policy of quality [in healthcare] of medical specialists and form a bridge

---

38 Burgers and Everdingen (n 5).

39 'Richtlijnendatabase' (n 20).

between science and practice.'[40] These three professional associations, although influential, only suggest a very partial picture of all organisations involved in guideline production. Guideline production in the Netherlands as a result is highly decentralised and fragmented with no clear leader and no hierarchical ordering of guidelines.

An important feature of these arrangements is that guidelines are produced by professional associations primarily for use by those professional associations. There are exceptions to this, where multidisciplinary groups work on guidelines across professional boundaries, such as guidelines made by CBO.[41] Nonetheless a primary driver for guideline production is provided by the wish to maintain (and improve) the quality of care provided by distinct groups of medical professionals. These guideline efforts seem to be mainly driven by the wish to self-regulate and are not primarily driven by government intervention. Competition between varies professional groups with the medical profession is also a part

---

40 'Kennisinstituut van Medisch Specialisten - Het Belang van Een Goede Richtlijn' (*Knowledge Institute of Medical Specialists - The importance of a good guideline*) <http://kims.orde.nl/het-kennisinstituut-voor-u/richtlijnen/het-belang-van-een-goede-richtlijn.html> accessed 11 June 2014.

41 'CBO-Richtlijn Voor Diagnostiek En Behandeling van Acute En Chronische Aspecifieke Lage Rugklachten | Nederlands Tijdschrift Voor Geneeskunde' <http://www.ntvg.nl.proxy-ub.rug.nl/artikelen/cbo-richtlijn-voor-diagnostiek-en-behandeling-van-acute-en-chronische-aspecifieke-lage/ingezonden-mededelingen#1225672> accessed 21 August 2014.

of this picture. Klazinga[42] and de Herk[43] show how guidelines are a response to changing government pressures, and in that sense, can be seen as an attempt to ward off the (implicit) threat by government bodies to more directly regulate the profession. Guidelines in the Netherlands are developed because of an independent wish to further regulate the professional standards, and improve the quality of care, while at the same time provide assurance to the government that through these efforts the profession can still be trusted to regulate its own affairs.

Guidelines in the Netherlands therefore have features of both professional self-regulation and government (mandated) self-regulation. Seeing this on a spectrum between the two extremes of 'pure' self-regulation and 'pure' government regulation, Dutch arrangements would lean towards the side of professional self-regulation.

One further player needs to be discussed, the healthcare insurers.[44] Around 36 different insurer's labels are active on the Dutch market, which represents around 26

---

42 Nicolaas Sieds Klazinga, 'Quality management of medical specialist care in the Netherlands. An explorative study of its nature and development' (Erasmus Universiteit Rotterdam 1996) <http://www.bmg.eur.nl/fileadmin/ASSETS/bmg/Onderzoek/Promoties/Promoties_1996/Klazinga_1996_web.pdf> accessed 24 September 2014.

43 R van Herk, 'Artsen onder druk. Het kwaliteitsbeleid van de medische beroepen in Groot-Brittannië, Nederland en België tussen 1970 en 1996, als gevolg van de interne en externe druk op de zelfregulering van artsen.' (Erasmus Universiteit Rotterdam 1997) <http://repub.eur.nl/pub/18416/971023_HERK,%20Rudi%20van.pdf> accessed 24 November 2014.

44 These health insurers are insurance companies who solely provide insurance for health care.

different insurance companies, of which the four largest conglomerates cover 95% of the market.[45] Healthcare insurers play no direct role in the drafting of guidelines. They do however influence healthcare professionals through policies that refer to these guidelines. If a patient is insured by a healthcare insurer he agrees to the specific policy the healthcare insurers has for reimbursing healthcare costs. Whether certain costs are reimbursed sometimes depends on whether the treatment provided by the healthcare professional conforms to guidelines. One example is the policy of the Friesland Health Insurer, who only reimburses bariatric surgery for cases of obesity if this conforms to the relevant medical practice guidelines.[46] This could potentially make guidelines into more than a quality instrument and extend its range to influence decisions on reimbursement of costs. Given the fact that guideline drafters will likely be aware of the way guidelines are used by healthcare insurance companies it is plausible they would be influenced by how insurers use their guidelines in practice. While healthcare insurers are thus not central to this research, for an overall picture of the Dutch healthcare system they need to be mentioned.

---

45 'Zorgverzekeraars in Nederland' (*Health Insurers in the Netherlands*) <http://www.zorghulpatlas.nl/zorghulpwereld/zorgverzekering/zorgverzekeraars-in-nederland/> accessed 25 May 2014; 'Overzicht Zorgverzekeraars in Nederland' (*List of Health Insurers in the Netherlands*) <http://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/rubrieken/verzekering/1302-overzicht-zorgverzekeraars-in-nederland/Overzicht+zorgverzekeraars+%28in+Nederland%29.pdf> accessed 25 May 2014.

46 De Friesland Zorgverzekeraar, 'Basisverzekering Online Verzorgd Polis 2014' (*Basic Health Policy Online Care Policy 2014*) <https://www.defriesland.nl/Images/Basisverzekering-Online-Verzorgd-Polis-2014-11650.pdf> accessed 26 June 2014.

*England*

In England, NICE functions as a central body for the production of guidelines, and has formal methods it employs in deciding who the actors are that are involved in guideline production. It explains on its website that the structure of guideline development has eight steps. After referral by the Department of Health, stakeholders can register their interest. These stakeholders can be national organisations of patients or carers, health professionals or representatives of the pharmaceutical industry, and they are consulted during the process. Then the National Collaborating Centre (NCC), an independent guideline centre funded by NICE, which consists of representatives of several royal colleges of medicine, prepares the scope of the guideline. Only then is the guideline development group established: consisting of medical professionals, representatives of patient groups and technical experts, mainly health economists but many of these may also be medical information specialists. Guidelines are therefore always made by multidisciplinary groups of medical professionals. A draft guideline is produced by the group and opened for public consultation, where only registered stakeholders can make comments. The NCC produces the guidance and finally NICE authorises the guideline and disseminates it.[47] This process is highly structured and formalised, and NICE therefore scores very high on the AGREE

47 'Developing NICE Clinical Guidelines' <http://www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/developing_nice_clinical_guidelines.jsp> accessed 26 May 2014.

scale, which measures the rigorousness of the methods employed in guideline production and is used by academics and other experts to compare guideline institutions.[48] A further important characteristic of NICE is the mandatory involvement of a health economist in guideline development, and an explicit requirement to take cost-effectiveness into account when developing a guideline as stated in the guideline manual used by NICE.[49]

This shows a picture of NICE replacing more informal attempts to make guidelines with a more rigid and formalised structure. Professional associations are not completely excluded from the process as they could influence who will be part of the working group in the National Collaborating Centre, and in this indirect way remain in control of the process. The arrangements are closer to a form of state regulation, as the professional associations can only participate in the guideline making process on the terms set by NICE. The mandatory involvement of a health economist for example also points in that direction. NICE has not completely replaced other bodies in making guidelines, there are still Royal Colleges who produce their own guidelines, but they often take the form of documents providing a more localized or contextualised version of a NICE guideline. It would be fair

---

48 G Leng and M Baker, 'NICE Accredits Guidelines Using International AGREE Criteria' (2013) 347 BMJ f4771.

49 'The Guidelines Manual' ch 7 <http://publications.nice.org.uk/the-guidelines-manual-pmg6> accessed 26 May 2014.

to say that NICE guidelines have the greatest authority among guidelines in the medical field in England.

*Comparing institutions in England and Netherlands*

While the historical trajectories of guideline development in the Netherlands and England start out looking very similar, with the advent of NICE the actors in the guideline field take on a different shape. In the Netherlands professional associations control the development of guidelines, and they invite members of their own associations to participate in producing them. While in England NICE sets priorities for guideline developments and members of associations have a more indirect role in the process through participation in Guideline Centres who provide the working groups developing guidelines. Whether these differences on paper turn out to be relevant differences when it comes to the process of development of a guideline is something this study will explore.

Another clear institutional difference is the fact that NICE makes it mandatory for its guideline group to include the patient perspective through the requirement that the guideline development group has a member from a patient organisation. In the Netherlands

patients are not always included in guideline groups, but other methods are sometimes used to establish the patient perspective such as questionnaires.[50]

The more fragmented nature of guideline production in the Netherlands opens up the possibility that several specialities might produce guidelines that are not fully joined up, for example, when the GPs might provide different advice from consultants in hospitals because they refer to different guidelines. This in turn might affect the drafting of guidelines if different parties aim to avoid those differences and strive for harmonisation. NICE guidance is less likely to lead to conflicting results in this way because it is always multidisciplinary in nature, where everyone in the chain of care is involved in the drafting process. For NICE guidelines a potential pitfall would be to ensure the *internal* consistency of the guidelines it produces. The potential impact of these differences will be included in the assessment of guidelines in this study.

Cost-effectiveness considerations and their role in the development of guidelines, and the way that these elements of the guideline development process are institutionally embedded, differ between the two countries. In England, NICE, through its guideline manual ensures a role for cost-effectiveness considerations in the development of guidelines. Any such reference is absent in the Netherlands. Potentially this difference

---

50 'Richtlijnendatabase' (n 20). An analysis of a selection of the guidelines in this database shows various ways of including the 'patient perspective' in guideline production.

might shape the development of guidelines substantially and this study will consider how these prima facie differences work out in practice.

## 3.3 The Regulatory Environment of Guidelines

We have identified the various organisations that play a role in the healthcare system and have the potential to influence, or work with, guidelines. The broader institutional and regulatory context in which 'the world of guidelines' is located are the institutions that make up the healthcare system: hospitals, regulators, the Department of Health, and parliament. A description of the relevant bodies will enable us to establish where potential relationships between guidelines and policy in the wider context of the healthcare system take place. In the Netherlands this means the healthcare system as it stands after the last major reform in 2006 will be our starting point, while in England the health care system after the reform of the Health and Social Care Act 2012 will be considered. In describing these two healthcare systems no attempt is made to be fully comprehensive, rather only those aspects of the healthcare systems that could plausibly influence the development of guidelines will be discussed. By placing a stress on both the institutions and their regulatory functions a good picture of these aspects of the context of guideline development will emerge.

*Netherlands*

This description will describe the situation as it was in the Netherlands up to 2014 when the research undertaken for this study started and will not address in detail more recent developments.[51] In the Netherlands the healthcare a system can be described as one of managed competition. Healthcare insurers play a central role as they contract care with providers to make sure the insured get the healthcare they are entitled to. Furthermore, there is a contractual relationship between the insured/patient and the healthcare insurer through an insurance policy. A further relationship is the relationship between the patient and healthcare professional, where the Act on the medical treatment contract plays a role.[52]

---

51 In relation to medical guidelines there are two recent developments that are noteworthy, the establishment of a quality institute as part of the Netherlands Health Institute and the codification of the medical standard as positive norm in article 7:453 of the Dutch Civil Code, which now includes a reference to the register of quality standards that is kept by the quality institute. These two developments signal a shift from reliance on professional self-regulation to more state regulation in the field of quality of care. That is, while in the past professional standards were purely a matter for professional bodies, now the healthcare institute has a role in authorizing standards and can as last resort require healthcare parties to draw up standards if insufficient action is taken. It is also made more explicit by the state that healthcare professionals should follow guidelines or explain why they deviate from it, through this explicit reference to the register of standards. These developments started in response to professionals failing to reach consensus on guidelines in various areas of practice. These developments happened after the guidelines that are discussed in this thesis had all been published.

52 Wet op de geneeskundige behandelovereenkomst 1994 (Stb 1994, 837).

This system of competition is regulated and managed through several bodies who are (to a greater and lesser extent) independent from the government: ACM, NZA, IGZ and Zorginstituut Nederland. The ACM (Authority for Consumers and Markets) is the competition regulator who works in conjunction with the NZA (Dutch Healthcare Authority), who is the specialised healthcare competition regulator. The IGZ (Health Inspectorate) oversees the quality of healthcare.

As discussed in chapter 1, the Zorginstituut Nederland (The National Health Care Institute) addresses various health insurance aspects such as the risk adjusted compensation payments to insurers[53], enforcement of the mandatory insurance requirement, and through its quality institute it facilitates medical guideline production and dissemination. It also is responsible for deciding on the limits of the insured package of care on behalf of the government, and makes cost-effectiveness judgements based on the principles of evidence-based medicine. This is important as it explains why cost-effectiveness considerations do not feature prominently in Dutch guidelines, as these issues are decided elsewhere by the Zorginstituut (and its predecessor institutions). Of these institutions only the Zorginstituut

[53] To achieve community-rated insurance premiums a sophisticated risk-adjustment payment is made between insurers through Zorginstituut Nederland, to compensate for 'bad risks' (i.e. people who have high health care needs), this task has recently been moved to the CAK (centraal administratiekantoor), these changes are not further detailed here.

has any active involvement in guidelines, but most of its work in this area has started after 2014, so did not play a role in the development of the guidelines in this study.

Finally, the Ministry of Health has 'system responsibility' for the overall healthcare system, making sure that the system delivers affordable high-quality care. But this 'responsibility' is rather different from that of the Secretary of State in England, and is not a direct legal obligation to provide for a comprehensive healthcare service, as in the case of England. The Ministry of Health might be described as a 'meta-regulator' encouraging and steering independent parties towards the goal, but not providing any healthcare directly. Politicians are at a greater distance from healthcare delivery, and this has been a deliberate choice in the design of the healthcare system as a result of the 2006 reforms.

The Dutch healthcare system might be described in terms of an attempt at striking a balance between the opposing interests of medical professionals that would strive to provide the maximum amount of healthcare for the highest possible remuneration and healthcare insurers who strive to keep the premiums they have to charge to their customers (all Dutch citizens) as low as possible to maximize their market share.[54] The Dutch

[54] Healthcare insurers are allowed to make a profit but have to reinvest this into healthcare (through lower premiums) and are not allowed to distribute this to shareholders or other external parties.

government through its regulators acts as referee and by providing detailed regulation of this 'market' creates a level 'playing field', to stay with an economic analogy.

What is interesting about these arrangements is the fact that none of these seem to directly affect the development of guidelines. This organisation of healthcare leaves the responsibility for producing, maintaining, and updating guidelines in the hands of professional associations. This seems to be the result of the fact that most of the regulatory interventions in the Dutch healthcare systems are concerned more with price than with issues of quality.

This leaves open the issue of financing of medical practice guidelines in the Netherlands. This financing comes from various sources and often has an ad-hoc nature. Some guidelines are financed through the Foundation for the Financing of Quality of Medical Specialists[55], some guidelines are financed through payments by hospitals for their use, such as the guidelines developed by the working group for infection prevention.[56] In some cases, such as those of the former EBRO-platform, guideline production is funded by subsidies from the Ministry of Health. Finally, most of the guidelines are financed out

---

55 'Orde van Medisch Specialisten - Stichting Kwaliteitsgelden Medisch Specialisten (SKMS)' <http://www.orde.nl/over-oms/wat-doet-de-oms/stichting-kwaliteitsgelden-medisch-specialisten-skms.html> accessed 11 June 2014.

56 'Werkgroep Infectiepreventie (WIP)' <http://www.rivm.nl/Onderwerpen/W/Werkgroep_Infectiepreventie_WIP> accessed 15 May 2014.

of the budget of the scientific or professional organisation producing the guideline, such as the NHG-Standaarden who are financed by the NHG itself.

*England*

The situation in England is described in relation to the most recent 2012 healthcare reforms. The Secretary of State and NHS England can give directions to NICE on which clinical guidelines to produce and in this way can potentially actively steer which guidelines are produced, however, the government cannot influence the content of a guideline. This aspect is left to professional self-regulation, as the guideline development groups employ medical professionals that negotiate on the content of the guideline.

The English healthcare system has features of competition, but this competition takes place within an 'internal market', it mimics features of a market within the bureaucratic system of the nationalised hospitals that exists in England. It is clear that the state plays a more active role in how this system functions, by steering actively how the 'market' functions within the existing structures through several arms-length bodies. The government is also actively involved in the production of guidelines through its control of NICE. This allows for the setting of the topics that need to be covered by guidelines. Furthermore, the government controls the budget of NICE and in this way is directly in

control of the financing of guidelines, and can potentially choose areas for guideline development of others.

Although it is clear that there are no healthcare insurers in the NHS, this does not mean that 'contract' does not play a role, this arrangement is established through NHS England and clinical commission groups (CCGs) that act as commissioners and so mimic a market situation. The Standard NHS-contract between a CCG and a provider of healthcare, say an NHS Foundation Trust, can require adherence to clinical guidelines.

NICE, however, does not have a monopoly on guideline production: the Royal Colleges and various associations of doctors do produce guidelines independent from NICE[57], and so does the General Medical Council in its publication Good Medical Practice[58] that provides ethical guidelines to doctors.

The English healthcare system is now based on the Health and Social Care Act 2012 (HSCA). The Health and Social Care Act 2012 opens with changing the role of the Secretary of State from providing a comprehensive healthcare service to *only securing that*

---

57 The Royal Colleges as well as associations of specialists and nurses.

58 General Medical Council, 'Good Medical Practice (2013)' (2013) <http://www.gmc-uk.org/static/documents/content/GMP_2013.pdf_51447599.pdf> accessed 9 January 2014.

a comprehensive healthcare service is provided for.[59] This might look like a small change in wording but it embodies a radical shift in the conception of how the NHS should be run. The Secretary of State now delegates the function of providing a comprehensive health service to the NHS Health Commissioning Board, clinical commissioning groups and local authorities. However, the Secretary of State retains the responsibility towards parliament for providing a comprehensive health service.[60] This set-up inevitably causes tensions between the independence of the arms-length bodies and the responsibility of the Secretary of State when something goes wrong, and the promise in the 2010 white paper to liberate the NHS (from day-to-day political control).[61] The Department of Health plays a central role in the NHS, after the 2012 reforms that nominally seemed to reduce its influence; it appears the Department of Health has only gained in importance. But before we consider the Secretary of State and the Department of Health we have to describe the arm's-length bodies.

The NHS Commissioning board, operating under the name of NHS England is the central commissioning body responsible for commissioning specialised nationalised

---

59 Health and Social Care Act 2012 s1: The Secretary of State must continue the promotion in England of a comprehensive health service designed to secure improvement— (a)in the physical and mental health of the people of England, and (b)in the prevention, diagnosis and treatment of physical and mental illness. (2) For that purpose, the Secretary of State must exercise the functions conferred by this Act so as to secure that services are provided in accordance with this Act.

60 Health and Social Care Act 2012 s1(3): The Secretary of State retains ministerial responsibility to Parliament for the provision of the health service in England.

61 Department of Health, *Equity and Excellence: Liberating the NHS.* (Stationery Office 2010).

services, primary care and care for offenders, and to oversee and the 211 new clinical commissioning groups (CCGs). It provides commissioning support to the CCGs and also authorises them, and can place conditions on their authorisation.[62]

These Clinical Commissioning Groups are the 'new form' of practice-based commissioning, formed of groups of general practices in a specified area, that commission healthcare for that area. They have been given a large part of the commissioning budget, and through their local knowledge are seen to be suited to commission services based on the needs of the local population.[63] In making decisions on commissioning care, the CCGs have to pay due regard of competition law, as section 75 of the HSCA 2012 requires, and are not allowed to engage in anti-competitive behaviour. [64] This brings us to NHS Improvement[65], and its extended role as a result of the Health and Social Care Act 2012. NHS Improvement now fulfils several roles: it protects choice for patients and prevents

---

62 'NHS England » CCG Authorisation' <http://www.england.nhs.uk/ourwork/commissioning/ccg-auth/> accessed 26 May 2014.

63 65 billion of the 95 billion budget, the remaining 30 billion is spend by NHS England.

64 S. 76 (1)Regulations may impose requirements on the National Health Service Commissioning Board and clinical commissioning groups for the purpose of securing that, in commissioning health care services for the purposes of the NHS, they— (a)adhere to good practice in relation to procurement; (b)protect and promote the right of patients to make choices with respect to treatment or other health care services provided for the purposes of the NHS; (c)do not engage in anti-competitive behaviour which is against the interests of people who use such services. (3) Regulations under this section may, in particular, impose requirements relating to— (a)competitive tendering for the provision of services;

65 Previously known as Monitor (from July 2012 until April 2016) and more previously as the Independent Regulator for Foundation Trusts (from January 2004).

anti-competitive behaviour, it licenses NHS healthcare providers, it sets the national tariff, and it has a duty to enable integrated care delivery. It is therefore a specialised healthcare competition authority[66] as well as the price-setting body in conjunction with NHS England. At the same time NHS Improvement licences a broad range of providers so as to require them to follow basic transparency and competition criteria.[67] Reducing healthcare inequalities and improving healthcare quality through integrated care delivery also falls to NHS Improvement, which might lead to tensions with its role as competition authority. The Health and Social Care Act 2012 simply states NHS Improvement should resolve these tensions and publish how they have achieved this.[68] These arms-length bodies all report to the Secretary of State for Health.

The Secretary of State has several levers to influence the arms-length bodies it has created: it sets the mandate for the NHS commissioning board[69], give guidance to NHS Improvement about its functions[70], provides guidance on improvement of quality of services[71], has wide powers to intervene when a body fails to act in the interests of the

---

66 With a duty to cooperate with the Competition and Markets Authority.

67 Health and Social Care Act 2012 s 94, standard conditions (for a licence).

68 Health and Social Care Act 2012 s 67.

69 Health and Social Care Act 2012 s 23.

70 s 63.

71 s 23.

health service[72], has a greater role in the reconfigurations of NHS (foundation) Trusts[73], and claims strong oversight over the National Institute for Health and Care Excellence that produces guidelines and quality standards.[74] This list of levers is far from comprehensive.[75] It is of course not self-evident that just because the Secretary of State has these powers it means he/she will use them, and it is too early to draw firm conclusions in this regard, but there are some suggestions that the Secretary of State might be micro-managing the NHS.[76] The last player on the government stage we have to consider is the National Institute for Health and Care Excellence (NICE), its role has largely already been covered in the section on the history of medical guidelines.

*Conclusions*

In both countries, the institutional environment is aligned with a system of managed competition, and the specialist competition regulators (NMA, NHS Improvement) now play a prominent role. However, there are differences in the way in which competition is

---

72 s 71.

73 s 174-178.

74 s 237.

75 It does not list powers of local authorities for example, nor does it list powers in any other Healthcare Act.

76 Denis Campbell, '"Control Freak" Jeremy Hunt Accused of Manipulating NHS for Political Ends' *The Guardian* (27 March 2014) <http://www.theguardian.com/society/2014/mar/27/jeremy-hunt-control-freak-manipulating-nhs> accessed 17 May 2014; Daniel Boffey, 'NHS Chief Admits to Dismay at Jeremy Hunt's Meddling' *The Guardian* (21 December 2013) <http://www.theguardian.com/society/2013/dec/21/nhs-service-jeremy-hunt-malcolm-grant> accessed 17 May 2014.

seen to operate in the two countries. In the Netherlands all hospitals are private not-for profit foundations, while in the NHS the hospitals are nationalised. Therefore, competition in the NHS has a more internal and bureaucratic nature compared to the Dutch healthcare system. These differences are significant insofar that they suggest that the government in the Netherlands takes a more hand-off approach to the healthcare system, and claims only a residual responsibility while in England government is more directly involved in shaping the system of managed competition. Therefore, labels like 'managed competition' might hide more than that they reveal, as these discussions are necessarily in the form of broad strokes.

These more generic differences are reflected in the approach taken to the development of guidelines in both countries. While the development of guidelines is in the hands of professional associations in the Netherlands, in England NICE, a statutory body has a large role. Potentially this institutional difference might have an effect on the development of guidelines. Where guidelines in the Netherlands might plausibly reflect the priorities of healthcare professionals, those in England might more closely align with the interests and priorities of the government of the day.

Guideline financing is another area of difference between the two countries. In the Netherlands there is no single source to finance guidelines. Sometimes professional associations pay themselves, sometimes private or government subsidies cover the cost. In

England the financing of guidelines is provided centrally through the NICE budget, and in this way, is guaranteed and controlled by the government.

## 3.4 Conclusion

The history of medical practice guidelines has shown that behind the common focus on 'evidence-based' medicine in the England and the Netherlands, the guideline process in the Netherlands is more fragmented, with a great variety of organisations being involved in the process without clear coordination. This is in contrast to the centralised approach in England through NICE. There is also a clear difference in ownership of guidelines in the Netherlands, with a process closer to the medical profession, compared to England where the state has more control. When considering the actors in both countries a similar picture emerges. In the Netherlands a greater number of diverse type of actors and organisations are involved in the guideline production process, while in England NICE imposes a more rigid model of guideline production with certain pre-established participants.

Guidelines in both countries can be described as attempts at professional regulation by medical professionals that play out along a spectrum between self-regulation and state-regulation, with the Netherlands tending towards self-regulation and England closer to state-regulation. In both countries these guidelines are made for use in a healthcare system,

and it is the regulatory and institutional environment of this healthcare system that forms part of the context of guideline development.

Both countries describe their healthcare system, in general terms, as one of managed competition. This involves a great many different arms-length bodies that involve themselves in setting prices for a 'market' of healthcare to function effectively. The effect this has on the development of guidelines differs. In the Netherlands these arrangements do not impact the room professional associations have to make guidelines. Rather the environmental factors are positive, in the sense that the government provides (some) subsidies for the development of guidelines and these guidelines are widely used in the healthcare system to support decision-making. In England the situation is similar, competition arrangements do not in themselves directly affect guideline development with the notable exception that the state has decided to create NICE to make guidelines. This means that in the guideline development process cost considerations are more prominent as the government requires NICE to take these into account. In the Netherlands, in contrast these decisions are not taken by guideline makers but by the Zorginstituut, that traditionally has the role of deciding what the extent of the insured care package is.

Finally, related to these differences in government involvement in guidelines is the way in which the financing of guidelines works. While in England there is a government budget to fund NICE and the guidelines it produces, in the Netherlands this funding comes from professional associations, subsidies or other sources. This study has provided some

of the historical background to guidelines, their current institutional and regulatory embedding within healthcare systems and the way they are financed, which we will further explore in the case studies that follow in chapter 4 and 5.

# 4 LOWER BACK PAIN: GUIDELINES IN ENGLAND AND THE NETHERLANDS

Lower back pain is a very common condition, around 60-85% of the population will experience a period of lower back pain during their lifetime.[1] Most of these cases are so-called 'non-specific lower back pain', where there is no identifiable physical cause of the pain. This is an area of medicine where a plethora of different treatments exist for this condition, but not all of them are based on solid evidence: an area where one would expect guidelines to fill a gap by providing advice on which treatments to avoid and which treatments to recommend. We will first consider the English guidelines on this topic before considering the Dutch guidelines. Both these guidelines present a path of care for patients and make several choices on which treatments to offer and in which order to provide them.

## 4.1 The English Guideline on Lower Back Pain

The focus of our analysis will be the 2009 guideline on persistent non-specific lower back pain (hereafter: back pain) drafted by the NCCP[2] and authorised by NICE.[3] The guideline

---

1 Ann C Papageorgiou and others, 'Estimating the Prevalence of Low Back Pain in the General Population: Evidence From the South Manchester Back Pain Survey' (1995) 20 Spine 1889; BF Walker, 'The Prevalence of Low Back Pain: A Systematic Review of the Literature from 1966 to 1998' (2000) 13 Journal of Spinal Disorders 205.

2 National Collaborating Centre for Primary Care.

3 P Savigny and others, 'Low Back Pain: Early Management of Persistent Non-Specific Low Back Pain' (London: National Collaborating Centre for Primary Care and Royal College of General Practitioners 2009) <http://www.nice.org.uk/guidance/cg88/evidence/cg88-low-back-pain-full-guideline3> accessed 13 August 2015.

was drafted by a group of medical professionals from a diverse background including GPs, physiotherapists, a spinal surgeon, nurses and patients. They were supported by a technical team provided by the NCCP consisting of a healthcare economist, information specialists and health fellows.

Before the guideline group was established, the chair for the group was recruited to work with a clinical advisor and the technical team in preparing the scope of the guideline. The chair and clinical advisor were approached informally by NICE and agreed to participate. The reason they were approached was because they had done work for NICE in the past on back pain and related areas, and were therefore well known, and well regarded in this medical field.

The scope sets the limits of the medical condition the guideline will provide recommendations on. In this case, the scope only included back pain from 6 weeks to 1 year.[4] So acute back pain that lasted less than six weeks was not included, nor was chronic back pain that lasted longer than a year. This decision was made during discussions between the chair and clinical advisor of the guideline group and the technical team at

[4] ibid Appendix A.

NICE. NICE had recently published a guideline on acute lower back pain, and wanted to draft a separate guideline for back pain that had a duration for longer than a year.

After the scope had been established it was clear what the expertise of the members of the guideline groups should be, and subsequently they were approached, informally, based on the expertise of the members of NICE; the chair of the guideline group had a role in this as well. Some guideline group members expressed an interest to be included before they were subsequently approached to take part; some of them cited reasons of professional prestige as motivating them, while others were driven by an interest to see first-hand how guidelines were developed. The final group consisted of a diverse group of participants: a GP, nurse, chiropractor, clinical phycologist, osteopath, acupuncturist, occupational therapist, two lay members and several academics in the field of back pain.

Once the group was formed they discussed the scope that had been set before, and agreed that the focus of the guideline would be on looking for evidence that might prevent back pain from becoming chronic pain. As one guideline group member said about this discussion:

> We were then able to re-discuss the scope. (…) But it was decided to go for 6 weeks to twelve months. (…) It was really to look at what literature was

> available (…) the interest was at that stage to improve early treatment (…) to look at how to prevent chronicity (…).[5]

However, not everyone was completely happy with this decision, as a guideline group member explained:

> it [the scope] does tend to cause problems. Particularly with the lower back pain guideline. (…) we were proscribed [by NICE] from including (…) return to work as an outcome measure (…) we weren't to look at anything but chronic [back pain studies] and it was predefined how chronic would be described, with something like lower back pain it is very difficult to have a clear dividing line between an acute [back pain] and chronic [back pain] presentation. The reason why this happened was because there was already a guideline for acute back pain, but [that guideline] was not going to be reviewed. As for return to work there was a working party looking at vocational rehabilitation, but that is very different from a treatment. There was no dissatisfaction but there was a lot of discussion. (…) it [the scope] can be quite restrictive because it is often drawn up without considering the evidence or the clinical problem [at hand]. (…)I think they [later] realised their mistake doing something that was a little bit too narrow. It served their purpose at the time, they didn't want to cover surgery, acute [back pain]. Now they realised that it has to be broader not to be contradictory.[6]

This guideline group member is clearly not convinced by the justifications about the necessity to make different guidelines provided by NICE for limiting the scope of the guideline. His ideas about integrated treatment for back pain could not fully be accommodated within the scope of the work as it was set. Nonetheless, the scope remained

[5] English back pain guideline group member 4 (Middlesbrough, 16th of March 2015).

[6] English back pain guideline group member 3 (Interview by Phone, 27th of March 2015).

unchanged, as it could not effectively be modified anymore at this stage of the development of the guideline.

Once the guideline group had been formed the process of searching for medical studies could begin. As part of this process *key clinical questions* were formulated to guide the search in the literature. The guideline group was involved in formulating these clinical questions and in that way can steer on which topic they look for in medical studies. The results of the literature search, based on these questions, were then considered by the guideline group. The English guideline group decided to only consider RCTs as evidence. If there were no RCTs that came up after a search in medical databases, or the quality of the studies that were found was deemed insufficient, the consensus of the group would provide the base for the recommendations.

This early choice in the process is an important one, as it excludes a great deal of evidence from consideration, namely all studies that are not RCTs, and as a result gives an important role to consensus. When asked why the choice had been made to only include RCT-evidence one member of the guideline group said:

> There was a huge number of stakeholders who had their favourite treatment they wanted to see in the guideline. (…) We will accept only RCTs evidence (…) unless they don't find any [RCTs]. We did that really to draw a level playing field between all the stakeholders. If RCT-evidence is present than we will consider it. (…) So we won't consider a whole volume of papers that weren't RCTs. (…) It has a simplicity that was clearly attractive. (…) When they [other healthcare professionals] later complained (…) we could stand up and say when we investigated the literature there

> was no RCT evidence to support your position.[7] (…) That was really one of the most important decisions [in the guideline drafting process].

This, however, did not mean that the discussion and recommendations did not include reference to other studies, the participants did discuss other studies they found relevant to the issue at hand. A guideline group member explains this as follows:

> It depends on the chair of the group. In our discussion we would also refer to other studies you had heard about. Normally that would be allowed and we would discuss those but we could only refer [in the guideline] to RCTs.[8]

This shows an interesting tension in defining evidence by members of the guideline group. In the text of the guideline, guideline members have agreed, only RCTs will be referred to, so externally only this type of 'high-quality' evidence is presented to the public. Internally, however, a broader definition of what is considered evidence is adhered to, as guideline members are free to discuss other studies they are aware of in arguing for or against advice for treatments on lower back pain. This provides a starting point to consider how evidence is negotiated and constructed through the debates in the guideline group.

A further selection process played a role, which was tied to the decision to limit the scope to patients with lower back pain for longer than 6 weeks and shorter than 1 year in

---

[7] English back pain guideline group member 4 (Middlesbrough, 16th of March 2015).

[8] English back pain guideline group member 2 (London, 20th of February 2015).

duration and to only accept RCTs. A guideline group member comments on this process of searching the literature in the following terms:

> That [the results of a search in the literature] includes all the studies they [the technical team] have selected. My impression is that [tend] to be quite highly selective based on scope. They tend to be more narrow in terms of their selection from what I have seen, then what you might do if you were performing a Cochrane Review [a type of systematic review]. So in practice it tends to be broader, if you do a systematic review as a research project you tend to be as broad as you can and select all the data in and analyse it. NICE tend to narrow down much more specifically within the scope question. So they are much more likely to exclude big studies that look like they might be in the area. (…) They may well exclude those [studies] because they do not fit closely enough in the 6 weeks to one-year bracket [the scope agreed].[9]

These comments are interesting, because at least in the eyes of this guideline group member, the evidence selection process seems to be too narrow. This opinion was not directly confirmed by other members of the group, but some did say that they would discuss more studies, than only those selected by the technical team.

These developments show how the process of forming a guideline group and defining both the scope and the key clinical questions introduces a process of selection of evidence, to an extent the selection of questions to ask, in the form of clinical question, steers what answers will be found. How evidence was constructed and used in making

[9] English back pain guideline group member 2 (London, 20th of February 2015).

choices for treatments in the English guideline will be considered next in greater detail by analysing the recommendations for various treatments in turn.

## 4.2 Evidence-base for English Lower Back Pain Treatments

There are several treatments for lower back pain the English guideline group considered. Together they form a pathway of care for patients that suffer from back pain for 6 weeks until 1 year. The treatments are offered in sequential order and increase from advice and pain relief to surgical intervention. Four treatments for back pain in this pathway are of particular interest: educational advice, exercise, manual therapy and acupuncture. Educational advice is seen as the first choice, the three treatments should be offered after this and are presented as equal, such that the choice of a treatment should be guided by the choice of the patient. Not all aspects of drafting a guideline would necessarily be controversial. There were a couple of issues the English guideline group was confronted with in establishing the evidence for treatments, and interpreting this evidence in light of the goal of the guideline. One of these was the availability of a large amount of (potential) treatments for lower back pain. The guideline would need to establish which of those to recommend and which to advise against. When this was established, the group then needed to work out what the best order was to provide these treatments in. By defining these two

elements a successful pathway of care could be defined for the group of patients that suffered back pain between 6 weeks and one year in duration.

*Evidence Use for Educational Advice*

The first area the guideline group discussed was educational booklets, and whether they help in the treatment of lower back pain. After considering that there are only three RCTs on this topic, and these do not show benefit, they concluded nonetheless that educational materials may play a role.[10] This shows how the fact that only RCT evidence is considered might lead to the realisation that this provides insufficient information to take a decision. This advice was reached by consensus, which suggests that for areas where good quality studies are rare consensus may play a big role in the decision-making process. This recommendation is not worded in very strong terms: 'No evidence of statistically significant benefit was found. However, the GDG [Guideline Development Group] agreed that educational materials may have a role.'[11] However, showing no evidence of benefit is not equal to no benefit seems to be the logic underlying this consensual position. This is a sentiment that is expressed by several guideline members as they explain that a large part of medical practice is not based on evidence at all, and an even smaller part on high-quality

10 Savigny and others (n 3) 67–68.

11 ibid.

evidence as defined by this guideline group, which comprises RCTs and systematic reviews of RCTs: 'much of what we do is not based on evidence at all, let alone based on RCTs.'[12]

Taken in this light, the advice that educational materials may have a role is not surprising, as it accords with existing practice in treating lower back pain. There is also no harm or side-effects associated with its use and in these circumstances, such a treatment can be advised despite the lack of 'evidence' of effectiveness according to the standard set by the guideline group members.

*Evidence Use for Exercise*

Advice and educational tools alone may not be sufficient and the patient may return to the healthcare practitioner still experiencing lower back pain. The next step in the English guideline is to offer exercise therapy. The advice to exercise is mainly based on the UK-BEAM trial[13], the largest trial of its sort in England.

The trial can be described as follows: a group of 1,334 patients were randomised in five groups. One group received best care by GPs which consisted of active management

---

12 English back pain guideline group member 4 (Middlesbrough, 16th of March 2015)

13 UK BEAM Trial Team, 'United Kingdom Back Pain Exercise and Manipulation (UK BEAM) Randomised Trial: Effectiveness of Physical Treatments for Back Pain in Primary Care' (2004) 329 BMJ 1377.

of the patient through advice on staying active and providing patients with a copy of the *back book*[14]. The other four groups were divided in such a way that one received exercise therapy in NHS practice, one received manual therapy in NHS practice, one received manual therapy in private practice, and one received manual therapy followed by exercise therapy. Discussion in the group centred around the question whether this RCT evidence was sufficient to recommend this treatment. The group was looking at the effect size of the treatment had in the trial. Even though the effect of exercise was shown to be small in the trial[15], according to the group there was a big enough basis to recommend exercise as an option.

When asked why this judgement had been made, the general sentiment seemed to be that even though the effects are small there is no alternative, otherwise you would end up not being able to recommend any treatment. A guideline group member explains this conclusion through an explanation of the pragmatic approach taken to the concept of 'meaningful clinical differences':

> It was a big pragmatic trial which was reasonably positive. (…) Which shows that in practice this shows to be effective so on the basis of that he [the chairman] was going to recommend manipulation [of the back]. That was because he took a pragmatic perspective, he wasn't so interested in what NICE has adopted, an insistence on a minimal important clinical

---

14 An educational booklet that gives advice on how to deal with back pain.

15 Even though there was shown to be a statistically significant reduction of pain, commentators doubt whether this was also clinically relevant.

> difference over placebo. Which I think is wrong because it is not clinically relevant, the difference over placebo.[16]

The reasoning behind this stance seems to be that there exists a tendency (in doing trials with large groups of participants) to have an average towards the mean. Which would mean that on average a participant in the trial would show a small effect but that this would obscure the effect that for a group of patients there could be a large effect and for some no effect at all. The group thus decided that despite a small average effect, exercise is an appropriate treatment because of the large group of patients with lower back pain complaints.

A further reason behind this step lies in the fact that for all treatments for lower back pain the RCT evidence, if it is available, is relatively weak in the sense that the effect sizes are relatively small. For lower back pain there is not one clear treatment that provides a 'cure.' One guideline group member talked about this dilemma in some detail:

> To a certain extent you could end up saying nothing about some of these things [treatments] because of the lack of evidence in certain areas. And that doesn't really help anybody. It is incumbent on the guideline group to give some steer, so have you give it your best shot, even in the absence of very tight evidence. More so than many other conditions. If you look at the effect

[16] English back pain guideline group member 2 (London, 20th of February 2015).

> sizes of the studies that are most often quoted [in the medical literature] they are very small.[17]

For most guideline group members, the fact that the evidence they reviewed in making their decision was not perfect did not move them to refrain from making decisions. Those decisions they argued had to be made, as this was the duty of those professionals that draft guidelines. So, they tried to reflect the uncertainty of their knowledge in the formulation of recommendations that are relatively broad, but nonetheless felt, as the quote above shows, compelled to make a decision regardless of a lack of solid evidence.

*Evidence Use for Manual Therapy*

If we compare the evidence on exercise as a treatment option with the evidence used for manual therapy the picture is very similar. The basis of the recommendation is the same UK-BEAM trial. As we have seen the trial established that manual therapy has a beneficial effect. The group therefore decided that a course of manual therapy should be offered as a treatment option. One guideline group members suggested that this might be due to a preference to keep existing treatments, and not 'upset' current practice:

> The discussion varies and the level of evidence varies because of necessity. I get a feeling that common practice has a hand in it as well because you

[17] English back pain guideline group member 3 (Interview by Phone, 27th of March 2015).

> have clinicians that represent that around the room. (…) They want to some degree reflect practice if possible.[18]

Another guideline group member disagreed with that assessment and stressed the importance of treating evidence as an 'equalizer', providing that level playing field that would limit the importance of any preconceived notions about treatments because of clinical experience:

> My job was to ensure that we were treating the evidence fairly. It was important that we were treating everything equally. Because (…) professionals don't read guidelines, they scan guidelines to see where their treatment is, and then kick off is it not there. (…) We had quite a lot of alternative therapies that were not routinely provided on the NHS, why shouldn't it be available if it works.[19]

This quote illustrates two important points about the factors that influence the decision-making process in a guideline. Firstly, it provides another example of the use of evidence not only to support the group in reaching a decision about which treatments to recommend but also which treatments to drop. Showing the necessity of a minimum of evidence before preferences based on clinical knowledge might be taken into account. The earlier decision to only allow RCT evidence thus did lead to treatments not being considered because they could not meet this threshold. This meant on the other hand that if a treatment could be shown to work, according to the definitions adopted within the group, it should be

---

18 English back pain guideline group member 2 (London, 20th of February 2015).

19 English back pain guideline group member 3 (Interview by Phone, 27th of March 2015).

recommended, no matter whether it was an alternative therapy or regular therapy on the NHS, and this would limit the influence of any clinical preferences that different members of the guideline group would have.

Secondly it illustrates the realisation by the guideline group that the guidelines will be used an assessed in a certain way be clinicians, and that they try to accommodate that potential critiques the guideline may receive. To that end it was seen as important to ascribe all the decisions in the guideline to evidence, and downplay as much as possible the influence of any other factors.

Manual therapy is defined broadly to include spinal manipulation, spinal mobilisation and massage.[20] This way of defining the treatment also has the effect of changing the organisation of care, because spinal mobilisation and manipulation are treatments provided by physiotherapists, osteopaths and chiropractors, who had a smaller role in the previous guideline. In this way the guideline does not only provide a recommendation to help practitioners choose a treatment option it also aims to change the organisation of care. However, my analysis suggests that it is not the evidence that forces these outcomes on the group rather it is the choices that are made in categorization of the evidence that provide the group with a basis for the recommendation. If manual therapy would have been defined differently, this would have changed the frame through which the

[20] Savigny and others (n 3) 103.

evidence would have been analysed. It is likely the resulting recommendation would not have been the same, even though the evidence drawn upon to support it would not have changed.

*Evidence Use for Acupuncture*

The next treatment that is recommended in the English guideline that will be discussed is acupuncture. The evidence base for acupuncture consists of four RCT's and a systematic review. One of those studies[21] is considered to provide the strongest evidence of the beneficial effect on reducing pain and reducing disability, even though the guideline group recognises that this study deviated from its protocol and had a high proportion of participants who did not complete the study. The fact that this study did have the correct population of acupuncture patients, people with back pain between 6 weeks and one year, is considered to be important. One of the guideline members stated that the RCTs conducted outside England provided stronger evidence for acupuncture than the Thomas paper:

> Yes this paper from York [the Thomas paper] was interesting. (…) in retrospect I am very pleased they did it. But when they did do it I thought this was an odd trial to do. It looked like this was propaganda for traditional acupuncture. An a+b trial, acupuncture + usual care versus acupuncture.

---

[21] Kate J Thomas and others, Longer Term Clinical and Economic Benefits of Offering Acupuncture Care to Patients with Chronic Low Back Pain (Gray 2005) <http://www.carolinashealthcare.org/documents/NorthEastInternalIntegrative/research-library/BackAndNeckPain/LongerTermClinicalEconomicBenefit.pdf> accessed 8 October 2014.

> That is only going to show a positive effect, it cannot really show anything negative. Interestingly, when they first wrote the protocol the primary outcome was at one year, when they finally wrote it up it was at two years. They had already started the trial by the time it moved. At 1 year the result was not significant. They had done 12 weeks of treatment at the beginning of that year not more. It became significant at the 2 year mark. Which I have never seen before in an acupuncture study. (...)It looks to me like instead of the acupuncture group improving the control group deteriorated more [and thus causing a significant effect after 2 years].

Clearly this guideline group member has doubts about the validity of the results of this study, in particular querying the result becoming significant at two years, while it was not at one year, cannot be a result of the treatment, because the acupuncture treatment only lasted 12 weeks. This guideline group member was however in favour of providing acupuncture based on German studies:

> I wasn't going to raise any issues, there was already good data from Germany that agreed in terms of cost-effectiveness. The GDG was not interested in the German data because it was outside the NHS. I was not going to say anything [about his doubts about the English acupuncture study] because I knew that the German data [was] much bigger. So I thought that was more reliable. (...) They [the guideline group members] wouldn't have known the details of this [the English] study. [22]

Most of the guideline members preferred the evidence gathered within the NHS to evidence from other healthcare systems, even if this evidence – in their view – was more convincing. This illustrates a tendency for guideline group members to give more weight

[22] English back pain guideline group member 2 (London, 20th of February 2015).

to evidence from their own country. This is understandable as they might be more familiar with the research – or as in this case have participated in the trials themselves – when this is conducted in their own country. However, this has the effect of further limiting/restricting what is considered as evidence.

*Evidence for Spinal Fusion*

The guideline also addressed spinal fusion, which is a surgical procedure aim to alleviate back pain. The guideline did not recommend spinal fusion directly, because surgical procedures where not in the scope of the guideline. Rather it advises practitioners to consider a referral when other treatments have failed. There was not a great deal of evidence supporting this treatment, but in this sense the treatment was not very different from exercise or manipulation that also relied on only a handful of studies.[23]

[23] A meta-analysis of three RCT's T Ibrahim, IM Tleyjeh and O Gabbar, 'Surgical versus Non-Surgical Treatment of Chronic Low Back Pain: A Meta-Analysis of Randomised Trials' (2008) 32 International Orthopaedics 107. And a systematic review of the same studies and 1 further RCT. Sohail K Mirza and Richard A Deyo, 'Systematic Review of Randomized Trials Comparing Lumbar Fusion Surgery to Nonoperative Care for Treatment of Chronic Back Pain': (2007) 32 Spine 816.

## 4.3 Achieving Consensus: Internal Workings of Guideline Groups

The English guideline group considers and debates evidence and many different reasonable opinions can be formed about the same scientific study by different members of the group. The process of arriving at recommendations is therefore more complicated than a simple mechanical application of a given result of a scientific study to a condition or the tallying up of the positive results of studies. The recommendations that are agreed on are the result of a debate among the members of the guideline group. This debate can take the form of questions about the methodological soundness of the scientific research, the clinical relevance of the scientific research, the acceptability of the choices to the outside world, such as government and healthcare practitioners, and its effects on the organisation of healthcare. This does not mean that guideline groups always invoke of all of these considerations/factors all the time. Some of the argumentation in the group is concerned with the more general questions how to approach the medical studies that are used in the discussion in the group, and the difficulties in matching results of clinical trials with the clinical practice of treating people with lower back pain:

> That's a standard [chronic back pain at 12 weeks], it's an epidemiological conceit that is put forward. If you look at the natural history of back pain it is a recurrent problem. Nobody is acute and then magically becomes chronic. It's a conceit that is used that causes all sorts of problems it's a convenience. (…) It is very very very rare for someone to present [to a GP] with a first ever case [of acute back pain] But when you are looking at the evidence if people say we only took people who had back pain for less than

> 12 weeks it is very difficult to appraise whether they actually did because almost by the nature of the condition people have had it before and this is probably the first time they consulted. It often goes from first consultation not from first report (…) these sort of things do muddy the water when you are trying to extract the data.[24]

This member of the guideline group presents his particular interpretation of a more general struggle for members of this guideline group to interpret medical studies within the context of the treatment of lower back pain. He points out how in the interpretation of medical studies one cannot be sure that the population of patients treated can really be meaningfully classified as either 'acute' or 'chronic'. This shows how in interpreting medical studies, and attaching values to the results of these studies the guideline group uses their collective knowledge and understanding of the nature of the disease (almost everyone has repeated episodes of back pain) and knowledge of the current state of treatment to arrive at acceptable recommendations.

Guideline group members also discussed in more general terms, sometimes explicitly, sometimes implicitly, the approach to treating lower back pain, and the underlying message they wanted to convey with the guideline:

> We are still very much on to a structural approach to a lower back pain, and now we moved to a cognitive approach, there is a new approach that [shows] it is a neurological condition. (...) It's not physical. The core component of the NICE guideline was self-management with reassurance, taking part in an active lifestyle, but the papers weren't published at the

[24] English back pain guideline group member 3 (Interview by Phone, 27th of March 2015).

> time. So we would have said something very different in two years. We would have loved to have said it because that was how we were practising at the time.[25]

This guideline group member, when asked about the underlying purpose of the guideline, spoke at length about the tension between those treatments where evidence could be presented for, and the 'softer' interventions, based on changes in lifestyle, where there was no evidence yet to consider although some members of the group felt they were more important. This shows how the absence of evidence can sometimes lead to the result that certain approaches or treatments cannot be included in the guideline. To introduce a new approach without any support in existing evidence does not fit in the NICE approach to developing guidelines.

The discussion moved on to consider the treatments in greater detail. This part of the discussion was structured early on by the decision to only accept RCTs as evidence, and this inevitably led to the need for 'consensus' when evidence was lacking, as the group had agreed they would not accept evidence other than RCTs. This also means that when the group needed to decide whether one treatment is better than the other, they needed to restrict this discussion to weighing the available RCT evidence. Seen in this light, the choice to regard manual therapy, exercise and acupuncture as equally valid options, and to give patients and doctors a choice in what to prescribe, enabled all parties involved to be

---

[25] English back pain guideline group member 3 (Interview by Phone, 27th of March 2015).

content with the guidelines. Conflicts based on the place of a treatment in the pathway for treating lower back pain could thus be avoided. Guideline group members placed these choices consistently in the frame of evidence:

> There was a possible choice of physical treatment, we couldn't clearly say that any one of them was better than the other. The only thing that we found there were a few studies suggested that combining them would be better than one by one. You could [not] really separate them out on head to head trials, (…) there was no difference in QALY's. It is not intended to give clinicians freedom, NICE is keen that the patient is involved in care, so this is an area where patient preference is an issue. [26]

Even though these discussions seem to be about evidence, it is plausible that part of the reasoning in the group is aimed at finding a position that as many members as possible can agree with. The fact that the medical studies do not seem to support a large difference between the various treatments makes it easy to reach the consensus that a choice between these treatments is appropriate.

The discussion on spinal fusion differed from those about the other treatments. Providing spinal fusion as a treatment was not part of the agreed scope for this guideline, as the scope excluded surgical procedures. The group could advise, however, about referral

---

[26] English back pain guideline group member 3 (Interview by Phone, 27th of March 2015).

to spinal fusion and they did so by advising that a referral could be considered for those that had tried all the previous treatments without success.

With the decision to make spinal fusion available for a select group of patients the group made a pragmatic choice. Given the fact that the guideline only deals with back pain that is shorter in duration then one year, the expectation was that very few patients would complete the pathway and be eligible for spinal fusion. This is because the guideline makers have defined the group of patients that could be referred for an opinion on spinal fusion as only those who have unsuccessfully have completed the other earlier steps in the guideline. Some members were critical of the inclusion of spinal fusion: 'Despite rather a lack of evidence it did get in the guideline.'[27], but a majority of the participants supported this decision.

It is likely that part of the motivation for this decision was, although none of the participants said this in so many words, to make sure that those in the group representing the interests of surgeons would agree with the guideline. If the guideline was written for those with chronic back pain, perhaps a different decision would potentially have been made, given the fact that patients with chronic back (for longer than a year) are more likely

[27] English back pain guideline group member 2 (London, 20th of February 2015).

to be referred for an opinion on spinal fusion than people who had lower back pain for a shorter duration.

What we can take from these various discussions in the guidelines group is that evidence was a guiding factor in making decisions in this guideline, a minimum threshold of evidence has to be passed to be able to include a treatment, but the decision about where to place a treatment in a pathway was based on clinical reasoning and political considerations that existed outside of the evidence. This is illustrated by the criticism this guideline received after adoption. The first main criticism of the English guideline was of the decision not to recommend injections for the back[28]. A further criticism was that this guideline only recommended treatments that one of the members of the guideline group provided.[29] These critics suggest that the guideline group was not basing their decisions solely on evidence but also leaned on achieving consensus in the group, and thus that treatments that were not 'represented' within the group were not recommended because they were not necessary to reach consensus. This criticism highlights the nature of the compromises a guideline group has to make, and the unenviable and difficult position it finds itself in. The fact that internal consensus is required in the group to develop a

---

28 Joan B Hester, 'Official Statement from the British Pain Society' [2015] The BMJ <http://www.bmj.com/rapid-response/2011/11/02/official-statement-british-pain-society> accessed 14 September 2015.

29 Christian Egeler, 'NICE Guidline on Backpain Is Misleading | The BMJ' <http://www.bmj.com.proxy-ub.rug.nl/rapid-response/2011/11/02/nice-guidline-backpain-misleading> accessed 4 July 2016.

guideline, can lead to external critique. If, however, the group decides to prioritise the expected critique of external partners over the need to reach consensus in the group, drafting a successful guideline will become difficult. A way out of this dilemma can sometimes be found in a very general formulation of recommendations, so general that all relevant parties can agree with (part of) it.[30] In this case the group responded to the criticism by pointing to the high level of evidence used for the guideline - only Randomised Control Trials – to support its choices.

Another dimension of this process of development of a guideline is that it might affect the organisation of care. When evidence is used to recommend treatments, it is unavoidable that something is said about who performs those treatments and where (in which healthcare setting): even if guideline makers do not set out with the intention of reorganising practice. This is one reason why it is difficult to strictly separate evidence-based decision from consensus-based decisions, and 'internal decisions' from external effects.

## 4.4 External Responses: Rejection of the Guideline

So far, we have considered the internal workings of the English guideline group on lower back pain. There are two ways in which external responses to a guideline might potentially

[30] An example of this approach is the Dutch guideline on lower back pain.

influence a guidelines development. After the guideline has been drafted an official period of consultation follows. During this consultation registered stakeholders can respond to the guideline, and this might lead to changes in the guideline. A further, less obvious way that external engagement might influence a guideline is through the *expectation* of critique by members of the group. If members of the guideline group already expect that certain decisions will be controversial they might try to avoid these decisions. In this way the guideline might be influenced by external parties, without this necessarily being visible in the text of the guideline.

It is likely that guideline makers might sometimes try to pre-empt critique on the guideline, as this will likely improve the acceptability and implementation of the guideline. A different way to put this is to assume that if the guideline drafters are correct in their assessment of the reactions of their peers the guideline will be uncontroversial, and importantly the document could safely be considered technical and thus free of political decisions. The development of the lower back pain guideline shows how both potential paths of influence can play a role.

The experience of the English guideline on lower back pain shows that guideline development is not always perceived as technical, and that the choices that are made are perceived by some as political and can be contested. In this case these perceived political choices do not lie in the decision to recommend a treatment, but rather in the decision not to recommend a treatment. This was indeed the case when the English guideline group

decided there was insufficient evidence to recommend injections in the back. The relevant section of the guideline is as follows:

> Searches were carried out to identify any form of injection for the lower back, however, only data on facet-joint, prolotherapy and intradiscal injections was identified. The GDG agreed that there was a lack of evidence to recommend the use of these treatments and agreed by consensus injections were of no benefit for this population.[31]

Why was this decision so controversial? The treatment that the guideline does not recommend is performed by a variety of healthcare professionals that are represented by the British Pain Society (BPS). As a result of being excluded the BPS were concerned that they will lose a substantial number of patients and funding because commissioners of healthcare would no longer fund the therapeutic injections in the back they provide. The BPS claimed that there was sufficient evidence to warrant injections in the back if one takes cohort studies and clinical case studies into account. They criticized the exclusion of this evidence in the guideline. A quote from their letter to the British Medical Journal (BMJ) illustrates this:

> 'We cannot agree with the exclusion of all injections for back pain. The BPS provided evidence to support the clinical effectiveness of such procedures. The Guideline Development Group (GDG) did not consider evidence presented from cohort studies and clinical case series in its deliberations on these, or any other, treatments. We think this was misguided. Although NICE has recommended further research into these procedures, we are most

---

[31] Savigny and others (n 3) 162–163.

> concerned that a significant number of patients will be denied this choice of treatment in the interim.'[32]

This statement taken at face value would be a discussion about evidence. The BPS was proposing to define evidence differently: they assume that a more inclusive definition of evidence will provide more support for their treatment. There is thus a contrast between the guideline group, who has chosen a narrow/strict definition of evidence, only RCTs, and the BPS who proposes a wider definition. There is more to this statement then just a complaint about evidence. The Society was mainly worried about the redefinition of the pathway of treating people with lower back pain. It is an issue of defining, and redefining professional boundaries through guidelines. Wilson and others in their analysis of the 'fall-out' of this guideline point this out as well and phrase this as follows: 'A key discourse emerged that enabled boundary work by doctors in interventional pain medicine – activity designed to protect and expand jurisdiction, assert professional authority, claim resources and protect autonomy.'[33]

Medical guidelines in this reading by Wilson et al. can thus be seen as a potential tool to reorganise professionals and so act as an instrument of control: they try to discipline medical professionals and to steer them. It is not surprising, given the interdisciplinary

---

32 Hester (n 28).

33 Nicky Wilson and others, 'Governing Healthcare: Finding Meaning in a Clinical Practice Guideline for the Management of Non-Specific Low Back Pain' (2014) 102 Social Science & Medicine 138, 144.

treatment of guidelines within NICE, that choices in guidelines will inevitably have some inter- and intra-professional effects. The importance of describing the BPS critique in detail lies in the fact that this is not a one-way street. In this case the guideline did not simply reorganise healthcare and professionals passively accept the change, rather they resisted and the way they chose to resist in this case provides an example of the limits of guidelines as regulatory tools.

The controversy did not end with the letter by the BPS, this proved to be just the starting point. The BPS, on the request of several members, took a vote at an extraordinary general meeting to oust their president. He was a member of the guideline development group, but not himself a pain specialist and tasked with providing independent expertise, and they stated that he had not served their interest well by defending the NICE guideline, while the society asked for the guidelines to be modified or withdrawn.[34] NICE reacted angrily to this development, it emphasised that the president of the BPS was not representing the society during his involvement in the guideline development but was expected to act as an independent expert. NICE feared that medical professionals would no longer participate in guideline groups if they felt pressured by outside groups.[35]

---

34 Wilson and others (n 33); Zosia Kmietowicz, 'President of British Pain Society Is Forced from Office after Refusing to Denounce NICE Guidance on Low Back Pain' (2009) 339 BMJ b3049.

35 'NICE Outraged by Ousting of BPS President | The BMJ' <http://www.bmj.com/rapid-response/2011/11/02/nice-outraged-ousting-bps-president> accessed 17 September 2015.

Nonetheless NICE did agree that a consultant in pain medicine needed to be included in the guideline group in the future.[36] In fact a consultant in pain medicine is now the lead on the new lower back pain guideline. The outside pressure on NICE has thus been successful and professional groups that felt excluded from the process are now included. The new 2016 guideline also doesn't recommend therapeutic injections in the back but does recommend radiofrequency denervation in certain cases. In addition, in contrast to the earlier guideline, it does not recommend acupuncture nor spinal fusion.[37]

What do these developments mean for the lower back pain guideline? How did this controversy affect the development of text of the guideline? Despite this controversy the text of the guideline was not changed as a result. The guideline group and NICE rejected these demands for change. So, in this sense the guideline was not affected by these external forces. The discussions in the guideline do, however, support the view that there might be an indirect influence. The decision by the guideline group to only use RCTs as evidence was partly taken because criticism of the guideline was expected. In their response to the criticism the guideline group also uses this choice as a defence and justification for maintaining their original position.

---

36 Wilson and others (n 33).

37 https://www.nice.org.uk/guidance/NG59/chapter/Recommendations#invasive-treatments-for-low-back-pain-and-sciatica.

All in all, in this case the engagement of external parties in the development process of guidelines only has a very limited influence on the text of the final guideline discussed in this thesis. It did clearly affect the text of the following, 2016, guideline in this sense the British Pain Society was successful in the long run, showing how the struggle to define professional boundaries is a long game.

## 4.5 Conflicts of Interest: Limited Relevance

One of the criticisms voiced about guidelines is the potential influence of the pharmaceutical industry on their creation.[38] This literature suggests that because of direct financial benefits, or indirect non-financial benefits (such as reputation, and paid and unpaid positions) guideline group members might be unduly swayed to favour certain drugs or treatments over others.

One potential avenue to establish whether the development of the guideline might potentially be affected by a conflict of interest of one or more of the members of the

---

38 Ben Goldacre and Carl Heneghan, 'How Medicine Is Broken, and How We Can Fix It' (2015) 350 The BMJ h3397; Ben Goldacre, *Bad Pharma: How Drug Companies Mislead Doctors and Harm Patients* (Fourth Estate 2012); Stelfox Henry Thomas, Choudhry Niteesh K and Allen S Detsky, 'Relationships between Authors of Clinical Practice Guidelines and the Pharmaceutical Industry' (2002) 287 JAMA 612; David Healy, *Let Them Eat Prozac: The Unhealthy Relationship Between the Pharmaceutical Industry and Depression* (1 edition, NYU Press 2006); Hans Melander and others, 'Evidence b(i)Ased Medicine—selective Reporting from Studies Sponsored by Pharmaceutical Industry: Review of Studies in New Drug Applications' (2003) 326 BMJ 1171; Gisela Schott and others, 'Does the Pharmaceutical Industry Influence Guidelines?: Two Examples from Germany' (2013) 110 Deutsches Ärzteblatt International 575.

guideline group is to look at pharmaceutical interventions in the guideline. The few pharmaceutical interventions that are advised are all off patent, and are no longer actively marketed by pharmaceutical companies. This study has found no evidence that a potential conflict of interest played any role in the decisions on pharmaceutical interventions in this guideline. This does not necessarily mean that the construction of the guideline was free of potential conflicts of interest. Potential conflicts of interest in this guideline could be related to other elements of the guideline development process.

A conflict of interest may exist when the people who conduct the RCTs are used as evidence, in the guideline, partly overlap with the people in the guideline group who have to critically asses that evidence. The fact that the people who conducted studies have to decide on including studies as evidence creates a potential conflict of interest, because they have to critically evaluate the weight of evidence they have themselves a vested interest in, being the authors of the work in question. At the same time their expertise, as authors, might be necessary and beneficial to correctly interpret the evidence.

An example in this guideline is the BEAM-study, which is cited as important evidence for some of the advised treatments for lower back pain. The fact that one of the guideline members was involved in co-authoring and designing this scientific study might lead to this evidence being favoured over other potentially stronger conflicting evidence. There are several mechanisms to minimize the risk of this happening built in to the guideline process: the fact that all the evidence is first assessed systematically by the

technical team, and given ratings prevents someone from relying on evidence that is not of sufficient quality. Furthermore, the fact that a large group of people debate this evidence also reduces the risk of undue influence by one person on the evidence base. If we consider the declarations made by the guideline group[39] they also list the authorship of articles on various aspects of lower back pain as potential conflicts of interest, so that the kind of scrutiny applied in this analysis is indeed possible.

This potential 'cognitive bias' by members of the guideline has not been supported in the interviews with guideline group members, or analysis of the guideline document. The fact that some studies are (co)authored by those that use them as evidence has not influenced the guideline. In sum, no support has been found in this study for the assumption that conflicts of interest have influenced the development of the lower back pain guideline.

## 4.6 Guideline Drafting in Institutions: Success Within Constraints

Through the study of the institutional context of guideline making we can more fully understand potential limitations and culturally dependent restrictions on the process of

[39] Savigny and others (n 3) Appendix F.

guideline development that are the are the result of influences from the government, healthcare professionals or the guideline making institution itself.

*Top-down Government Pressures*

If we consider the top-down pressures first, the remit that the UK Department of Health sets for NICE seems to be a form of direct government intervention in guideline making. In theory by setting the remit the government can control the topics which NICE has to make guidelines on. In practice this is not the case as NICE itself has identified/ established a large library of guidelines they want to produce. So the remits in reality are set by NICE and no active intervention by the government takes place[40] A senior official in NICE explained this in more detail: 'No those remits are discussed with the department, but [now] we have so many topics that we want to make guidelines on. We have put them in a library and NICE decides through its work plans how to prioritise the topics in the library.'[41]

While there is no direct government intervention in guidelines, there is an expectation that care will be provided according to NICE guidelines, although this is not a strictly legally binding requirement. Clinical Commission Groups (CCG's) have to take technology appraisals into account and provide funding for those technologies

---

40 The exception are the recent safe staffing guidelines, they however don't form part of this study.

41 Senior NICE official (London, 22th of May 2015).

recommended but not for treatments recommended in guidelines.[42] Given this background it is not surprising NICE guidelines make extensive use of cost-benefit analysis using detailed breakdowns of Quality Adjusted Life-Years (QALY's) to be able to recommend certain treatments. Given the expectation (but as noted not a legally binding requirement) that recommending a treatment will lead to funding of a treatment the guideline groups have a duty to analyse the likely costs for the NHS of their recommendations. In the case of the English lower back pain guideline there is no evidence that this has had any effect on the recommendations as all advised treatments have a QALY that is below the threshold identified by NICE.[43]

*Bottom-up Stakeholder Pressures*

What influence, if any, did pressures from healthcare professionals have on the development of the guidelines? In the case of guidelines produced by NICE, comments on the drafts of guidelines can only be given by registered stakeholders.[44] Those registered stakeholders are NHS trusts, pharmaceutical companies, and scientific and professional associations of healthcare professionals. For the English guideline on lower back pain, in addition to these opportunities for influence, the professional associations were also

---

42 The National Institute for Health and Care Excellence (Constitution and Functions) and the Health and Social Care Information Centre (Functions) Regulations 2013.

43 Apart from spinal fusion if a decision was made for immediate surgery, this was however not recommended in the guideline.

44 https://www.nice.org.uk/get-involved/Stakeholder-Registration

directly involved in selecting the chair and the members of the guideline group. Since then the procedures have changed and NICE advertises these positions through open competition. Through these channels healthcare professionals can try to influence the content of guidelines.

We have already seen that through this process the British Pain Society was very vocal in their opposition to the guideline. They were vehemently opposed to the guideline but despite all this protest the text of the guideline did not change as a result of their representations.

*The Institutional Context of the Guideline*

An important element of the development of guidelines, which came to the fore in this case study is the specific institutional environment within a guideline making institution. Although this guideline was written for NICE, but the guideline group was nominally independent from NICE control, and had complete freedom to decide on the text of the guideline. There is nothing to suggest NICE did not respect this independence in practice. In fact, the opposite seems true, they maintained scrupulous barriers between the guideline making process and their internal procedures.

There is, however, the institutional culture and procedures of NICE that does influence and restrict the actions of the lower back pain guideline group. Two aspects of this institutional environment are particularly relevant. Firstly, NICE has developed a very

detailed guideline manual that describes in detail how to consider evidence and what elements to consider in drafting guidelines and it employs a technical team that applies this manual. This guideline manual has had a direct effect on the functioning of the guideline group and the development of the guideline. Because of these procedural rules NICE could control the scope of the guideline. We have seen that some members of the guideline group complained about this scope, or at least afterwards expressed regret that they did not get the opportunity to make an 'integrated' guideline that covered lower back pain fully.

It is not surprising that NICE wanted to control the scope of the guideline, it has a vested interest in making sure that its guidelines are consistent, and it has only limited resources. These choices did affect guideline making groups, despite their independence being preserved are given a task within set boundaries. The influence of NICE is not restricted to the scope NICE also influences who the members of the guideline group will be. The scope identifies in which areas the guideline group will need expertise and members of the guideline group are subsequently sought out who have this expertise. This means that NICE can include certain experts but also exclude them, and could be criticised for those decisions. As happened in this case, as the discussion in section 4.4 on the exclusion of a dedicated pain specialist from the guideline group shows.

Second, in addition to these elements of influence of the institutional rules of NICE another element is the operation of the technical team within and with a guideline group. It is this technical team that searches the medical literature looking for medical studies that

are then classified as counting as evidence or not. Constructing what counts as evidence is therefore an interplay between the guideline group and the technical team. While the guideline group applies a 'broad brush' by asking for only RCTs to be included, the technical team fills in the details and decides which studies should be classified in that way, and whether they fulfil the criteria to be included. In this way the institutional rules and procedures of NICE prescribes and circumscribes what counts as 'evidence' for NICE guidelines.

NICE procedures are explicitly aimed at ensuring that guidelines are based on the highest quality of evidence available and they invest a great deal of money in enabling the systematic searching and appraising of large amount of scientific evidence. The paradoxical result is that because of the diligent and high-quality appraisal process, as done by the technical team, the evidence base used for the recommendations is often small, as only a few studies make it through this rigorous process (at least as far as back pain is concerned).

*Institutional Influences Considered*

When considering the various institutional factors affecting the development of guidelines it is noteworthy that the most prominent factors influencing guidelines are those *within* the guideline making institution. The development of guidelines is only marginally affected by pressures from government or external critique from professionals. In contrast, internal

cultural norms and procedures through their influence on the medical studies used to make decisions appear to directly shape the framing and selection of evidence used in the development of the guidelines.

*Costs: A Complicating Factor for the English Guidelines*

Guidelines drafted by NICE include an assessment of cost-effectiveness of the treatments proposed in the NICE guidelines. In the NICE guidelines the cost of treatments are expressed in QALY's (quality-adjusted life years).[45] Without providing a great amount of technical detail as to the calculation of these QALY's for our purposes it is important to note that NICE has defined a threshold of £20,000 below which a treatment would be considered cost-effective. It is important to consider cost effectiveness, as they do feature in the discussions of the members of the guideline group, and so potentially lead to decisions being changed because of these considerations.

Given the fact that the cost of most treatments in the guideline are well below the threshold of £20,000 per life year gained, a treatment below the threshold could be recommended without having to balance the evidence on clinical effectiveness with that of cost-effectiveness. This was summed up succinctly by one of the guideline group members:

---

45 For more details on QALY's see: 'Guide to the Methods of Technology Appraisal 2013 |' <https://www.nice.org.uk/article/PMG9/> accessed 4 July 2016.

‘Most of the treatments are relatively cheap so you not going to break the bank if you recommend it. And patients expect some sort of treatment.’[46]

The only treatment for which the guideline group did have to consider the cost-effectiveness was spinal fusion because the economic modelling performed showed that if this treatment was given to all patients the treatment would be more expensive than £20,000 per QALY.[47] The guideline did not have to confront this issue directly as the scope of the guideline did not include the power to recommend surgical procedures. Instead the recommendation made in the guideline was one on referral. And this was only deemed appropriate when a patient would have followed all the previous steps in the guideline without success. Given these constraints on access there was no reason the exclude spinal fusion, given the expectation that the uptake – in the confines of the given guideline scope – would be marginal, and therefore the costs of this option would also not be too high.

This interplay between cost considerations as expressed in the economic models used by NICE and the consideration of clinical relevance or quality considerations is strong in NICE guidelines. In every assessment of evidence, a reference is made to the costs of treatment. For the reasons given, in the case of the lower back guidelines it did not

---

46 English back pain guideline group member 2 (London, 20th of February 2015).

47 Savigny and others (n 3) 217.

materially change the recommendation, it is however a factor that potentially influences the development of other guidelines.

## 4.7 The Dutch Guideline on Lower Back Pain

The Dutch guideline on lower back pain that is the focus of this study is the one made by the NHG in 2005.[48] This guideline is made by a group of GPs, epidemiologists and healthcare researchers. The members were approached informally by the NHG, based on their knowledge of the topic and preference was given to those GPs with a doctorate. At least one of the guideline group members put himself forward as he explains: 'because this was a GP guideline [I would not normally be approached] so I decided to put myself forward, when I knew that a new guideline update was being considered.'[49] The final group consisted of several GPs, a back-pain specialist, an epidemiologist and a healthcare researcher. The guideline process was supported by a scientific editor from the NHG who was responsible the drafting of the guideline of versions of the guideline for discussion among the guideline group members.

---

48 'NHG-Standaard Aspecifieke Lagerugpijn | NHG' <https://www.nhg.org/standaarden/volledig/nhg-standaard-aspecifieke-lagerugpijn> accessed 21 August 2014.

49 Dutch back pain guideline group member 3 (Rotterdam, 5th of December 2014)

The guideline group members all read and appraised the evidence literature themselves, with no predefined restrictions on type of literature they could consider. Several group members committed on how this process was conducted:

> We had to search literature ourselves, and provide a summary for the group. Everyone did a specific task. We ranked [the evidence] using common sense. We followed the evidence-based ranking, with meta-analysis having the highest weight and the lowest when there is no evidence. We [would] discuss the statement about the evidence. And then given the evidence [we would discuss] what is the recommendation.[50]

This guideline group member explained how the group has set out a plan which involved guideline group members, based on their own preferences, setting out to find literature on specific aspects of the guideline (diagnostic, treatments etc.) and then turning the evidence found into a summary for the group to discuss. This group member suggested that in general the evidence-based pyramid would be the starting point in undertaking such a search. Another guideline group member confirmed this, describing the process in the following way:

> They [the NHG] have a guideline on how to make guidelines, it was based on the EBRO. The stress was to define a question, and to search according to the well-known evidence based pyramid. This was the starting point. So we tried to develop a couple of questions, and that's why the [NHG] invite

[50] Dutch back pain guideline group member 3 (Rotterdam, 5th of December 2014)

> people in those fields, and they also rely on those people, not on their decisions, but their knowledge of the medical field.[51]

As this guideline group member points out, the fact that the guideline group members were asked to do searches in the literature themselves led the NHG to the decision to also invite back pain specialists from other medical disciplines. These members could then be trusted to know the relevant literature in their field well and support the group in finding and selecting it.

The guideline includes both acute, sub-acute as well as chronic back pain in its remit. The guideline is restricted to the treatment of back pain by GPs and referrals made by them. This restriction is seen as self-evident by all participants: 'The common interventions that where in the field of GP-care that's what you focus on.'[52]; 'It [the guideline] is really for the support of the GP, that's the bottom line.'[53]

---

51 Dutch back pain guideline group member 2 (Amsterdam, 5th of December 2014)

52 Dutch back pain guideline group member 3 (Rotterdam, 5th of December 2014))

53 Dutch back pain guideline group member 2 (Amsterdam, 5th of December 2014)

The reason for initiating this update of the original NHG back pain guideline was because of the regular process of updating guidelines by the NHG, where guidelines are reviewed every 3 to 5 years to see whether they need updating.

The Dutch guidelines on lower back pain will be assessed using the analytical questions used in the assessment of the English guideline: focusing on evidence, consensus, external responses, conflicts of interest and institutional factors.

## 4.8 Evidence-base for Dutch Lower Back Pain Treatments

The Dutch guideline presents several pharmacological and non-pharmacological treatment options for lower back pain. If we survey the options that the Dutch guideline presents under the heading of sub-acute and chronic back pain the following treatment options: education, exercise therapy, cognitive behavioural therapy and for chronic pain referral to a specialised multidisciplinary centre, and pharmacological treatments with paracetamol, NSAIDs and weak opioids. On first look the treatments presented in this guideline are on some points similar to English treatments, while other aspects of the treatment regime are different. Behind these similarities and differences there is a difference in the use and treatment of evidence, this is highlighted by the use of ‘narrative footnotes’ that are part of

the guideline that contain the assessment of evidence underlying the different treatment options.

*Evidence Use for Educational Advice*

In the Dutch guideline the first step after diagnosing a patient with back pain should be providing education on how to deal with back pain. The guideline refers to a study with the English 'Back Book'[54] that shows – in their words – 'that use of this material lead to an improved understanding of the condition, and less physical limitations.'[55] They recommend providing patients with the Dutch translation of the 'Back Book'. The fact that the study they cite to support this recommendation does not show *statistically significant* improvements in pain or disability is not mentioned. This contrasts with the focus in the English guideline on the fact that this study showed no statistically significant benefits on pain and disability and that *nonetheless* educational materials (like the 'Back Book') could be recommended.

---

54 A Kim Burton and others, 'Information and Advice to Patients with Back Pain Can Have a Positive Effect: A Randomized Controlled Trial of a Novel Educational Booklet in Primary Care' (1999) 24 Spine 2484.

55 'NHG-Standaard Aspecifieke Lagerugpijn | NHG' (n 48).

This example helps to understand an important element of the use of evidence in guideline making. Here we have a case of two groups using exactly the same evidence[56] to reach the same conclusion: the recommendation to use educational materials. They base this on two very different readings of that evidence, which provides support for the contention that it is not the evidence but is the medical study itself that presents a finding. Rather it is in according meaning to some aspects of the study, to the exclusion of others that the Dutch guideline group draws a specific conclusion from the available material.

It is important to stress here once again that this study does not suggest that any of the interpretations of the evidence are better or worse, they are simply different. They are also necessarily different because the context is different, and the questions are different. If you ask different clinical questions when reading the same medical study you will likely draw different conclusions.

In this case the final recommendation in the two countries is for all intents and purposes the same but the Dutch conclusion is 'evidence-based', the other is consensus-based. The paradox is that the English guideline group could suggest the Dutch group was wrong to use this study as evidence because it provided no statistically significant conclusions, while the Dutch group could suggest the use of consensus was unnecessary

---

56 This is strictly speaking not completely correct as both groups cited two other studies (that differ from each other) to support their conclusions, but their assessment of those studies is not analysed here.

because the study provided clinically relevant results. This in a small way shows how evidence in itself only acquires meaning through a process of interpretation, it is not evidence of anything until it is deemed to fit within a frame of understanding, and this frame is created by guideline groups.

*Evidence for Therapy*

The next treatment that can be offered in the Dutch system, if education is not successful and the pain remains, is therapy. Therapy is described broadly as including exercise therapy, cognitive behavioural therapy and manual therapy. A preference is expressed for therapies that activate the patient - massage and other passive therapies are discouraged.[57] The guideline group states that it is unclear which specific exercise or cognitive therapy is effective so suggests that the wishes and choices of the patient should play a role and a varied exercise programme should be chosen.[58] The guideline group feels it cannot make any specific recommendations about therapies to provide. Partly this choice might be due to the limited evidence the guideline has to rely on. They cite four systematic reviews on the topic that all can only give quite generic recommendations on appropriate treatments.[59]

---

57 'NHG-Standaard Aspecifieke Lagerugpijn | NHG' (n 48).

58 ibid.

59 ibid.

A guideline group member remarked on their use of Cochrane reviews[60] on those therapies in relation to evidence-based medicine:

> Cochrane reviews were considered the best evidence available, (…) [they have a] transparent methodology, other reviews had lower weight. (…) We then discuss the conclusion of the review, it often showed a small effect, and we then took this finding to be true.[61]

This use of systematic reviews was adopted as a good starting pointing for the group as they lacked the time and resources to review individual studies, so reading a systematic review of groups of studies would reduce time. In addition, Cochrane Reviews had (and have) a good name and are seen as a trustworthy source of medical data.

Nothing would stop the group from making more specific recommendations, based on consensus among them if they feel the evidence is lacking. That they did not choose to do so was because it was not possible to reach consensus on specific forms of treatment. One of the guideline group members summarized the difficulty of recommending a specific therapy for back pain in the following terms: ‘doing something is better than doing nothing.

---

60 Cochrane Reviews are a specific form of systematic review, provided by the Cochrane collaboration which is perceived as an independent and authoritative provider of systematic reviews in medicine.

61 Dutch back pain guideline group member 3 (Rotterdam, 5th of December 2014)

If you want to give a succinct summary of the literature that is about the shortest summary you could give.'[62]

Another guideline group also pointed to the importance of existing practice in the field of back pain, to say that there was no sufficient evidence to exclude any treatments:

> We also looked what is common in the Netherlands: to refer to the physiotherapist or exercise therapist Mensendick.[63] If locally a GP works with a Mensendick therapist that is fine, because the evidence doesn't really discriminate [between different forms of exercise therapy].[64]

In reality there were large disagreements between the members in the group about the best therapy/treatment. So much so that two members decided to leave the group during the guideline formation process. One of the guideline members commented as follows: 'the discussions in the beginning were very animated. Everyone defending his own favourite treatment. At some point two of the members decided for themselves they didn't want to continue and left the group.'[65]

---

62 Dutch back pain guideline group member 2 (Amsterdam, 5th of December 2014))

63 A physiotherapist who uses this specific type of exercise therapy.

64 Dutch back pain guideline group member 3 (Rotterdam, 5th of December 2014)

65 Dutch back pain guideline group member 1. (Utrecht, 28th of November 2014)

Eventually the group could find a common consensus by presenting treatments in a generic way and not prescribing that one treatment was preferable over another. In this way all members of the group could agree with the choices made. A guideline group member explains that taking a more high-level view of treatments could solve some of the disagreements: 'we just need to know do we have to refer or not, and what the physiotherapist does [does not matter because] he has his own guideline for that.'[66]

Leaving the GP and the patient a free hand as to what type of exercise they prefer but diminishes the level of 'guidance' the guideline provides. This was necessary to reach consensus in this group, and it reflected the fact that the available medical studies did not provide a sufficient basis to recommend more *concrete* therapies, in the opinion of the members of the guideline group.

*Evidence Use for Chronic Back Pain*

A similar choice was made in the recommendations for chronic pain. The guideline group recommended referral to a specialised multidisciplinary centre if other interventions are not successful.[67] That choice is also a pragmatic one based on consensus in the group. The evidence for the treatments given in those centres was considered to be unconvincing and

66 Dutch back pain guideline group member 2 (Amsterdam, 5th of December 2014)

67 'NHG-Standaard Aspecifieke Lagerugpijn | NHG' (n 48).

no recommendation could be made for either corsets, injections in the back or spinal fusion. The fact that this guideline was drawn up by GP to be used by GPs, helps them avoid the controversy of having to define the pathway after referral. The fact that the evidence for the treatments they refer patients too is not convincing does not affect those groups of healthcare professionals directly. So, spinal surgeons, people specialising in pain medicine etc., are not affected by this guideline. When asked why the group choose 'multidisciplinary interventions' and referral to a multidisciplinary team as treatment options for chronic back pain, it became clear that this choice allowed enough flexibility for all parties to interpret multidisciplinary in their own way. One of the group members explained this as follows:

> yes you could say these words were deliberately vague. That was because we had different treatments then, a centre in the west would do it different from one in the south. We had to cover them all with this recommendation.[68]

Another guideline group member adds: 'because of the diversity of forms [of treatment], and the lack of hard evidence you have to let it go, so people can find their solution locally.'[69] The need to choose this wording was thus not a factor of a decision based on evidence, as the guideline discussed and detailed that no judgement can be made about the evidence for these treatments, it was rather a clinical necessity to give GPs this choice, as

---

68 Dutch back pain guideline group member 1 (Utrecht, 28th of November 2014)

69 Dutch back pain guideline group member 2 (Amsterdam, 5th of December 2014)

otherwise there would be no treatments for chronic lower back pain (and implicitly that was seen as a bad thing).

*Evidence Use for Acupuncture*

Apart from the treatments the guideline recommended we also have to talk about the treatments the guideline did not mention. In contrast to the English guideline, the Dutch guideline does not mention acupuncture as a potential treatment. There are broadly two reasons for this: evidence for acupuncture was not included in the search for evidence, and nobody with expertise in acupuncture was part of the guideline group. The search strategy used to find the evidence for this guideline has been published[70] and acupuncture has not been included as a potential treatment and has not been searched for. The guideline group members denied this was a deliberate choice, but some did point out that acupuncture is an 'alternative treatment' and so not something 'regular medicine' should promote. One of the members said:

> [it is] alternative medicine, so we won't do it. It is never expressed in those stark terms, I think, but it is an implicit starting point. (…) it is funny that you should ask [this] because it reminds me that you start a guideline

---

70 NHG, 'Zoekstrategieen Lage Rugpijn' <https://www.nhg.org/sites/default/files/content/nhg_org/uploads/zoekstrategieen/lage-rugpijn_m54.pdf> accessed 17 November 2014.

> making process with unspoken assumptions that are often not very clear (…) we only talk about regular medicine.[71]

This shows that one could say that there was an implicit choice not to include any reference to it, based on a mutual implicit understanding that only regular medicine would be discussed and that acupuncture is not regular medicine. It also highlights how through a process of selection one can (deliberately or not) exclude certain aspects and therefore conclude there is no 'evidence' to support certain treatments.

## 4.9 Internal Consensus Building: How Results were Reached

According to my interviewees, within the guideline group, at least in the early phase of constructing the guideline, there were heated discussions. A guideline group member expresses this in the following terms:

> It was quite chaotic [the development of the guideline] There was a lot of argument, people left the group new people joined, I never experienced that before (…) The first meeting was called to study the existing standard and formulate criticism (…) In this first meeting people set out their red lines, this person had a PhD on this treatment, another had a PhD on that

[71] Dutch back pain guideline group member 2 (Amsterdam, 5th of December 2014)

> treatment, (…) that immediately caused a lot of friction so at the third meeting people left because they could just not work together.

The discussions centred mainly on what the best treatment was for lower back pain. With various participants defending various different treatments related to their varied background in treating lower back pain. As these differences were fundamental, and the group was unable to bridge the differences of opinion, two members decided to leave the group at this stage.[72] This decision did not, however, lead to the abandoning of the guideline. Instead two new members were recruited, and in the end the group was successful in achieving consensus. A guideline group member commented on achieving consensus in the new guideline group in the following terms:

> We ended up with a lot of consensus points in the guideline that we had to discuss. And we discussed about it all and formulated a position [as a group].[73]

Asked how such 'consensus' was achieved I was told that 'At the end of a long discussion one guideline group member would summarize the discussion and we will come back to it next time.'[74] So the method employed was to have long discussions, where debates would

---

72 Regretfully I have not been able to identify those participants, and thus have not been able to hear their side of the story. It is important to note that in describing these events I do not wish the give any value judgement.

73 Dutch back pain guideline group member 1 (Utrecht, 28th of November 2014)

74 Dutch back pain guideline group member 1 (Utrecht, 28th of November 2014)

be summarised and fed back to participants in a subsequent session until it was decided that everyone could agree on a formulation of a recommendation.

One of the reasons the new group of guideline group members were able to work out a consensus on the treatments for lower back pain was because they abandoned the option of declaring one treatment to be superior to another. And one of the guideline group members links this to debates about clashing ideas about the model of back pain. Specifically applying this to the weighing of medical studies: 'When there is no hard evidence, you see that you see a different weighing from different believers. So you see then if you believe it, you will weigh this [study] highly.'[75] Another guideline group member also related this to discussions about different ideas about the condition:

> Yes that is [about] our vision on what you should do with this condition, sometimes even stronger then evidence that points in a different direction. Eventually it has to fit with the model that you have in your head about the disease.[76]

It was through establishing a shared vision of the disease that the group could accommodate different perspectives. As noted earlier, the fact that the treatment options were described in relatively open terms made it possible for everyone to agree. In this way a statement on

---

75 Dutch back pain guideline group member 2 (Amsterdam, 5th of December 2014)

76 Dutch back pain guideline group member 3 (Rotterdam, 5th of December 2014)

whether one treatment was better than another treatment was avoided. A specific example of how this was achieved was through the description of the possibility of referral to an interdisciplinary centre. This decision was purposefully drafted in an ambiguous way, no explanation was provided in the guideline as to what would constitute an interdisciplinary centre. This ambiguity was a useful tool for the guideline group because of the differences in treatment of lower back pain in various parts of the country, where a choice for a specific centre or specific method would risk exclusion of many current providers of treatment for people with chronic back pain. One guideline group member was quite frank in explaining with reference to international studies on back pain that: 'We don't have a clear definition of what a multidisciplinary centre is.'[77] In the case of this multidisciplinary intervention this ambiguity was thus a tool to help achieve consensus and make sure the guideline would accommodate for reginal diversity in treatments within the Netherlands at the time.

In sum, the guideline group thus used different overlapping methods to achieve an acceptable result. Evidence was used at a starting point in discussion but this was complemented with knowledge of clinical practice, and understanding of the condition and a desire to compromise on the specificity of recommendations to be able for everyone to

[77] Dutch back pain guideline group member 2 (Amsterdam, 5th of December 2014)

agree. The fact that to reach this consensus that guideline in some areas became less specific was seen as a necessary trade-off.

## 4.10 External Responses to the Guideline: A Deafening Silence

The Dutch response to the guideline was limited. One critical response came in the form of a letter to *Huisarts en Wetenschap*[78] in which the authors conclude the guideline did not conform with daily practice and gives a false sense of security. The letter argued for a further classification of non-specific lower back pain patients before providing them treatment would provide higher quality treatment and lower costs[79], referring to a RCT to support their position.[80] In a response in the same journal the guideline drafters dismiss the complaints as unfounded, based on a misreading of the literature.[81] We can conclude that

---

78 One of the most important academic journals on care in general practice in the Netherlands published by the NHG.

79 'NHG-Standaard Aspecifieke Lagerugpijn, Een Herziening? | Archief | Website van H&W | Huisarts En Wetenschap, Het Maandblad van Het Nederlands Huisartsen Genootschap.' <http://www.henw.org/archief/volledig/id1815-nhg-standaard-aspecifieke-lagerugpijn-een-herziening.html> accessed 20 October 2014.

80 Julie M Fritz, Anthony Delitto and Richard E Erhard, 'Comparison of Classification-Based Physical Therapy with Therapy Based on Clinical Practice Guidelines for Patients with Acute Low Back Pain: A Randomized Clinical Trial' (2003) 28 Spine 1363.

81 'NHG-Standaard Aspecifieke Lagerugpijn, Een Herziening? | Archief | Website van H&W | Huisarts En Wetenschap, Het Maandblad van Het Nederlands Huisartsen Genootschap.' (n 79).

the Dutch guideline was not as controversial as the English guideline turned out to be. There are several reasons why this might be the case.

Firstly, this Dutch guideline was able to neatly sidestep the more controversial aspects of the treatment of lower back pain, such as injections in the back and spinal fusion as these elements of the treatments are not covered by GPs. Secondly, this guideline was potentially not as ambitious as the English guideline: the treatments it prescribed are relatively openly formulated and provide little in the form of guidance other than stating that 'passive' treatments should be avoided. Thirdly, this guideline was made by the NHG, which as the professional association of GPs ensures in several ways that this guideline is supported by its members. For example, 50 randomly selected GPs are sent the guideline and are invited to make comments, and an authorisation committee consisting of professors of general practice have to authorise the guideline before it becomes final. In this way the chance of a guideline being controversial are limited.

Moreover, this article appeared after the guideline was published, and did not result in any changes to the text of the guideline, the impact of these external commentaries on the guideline was therefore minimal.

## 4.11 Conflicts of Interest: A Limited Role

A direct or indirect influence by the pharmaceutical industry has not been established in the research into this guideline. Pharmaceutical influence would in any case have been unlikely given the fact that the only medication that forms part of the guideline is no longer under patent, so it is no longer profitable for the companies who hold those patents to pursue aggressive marketing. As these drugs are now generic, anyone can produce them and the costs of these medications are low, and so are the profit margins. Pharmaceutical companies direct their marketing towards drugs that are patented as during this period they can make a profit, aimed at recouping the costs of developing the medication. In addition, the potential for guideline group members to influence the pharmaceutical choices in the guideline was limited given the decision to follow a different generic NHG guideline on pain relief and apply this to lower back pain. For other medication that fell outside of the remit of the NHG pain relief guideline the group decided not to recommend any (this concerned muscle relaxants and antidepressants). It is unlikely this aspect of the guideline was in anyway unduly influenced.

This is not the only potential influence on the guideline as the result of a potential conflict of interest, it is worth looking at the financial and non-financial interests of the group members. One of the guideline group members has a financial interest in a clinic that provides a certain type of injections in the back. The position of this member was scrutinised by the head of the division for guideline development of the NHG, and he was

accepted as a member of the group. The potential influence of his position on decisions in the group was clear from the outset. The interviewed group members all stated that they felt there was no undue influence on their decision-making. Given the fact that the group discourages referral to monodisciplinary specialists who provide injections in the back[82], and in general observe that there is limited evidence to support this treatment, any influence of this specific guideline group member on the guideline to service his personal interests are absent.

The only potential influence or conflict of interest that affects this guideline to a degree is the fact that some guideline members have also (co)-authored publications they cite as evidence. This potential for a 'cognitive bias' is thus a possibility, where through the preferences of the members of the group for using and citing work they have authored themselves the 'evidence' they base their decisions on maybe distorted. For several of the Cochrane Reviews used in the guideline some of the guideline group members are co-authors. Given the extensive vetting process that is adopted for making Cochrane reviews, it is unlikely that the fact that guideline members have worked on the evidence they cite will unduly 'corrupt' the evidence used. What is likely, however, is that the guideline group chooses to cite documents they are familiar with. In this way a conflict of interest, having co-authored a medical study, might lead to a bias in the process of selecting evidence. Even

82 'NHG-Standaard Aspecifieke Lagerugpijn | NHG' (n 48) 768.

though such a bias cannot be excluded a large part of the Cochrane reviews they rely on are all co-authored by one high profile person, who is one of the most prominent epidemiologists in the field of back pain, and a figure that is well known to the members of the guideline group. So even if these guideline group members would not have happened to be involved in this study, these studies would nonetheless have been selected, because of the prominence they have in the field of back pain. So in this way these 'cognitive biases' have not materially affected the text of the guideline. We can conclude that potential conflicts of interests have not played a role in influencing the Dutch guideline on lower back pain.

## 4.12 Institutional Factors: Pressures to Conform

When analysing the Dutch guideline on lower back pain in its broader regulatory context we should take into account two interrelated viewpoints: how is the guideline positioned from a top-down perspective from government to profession, and how is the guideline positioned in a bottom-up perspective from individual healthcare practitioner vis-à-vis the guideline. We can also examine the meso-level question of how guidelines are positioned within the institutions that produce them.

*Top-down Government Pressures*

If we consider the top-down perspective, it is important to clarify the relations between the various organisations involved in aspects of guideline making in the Netherlands. The government is involved in controlling certain aspects of guidelines. The most important organisation in this regard is the Zorginstituut Nederland that, as explained in chapter 2, provides judgements on what is insured under the private insurance policies Dutch citizens are required to have. So, the fact that the guideline on lower back pain recommends several treatments does not mean that this is part of the package for which people are mandatorily insured.

This is can be illustrated by looking at physiotherapy: the guideline recommends this as a possible treatment for lower back pain but it is not covered by insurance, so patients have to pay for this treatment themselves. This also explains why there no explicit cost-benefit analyses to be found in Dutch guidelines, as guidelines do not in any way create rights for patients to receive reimbursement for a certain treatment. In this sense there is a greater distance between government bodies and guideline making organisations. It is the government (in the form of the Zorginstituut) and elected politicians who have the

final say on what is covered by health care insurance and what is not, but they use guidelines to help them decide instead of influencing those guidelines themselves.

For the lower back pain guidelines this means there was no government involvement in making the guideline, nor does the guideline refer to government rules in any way. It seems that at least for the Dutch guideline the profession was independent in formulating the optimal treatment strategy. These efforts were, however, embedded within the broad legislative guidance provided in the Law on the Medical Treatment Contract[83], that care of a good standard conforming to the norms of the profession should be provided. And it is this legislative backdrop that indirectly steers the profession towards developing its own professional standard, and guidelines are one instrument in doing so.

There are other ways, however, for the government to exercise control and secure a high quality of care. Medical tribunals can discipline doctors and they often use medical guidelines to establish what care of a good standard is. These medical tribunals consist of a mix of judges and medical professionals so that control of the quality of the profession is not completely in the hands of the profession itself. Nonetheless, to a large extent having medical guidelines to back up what the profession considers a good standard of care results

[83] Wet op de geneeskundige behandelovereenkomst (Stb 1994, 837).

in the profession effectively controlling what good care is. In essence the profession creates to a large extent the norms they then have to comply with. Medical tribunals follow guidelines to enable them to formulate what the expected course of action of a medical professional is so they can assess whether deviations from that standard are tolerated.[84] The claim that guidelines diminish the professional autonomy of medical professionals is not something that this thesis would argue for, but it is correct to say that by making their procedures more explicit the profession gives part of their knowledge monopoly away and enables forms of control and accountability. Van Everdingen predicted this almost thirty years ago in his thesis on consensus-based guidelines when he said:

> 'So far consensus-guidelines have not been used in legal procedures. That they will be used in the future seems quite likely. Through a correct formulation of the guideline text this will more likely strengthen the legal position of the doctor rather than weaken it.'[85]

Guidelines have grown to be instruments of accountability that can be used against doctors, but as Everdingen predicted correctly they can also strengthen the position of a doctor by giving him an argument to support his actions when they are controversial. So, what exactly is the position of guidelines as a form of regulation and how should we characterise this form of regulation precisely? I would suggest that the Dutch guidelines are a specific form

---

84 JJE van Everdingen, *Consensusontwikkeling in de geneeskunde* (Bohn, Scheltema en Holkema 1988).

85 ibid 56–57.

of mandated self-regulation, somewhere on the spectrum between 'pure' self-regulation and fully delegated rule-making: this is self-regulation in name but not in substance.

Guidelines in the Netherlands are a result of a symbiotic relationship between professional bodies like the NHG and the government. The government needs guidelines, as they break the information asymmetry between those inside and outside the profession. The government gets an insight into what healthcare professionals define as good care. As a result, they are able to regulate the profession through various mechanisms. But guidelines are positive for the professional organisations as well. It gives them a way of trying to control their members, in case of the NHG GPs, and steer them in the direction of better care. Whether they are always successful in that aim is not part of this study: research suggests results are mixed, but in general the acceptance of NHG guidelines is high.[86] Some guidelines of course are more directive than others. The lower back pain guideline provides relatively few directives apart from the statements that MRI's should not be used for lower back pain unless a specific cause of back pain is suspected. As we have seen the external response to this guideline was muted. There was little in the way of professional steering from the NHG in this guideline. It seems the main reason to revise the previous guideline was the general policy that guidelines should be updated after 5 years. In the Netherlands GPs try to influence their professional association and look for guidelines that are helpful

---

86 Richard Grol and others, 'Implementatie van NHG-Standaarden: Succes of Probleem?' (2010) 53 Huisarts en wetenschap 42.

in their daily work, solving practical dilemmas they face in providing treatments for their patients.

*The Institutional Context of the Dutch Guideline*

An important element in the regulatory context that easily is overlooked is the position of guidelines within the guideline making body. A part of this institutional context is the larger culture and values of the NHG. More specifically of particular relevance to the development of guidelines are the *perceptions* that guideline group members hold (whether true or not) about the culture of the NHG. The decision for example that acupuncture would not be considered was not an explicit decision made by the group, partly as a result of the values of the NHG as one guideline group member explained:

> The NHG doesn't want to promote this [acupuncture] or homeopathy, as long as it is not evidence-based effective it will not be put in the guideline.[87]

This guideline group member described a generally held view that reflects perhaps a broader norm in Dutch healthcare that acupuncture does not belong in the category of 'regular medicine.'

---

87 Dutch back pain guideline group member 1 (Utrecht, 28th of November 2014)

There are further institutional mechanisms within the NHG that limit and steer to an extent the room a guideline group has in drafting a guideline. The NHG has installed an advisory committee of GPs that are (or strive to be) representative of GPs in the Netherlands that suggests topics for medical guidelines to be developed by the NHG.[88] Furthermore all guidelines are sent to a sample of 50 GPs and their comments are incorporated in the final guideline. This might have some moderating effect because guideline groups will only be able to produce guidelines on topics that this advisory committee recommends being addressed or are otherwise put on the agenda by the NHG. Furthermore, this round of commentary by 50 GPs will have a moderating effect on the guideline, where larger deviations from current practice might lead to more comments being made. More important in practice is the so called 'authorisation committee.'

This committee that consists of professors of GP-care in the Netherlands has the final say over whether a guideline will be authorised, and thus be published as an official NHG guideline or not. This process consists of one or several meetings where the whole guideline group has to appear before the committee to defend the guideline and fully justify all the choices made. More ambitious guidelines that drastically want to change existing practice will have a harder time convincing this committee the guideline should be adopted.

---

88 The NHG advisesraad standaarden https://www.nhg.org/nhg-adviesraad-standaarden-nas [The NHG advisory board for medical guidelines]

One of the guideline group members summarises the position of several members of the group in explaining the process:

> A guideline has to be supported by the profession. The authorisation committee looks [at the guideline to see] if everything was thought through properly. There are always discussion points where we go through the arguments again for taking a position. And then they agree. But if they say you have to think about these points we would have to go back and look at it.[89]

This is clearly more than just a simple rubber-stamping exercise, the guideline group members all have to be present to be able to talk in detail about all the aspects of the guideline. In this way the NHG retains institutional control over the guidelines it publishes.

One of the ways this manifests itself in general is the more risk-averse strategy adopted by the NHG. The *NHG-standaarden,* I would argue are reluctant to adopt new treatments and require a relatively higher threshold before adopting new treatments. This is because of the moderating influence of the authorisation committee that moderates large changes in guidelines, and so acts as a brake on guideline groups. In general GPs in the Netherlands are seen as conservative, they provide lifelong care and they are reluctant to adopt new treatments because they might find out the treatments are unsafe a couple of years later, and they will still be responsible for the affected patient. A Dutch guideline group member said: 'Yes I guess you could say that we [GPs] are conservative. In most

---

[89] Dutch back pain guideline group member 3 (Rotterdam, 5th of December 2014)

cases it is something like, oh there is something new what should I do? Let's wait and see what happens before we try it ourselves.'[90]

## 4.13 Comparison of Guidelines: Difference or Similarity?

This chapter has so far reviewed both the English and Dutch guideline on lower back pain. As we shall see the guidelines present a complicated picture of similarity in recommendations despite differences in procedure, context, evidence and interest. At the same time, we can observe differences in recommendations where evidence is very similar.

*Evidence Interpreted*

If we take the findings from both our guidelines together it becomes clear that evidence is used in different ways. In the Dutch guideline the authors are interested in illness behaviour, and so they stress the findings of the study with the 'Back Book' that the subjective perceptions of disability decrease especially for those who are risk-averse. The English guideline is more interested in the results on pain and disability, the objective results as it were, and so they report those (less favourable) results of the study.[91] This is a

---

90 Dutch back pain guideline group member 1 (Utrecht, 28th of November 2014)

91 In the case of the English guideline this seems to be a result of the following clinical question this was used to search the literature: what is the effectiveness of prepared patient information material compared to no information or alternative information on pain, functional disability or psychological distress? In reporting the results however, the element physiological distress is not reported and to focus is on pain and disability.

good example to show that it takes a lot of effort to disentangle the evidence in 'pure form' and the clinical consideration of that evidence. Only by looking at the selections and omissions made in using and not using particular results a study do we see the selection process that is an integral part of the adjustment process to make a guideline fit in a clinical context.

If we return to the various treatment decisions taken in the English guideline, the repertoire of science plays an important role. In the description of the decisions taken it is clear the guideline group carefully describes the various strengths and weaknesses of the studies used. This is especially the case when acupuncture is discussed, the controversy surrounding this treatment is greater than for treatments that have been recommended in previous guidelines, which might explain the greater stress that is put on the robustness of the scientific evidence base in this case.[92] From the interviews with the guideline group we could learn that great stress was also placed on the robustness of the evidence considered, and the attempts that are made to account for evidence that might have methodological flaws or shows only small benefits. The technical team and the definition of the research questions they use in their search through the literature have a marked influence on the repertoire of science as Moreira defines it.

---

92 This decision has indeed attracted criticism which was discussed under external politics.

As we have seen the fact that research questions are defined narrowly, may lead to search terms that only turn up a part of the literature, and this literature then undergoes a further filtering: examples include the exclusion of reviews due to the heterogeneity of the studies involved, a sample size that is too small etcetera. It is only after this filtering process has been completed that the *remainder* of studies are discussed and the process of considering the robustness of these studies takes place.

If we compare the English case with the way the Dutch guideline group treats evidence we see that there are some important differences in the way evidence is handled. The Dutch guideline group adopts a less strict definition of evidence and doesn't restrict itself to considering RCT's, it also doesn't employ the GRADE-methodology, nor can it rely on technical assistance from a team of methodologists. The potential evidence the Dutch guideline group could have considered is therefore broader, and one might therefore expect that more low-quality evidence would be used. This, however, is not the case as the guideline group relies heavily on Cochrane reviews, which are systematic reviews seen as the highest possible quality of evidence.[93] This is logical if one considers that the Dutch guideline group had to read and appraise *all* evidence themselves, there was not enough

[93] They are at the top of the evidence-based medicine pyramid discussed in chapter 2.

time to consider separate randomised control trials, other than for those areas where no systematic review was available.

The differences in the treatment of evidence between the two guideline groups lies not in the decision of what evidence to select, but on how to appraise it. The Dutch guideline group is more strictly focussed on the clinical relevance of certain evidence for GPs, and pays less attention to the strict requirements of evidence-based medicine. An example of this is provided in the treatment of the same evidence for educational materials: whereas the Dutch guideline group stressed the clinically relevant results, the English group stressed the statistically relevant results. It is easier for the Dutch guideline group to stress clinical relevance because they have to write a guideline from the perspective of only one profession, whereas the English guideline group has to incorporate the perspective of several different professionals. Early on in the guideline drafting process members of the Dutch group left, the subsequent decision to see back-pain treatments as interchangeable helped achieve consensus. This means that the Dutch eventually found it easier to reach compromises, and easier to tailor their message to a specific audience, the Dutch GP.

If evidence is not the reason the guidelines are so similar what is? The fact that the guidelines are so similar is due to the fact that they are drafted in quite general terms to make sure that they confirm to a great extent to the current practice in the treatment of

lower back pain in and the evidence is fitted to this reality.[94] The route towards this result differs: the use of consensus when high quality evidence is lacking in England, compared with a pragmatic interpretation of evidence based on clinical preferences in the Netherlands - but this did not lead to very different outcomes. The areas where we see a clear difference support this: the inclusion of acupuncture in the guideline and a greater role for forms of manual therapy in the English guideline, are also the areas where there is greater difference in the practice of treatment of lower back pain between the two countries, and where the NICE guidelines are perceived to be the most controversial.

*Internal Consensus Building: A Careful Process of Creating Acceptability*

The discussions in guidelines groups can be passionate and heated but in both groups consensus was reached in the end. Guideline members have different opinions about the clinical relevance and context of guideline decisions. They can differ passionately about the relevance of results from clinical trials. Is a small treatment effect enough if the finding applies to a large population of patients? Are certain therapies more effective than others, and if so should we discourage doctors from referring to those therapies that are less

94 The areas where the two guidelines differ (acupuncture) is also where a deviation from current practice is recommended.

effective? The answers to those questions are complex and the evidence derived from scientific articles cannot provide simple answers to those questions.

In reality guideline groups adopt a pragmatic approach to these questions and strike compromises. Differences in guideline text appear because those compromises work out in different ways. The Dutch guideline group has chosen to give relative general recommendations, or in the case of multidisciplinary treatment use deliberately vague terms to reach their consensus. In the English guideline, the group introduced the choice of the patient as a factor in choosing a treatment and so neutralised any differences of opinion over which of the several treatment options where more effective. In this way both guideline groups managed to achieve consensus and incorporate evidence in a balanced way.

Words have no fixed meaning in and off themselves. They are in these careful processes of consensus-building imbued with specific meanings that help to bridge differences in understanding between the various parties around the table.

*External Responses: From Silence to Upheaval*

The external responses to the two guidelines are very different. Whereas the English guideline led to a storm of criticism from various professional bodies, the Dutch guideline did not cause such a stir. This seems to be related to the perceived controversial nature of the recommendations adopted in the respective guidelines. On the one hand there were

those that complained about the exclusion of certain injections in the back as a treatment for lower back pain, while others complained about the inclusion of acupuncture as a treatment option in England. The guideline group is in a difficult position as the different 'worlds of a guideline'[95], as Moreira defines them, inevitably will clash. If the guideline group is determined to present an ambitious guideline that actively changes service delivery, it is inevitable that this will cause a measure of resistance. But if the guideline group chooses to present a conservative guideline that just duplicates already existing practice, what is the rationale for making the guideline in the first place? There might be no need for a guideline as practice already conforms perfectly to evidence-based recommendations? The guideline group will always have to navigate between those two extremes different diseases have different profiles in this regard. Lower back pain is an area where our knowledge of the disease in the last 20 years has not dramatically changed, for diabetes we shall see that this is completely different. The extent to which a guideline is willing to adopt positions that might be seen as political is not only related to their level of ambition but also premised on consensus within the group. This is an important reason

---

95 Tiago Moreira, 'Diversity in Clinical Guidelines: The Role of Repertoires of Evaluation' (2005) 60 Social Science & Medicine 1975, 1976.

the Dutch guideline was less controversial: there was no consensus to adopt very specific treatment recommendations.

*Conflicts of Interest: None Established*

As we have seen the fear that guideline groups have conflicts of interests that negatively influence their recommendations has not been borne out in this case. This study has found no evidence that the pharmaceutical industry had any influence on either of these guidelines. Personal or financial conflicts of interest where adequately dealt with.

The only remaining potential conflict of interest lies in the fact the guideline group members are sometimes involved themselves in the trials they use as evidence. Guideline members are invited to participate in guideline drafting precisely because they are experts in the field at hand or otherwise have methodological expertise that is of use. It is therefore unavoidable that guideline members will be confronted with studies that they conducted themselves. As long as there are sufficient safeguards in the form of a balanced representation of members in the group this is unlikely to influence guidelines in a negative way. This shows that not all guidelines are equally prone to influence and corrupting conflicts of interest, in the case of the lower back pain guidelines no evidence was found to support this.

*Institutional and Regulatory Factors: Competing Pressures*

We have so far analysed extensively how guidelines are drafted by looking at the use of evidence, analysing the internal politics of guideline groups and gauging the direct reactions to a guideline. It remains necessary to place guidelines in a broader regulatory context.

Considering the top-down regulatory pressures, one would expect to see a rift between the two countries as on paper their healthcare systems are based on two different models. The Dutch system is based on insurance with private hospitals and elements of market competition and so one would expect limited government interference with guidelines, compared to the NHS with a system of nationalised hospitals and care funded through where one would expect more government interference. In practice, however, the differences are not as stark for lower back pain guidelines.

In the Netherlands, as expected, direct government control of lower back pain guideline production is almost completely absent but the government does have influence on how guidelines are used *after* they are produced. Through rules on remuneration, medical tribunals and rules about contracts between healthcare insurers and healthcare professionals the government uses guidelines outside of the context in which they were originally drafted and for other purposes then originally intended. The Dutch government doesn't directly influence guideline drafting, but through these other mechanisms does

informally require the profession to satisfy a standard of good care, which guidelines help deliver.

The English guidelines are drafted under the auspices of NICE and the remits for those guidelines are set by the Department of Health. As we have seen this is just a formality in reality and does not mean that the government is actively involved in setting topics for guidelines. Arguably using NICE is a good way for the government to distance itself for the guideline making process. One element of guideline making the government does provide guidelines for is the cost-effectiveness of treatments. We have seen that it is NICE that operationalizes this through the use of Quality-Adjusted Life Years so in this respect the government has control over guidelines. The necessity to include a form of cost-benefit analysis in guidelines also has the potential to change the choices the guideline group makes. This is a way to avoid politically controversial decisions on (pathways of) medical treatments for the government.

For bottom-up pressures the expectations would be different. There is no reason to expect differences in pressures from healthcare professionals on guideline makers in the two countries as their interest and composition are very much the same. Here, however, we do observe a difference because of the different institutions making the guidelines.

For the lower back pain guideline in particular the issue come to the fore in the criticisms expressed by the British Pain Society. This criticism seems to be based on a lack

of a feeling of ownership (and thus support) for this particular guideline, despite the President of their Society being involved in the guideline they felt he did not represent their interests in developing the guideline. Whereas other healthcare professionals involved in the delivery of care for patients with lower back pain were part of the guideline group they felt excluded. Such issues of ownership were less of a problem in the case of the NHG, where only the profession itself, the GPs, needed to feel ownership of the guideline recommendations.

If a guideline changes in a major way this is a good opportunity to explore what has driven this change and on what evidence this change is based on. On this basis we can conclude that the Dutch guideline group has a low risk appetite with only minor changes to the lower back pain guideline, and a view that the evidence for lower back pain is not of a high quality.

Given these constraints it seems surprising that both countries have drafted lower back pain guidelines that are so similar. A closer look reveals that while the results only differ in certain respects, the way in which the result was reached was different. In summary, the way evidence was used in the Dutch and English guidelines was based on the perceived importance of clinical knowledge, the relevant regulatory environment, and the relative difficulty in obtaining consensus in the guideline group.

# 5 DIABETES TYPE II: GUIDELINES IN ENGLAND AND THE NETHERLANDS

Diabetes type II is a serious chronic condition that can have debilitating complications and affect patient's quality of life severely. Simply put it is characterised by a heightened level of blood glucose, because of an inability of the body to process sugar or glucose[1] sufficiently. This has several consequences as it increases the risk that (severe) complications will ensue such as stroke, heart attacks, blindness, kidney failure, neuropathic pain[2] and amputations. Diabetes treatment is not (yet) aimed at curing diabetes, but the aim of treatment is to normalise the blood glucose and lipid levels in the blood as much as possible so that the risks of developing these complications is reduced, successful treatment would find the optimal point between reducing the risks of the disease and the harm from the side effects of the treatment.[3]

Diabetes type II is distinct from diabetes type I: type I is a condition with a predominantly auto-immune component where the patients are not able to produce any insulin and are dependent on insulin permanently. Diabetes type I will not be covered in this thesis. For patients with diabetes type II the treatment of diabetes type II is evolving fast, and new medication is coming on the market every year. The challenge in drafting

---

1 Throughout this chapter blood sugar and blood glucose will be used interchangeably.

2 Pain as a result of nerve damage, a complication of the disease.

3 An important development here is the recognition that this optimum differs between patients.

medical guidelines is to provide up-to-date advice while being mindful of the cost of new treatments versus the health benefits gained.

The treatment of diabetes recommended in guidelines is variable but often involves one or more of the following steps: providing patients with information/advice to lose weight; a diet; a blood glucose lowering drug, moving on to two or even three blood glucose lowering drugs and finally insulin by injection possibly in combination with blood glucose lowering medication. This treatment pathway is thus one that increases in severity to help control the blood sugar levels of a diabetes patient. The treatment of diabetes is however related to how long a person has had diabetes, the age of the patient and the other healthcare conditions a patient might suffer from. In designing diabetes guidelines the guideline group members thus have to decide when to intensify treatment and which combination of drugs should be provided at the different stages of treatment.

In researching the drafting of Diabetes type II guidelines a particular focus is placed on blood glucose lowering medication. This means other aspects of diabetes care will not be covered in detail. This choice has been made for two reasons. Firstly, there is not enough space to cover all aspects of diabetes care in great detail: as diabetes is a chronic condition with a great many side effects its medical treatment reaches over many medical disciplines. A further reason for picking out only a few aspects of diabetes care is the fact that the scope of the Dutch and English guidelines varies considerably. It is helpful to only discuss those aspects of diabetes care that the English and Dutch guidelines pay attention to and

formulate recommendations on. This means that most of the complications of diabetes care that are treated in hospitals, and other second-line services (such as diabetic foot care and eye care) will not be covered.

## 5.1 The English Guideline on Diabetes Type II

In this section we will discuss two diabetes guidelines developed by NICE: the 2008 guideline developed by the National Collaborating Centre for Chronic Conditions and authorised by NICE[4], and the update of this guideline by a so-called short guideline produced by NICE in 2009.[5] The focus is placed on the 2008 guideline, and reference is made to the short guideline when relevant for parts of the original guideline that have been updated.

## 5.2 Evidence-base for English Treatments for Diabetes Type II

The 2008 guideline provides guidelines on the treatments for the whole trajectory of the diagnosed diabetes type II patient. The elements that will be discussed under this heading

---

4 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London, *Type 2 Diabetes National Clinical Guideline for Management in Primary and Secondary Care (Update)* (Royal College of Physicians 2008) <http://www.ncbi.nlm.nih.gov/books/NBK53885/> accessed 14 April 2015.

5 NICE, 'Type 2 Diabetes: Newer Agents' <http://www.nice.org.uk/guidance/cg87/resources/cg87-type-2-diabetes-newer-agents-a-partial-update-of-cg66-short-guideline2> accessed 21 October 2014.

are: education and lifestyle advice, blood glucose lowering drugs, and target levels of blood sugar.[6] Other elements of diabetes care will not be covered as discussed above.

The 2008 guideline as the first element of treatment outlines lifestyle advice and education for diabetes patients. This is significant as it involves non-pharmacological interventions that have the potential to prevent people having to use medication. Paradoxically this is the area in the guideline where there is only a small amount of evidence to base decisions on. The guideline group based its decision on a health technology assessment report[7] that reviewed 14 studies, but concludes that they are of a poor quality and it is hard to draw firm conclusions from them. A similar picture emerges when it comes to lifestyle advice: only four studies are included and these all have a different design, which only allows for the drawing of tentative conclusions. As a result, the recommendations formulated in the guideline in this area are of a general nature such as: 'Provide dietary advice in a form sensitive to the individual's needs, culture and beliefs being sensitive to their willingness to change, and the effects on their quality of life.'[8] These

---

6 Technically we would need to speak about levels of HbA1c, as this is the measure used to express the average blood sugar (or better blood glucose levels) over a period of time. For the sake of clarity blood sugar is used as the preferred term.

7 E Loveman and others, 'The Clinical and Cost-Effectiveness of Patient Education Models for Diabetes: A Systematic Review and Economic Evaluation' (2003) 7 Health Technology Assessment <http://www.journalslibrary.nihr.ac.uk/hta/volume-7/issue-22> accessed 29 April 2016.

8 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 38.

general recommendations by the guideline group are understandable given the fact that it is unclear which type of education and/or lifestyle advice is most effective.

*Blood Glucose Lowering Medication*

Blood glucose lowering drugs form the second step in the guideline: if education and lifestyle advice does not control the blood sugar levels of the diabetes patients to a sufficient degree, blood glucose lowering medication is needed to reduce the chance of the micro- and macro-vascular complications of diabetes type 2. The central question facing the guideline makers was to advise which medications to prescribe and in which order, because if blood glucose is not sufficiently lowered with one blood glucose lowering drug, intensification of the treatment is undertaken and one or more blood glucose lowering drugs are added to the treatment regime. Diabetes management is complicated, there are a multitude of choices to be made in prescribing medication. This creates a dilemma, as on the one hand one could argue that every patient is unique and diabetes management should thus be the ultimate tailored treatment, but on the other hand there is the idea that diabetes care should be protocol driven to achieve consistent results in a large population and to assure everyone receives the best quality of care for the lowest societal cost. In analysing the use of evidence, and how the guideline group appraised and used evidence in the two guidelines, this tension emerges consistently. And the guideline group had to make decisions that fall within this spectrum, where sometimes a uniform approach might be

prominent and other times a tailored advice is given. Guideline group members might not describe their reasoning in these terms but this clearly played a role in their deliberations.

In systematically discussing the evidence used in the guideline and the choices made in the guideline references several classes of medication for diabetes type 2. In this study no mention will be made of the brand names of these medications, but will refer to them by their generic name, which is the name for the active substance in a drug.[9]

It follows too far for the purposes of this study to explain in detail all the differences between these drugs but in a general sense one can say that each class of drugs employs a different mechanism to lower blood sugar levels, and insulin as a treatment eventually completely replaces the insulin produced by the body to help lower blood sugars. Prescribing medication from a different class is therefore a choice that makes a difference, even though all these drugs are employed for the same purpose of blood glucose control. Within these broad classes of drugs there are individual drugs, complicating the choices that a clinician has even further, depending on whether one views the drugs within the

---

9 The classes are: Biguanides, Sulfonylurea, Meglitinide, Thiazolidinediones, Alpha-Glucosidase inhibitors, GLP-1 mimetics, DPP-4 inhibitors, HPV-insulin, Biphasic Analogues and Long-Acting Insulin Analogues. This leaves out the Sodium-glucose-co-transporter-2-inhibitors that are discussed in the Dutch guideline but not included in their treatment plan, as these earlier English guidelines do not include these newest drugs.

classes as interchangeable or not. In describing the various drugs and their relative worth the focus will be placed on the reasons for preferring one drug over another or the reasons for making exceptions for specific groups of diabetes patients. There are 18 diabetic drugs, of which some belong to the same class of drugs that will be discussed in this case study. There are many different ways in which they can be combined in providing diabetes treatment, which underscores the complexity of this area of medicine.[10] In describing the evidence base for the choices in the treatment of diabetes reference will be made to these drugs to explain the options available to the guideline group in making their decisions.

*Blood Glucose Lowering Drugs*

In choosing the first treatment offered to patients that require blood glucose lowering drugs the guideline group makes reference to studies that show that Metformin lowers morbidity and mortality and has a cardio protective effect.[11] The choice for Metformin is largely a result of the UKPDS study, the largest study with diabetes patients that is available.[12] In this study the use of Metformin/Sulfonylurea/insulin was compared with diet. It showed

---

10 And even this is a simplification as most of these drugs can also be taken in combination form (metformin and pioglitazone for example) and/or are presented in several different formats (different strengths as well as different forms of 'release' of the active substance over the day, so called modified release drugs). Which brings the total amount of drugs that can be prescribed up to around 30 at the time of this guideline, and this has grown continually since.

11 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 85.

12 Irene M Stratton and others, 'Association of Glycaemia with Macrovascular and Microvascular Complications of Type 2 Diabetes (UKPDS 35): Prospective Observational Study' (2000) 321 BMJ 405.

that Metformin was superior to the other drugs in reducing mortality and morbidity. There was little discussion in the guideline group on this issue, and the evidence for metformin. The primacy of Metformin was uncontroversial as the UKPDS results were seen as convincing evidence. Added to its proven effectiveness was the fact that the drug was also highly cost-effective as well as safe.

There was a limited role in the guideline for Sulfonylurea as a first line treatment for diabetes in three very specific clinical circumstances: for diabetes patients who are not obese, if metformin is not tolerated, or in case of severe hyperglycaemia.[13] These were all quite specific clinical circumstances that confirmed that in specific cases if Metformin really was not an option Sulfonylurea could be prescribed. This did not alter the overall picture the guideline group draw from the evidence that metformin was the preferred drug.

In deciding which drug to recommend as the second step, adding a further blood glucose drug to the one already prescribed the group had a wide array of choices and evidence to rely on. In an 'ideal' world the group would be able to compare studies using metformin as a first treatment (or Sulfonylurea) and compare this within a RCT with patient

---

[13] National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 87.

that use metformin and the 4 different classes of drugs the group considered as second step treatment: Sulfonylurea, Meglitinide, Thiazolidinediones, and Acarbose.

The reality is that this level of evidence was not available to the guideline group. Not all these comparisons between various drugs are covered in RCTs as most RCTs compare a particular drug with placebo or with another drug in the same class (for example comparing Glimepiride with Gliclazide both drugs in the Sulfonylurea class). One of the guideline group members explained limits in the evidence considered:

> Every study that was put in front of us [contained] people who did not do dreadful [did not have severe complications], so say 8.5%Hb1Ac [blood sugar level], (…) people with Hb1Ac >11 might have responded differently [to the drug]. (…) The studies were not set-up to support guideline development (…) but to satisfy the FDA [Food and Drug Administration] and EMA [European Medicines Agency].[14]

This point is an important one, and something that is a specific factor that affects the evidence-base of the diabetes guideline as a whole. As this guideline group member explained, more extensively pharmaceutical companies that produce the drugs used to treat diabetes don't have to show that their drugs are superior to the ones that are already on the market, they merely have to show that they are safe, and so their comparator group in a medical trial is very often a placebo. The information that can thus be gained from these studies is that the drug is safe, and indirectly its effectiveness can be compared by looking

[14] English Diabetes guideline group member 3 (Leeds, 20th of May 2015)

at other RCTs that compare different drugs to placebo. But, and this is the central point, this does not answer the question whether one drug is superior or inferior when prescribed with other drugs. It is precisely this question that guideline groups had to answer.

The group used the evidence that was available to compare the different classes on drugs on several characteristics: glucose levels, weight, lipid profile[15] and adverse events. In making the decision on the preferred second line treatment these measures of effectiveness were taken into account, as well as the cost of the drug, and its clinical profile. This multi-layered decision-making process did not point to uniform solutions for all diabetes patients, several examples from the guideline can illuminate how this complicated weighing exercise was carried out. The guideline text gives an indication of the approach the guideline group endorsed in this area:

> Because such people vary in attributes (such as body weight and insulin sensitivity) which can affect choice of medication, and because some medication side effects can have consequences for aspects of daily living (such as driving motor vehicles), blanket recommendations cannot be made for everyone with Type 2 diabetes.[16]

---

15 Not further explored in this thesis.

16 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 53.

In weighing the position of the various drugs specific elements of the evidence stood out, that tipped the analysis of the medical studies in a certain direction depending on the particular position of this drug in the treatment regime of diabetes.

For Sulfonylurea an important question confronting the guideline group was whether the side-effects of Sulfonylurea might restrict their use. The GDG group decided that the side-effects would be a particular problem for specific occupations such as bus drivers. Here one sees an example of an attempt by the guideline group to bridge a gap between providing one type of advice for everyone, and medicine tailored to the individual. The concerns about side-effects were not enough to say that Sulfonylurea should not be used at all: so, by distinguishing sub-groups of diabetes patients for which the drug is not suitable a balance was struck between the benefits and side-effects of the drug. Distinguishing these sub-groups was an important tool for the guideline group to establish a tailoring of treatment, making sure that patients who differ on relevant characteristics receive the diabetes care best suited to them.

The guideline group than compares the Meglinitide class of drugs with the Sulfonylurea class of drugs as the working mechanisms are similar. It considers that Repaglinide[17] is not superior to Sulfonylurea: due to the fact that it has to be taken three times a day and is more expensive than Sulfonylurea, this lead the guideline group to the

[17] One of the specific drugs in this class of medication.

conclusion that it should not be recommended for the general population of diabetes patients. One factor that presented challenges for the guideline group was that there was only one RCT directly comparing a Meglitinide with a Sulfonylurea.[18]It seems that the fact that the drug had to be taken three times a day was a major barrier.[19] From a clinical point of view it was difficult to endorse Repaglinide over Sulfonylurea as the uptake in clinical practice was very low, and there was limited experience among GPs and other diabetes experts in using the drug.

In this case it was not the evidence that drove a recommendation, because by all accounts the difference between Repaglinide and Sulfonylurea is marginal, it was the added complexity of administering the medication in clinical practice. It is interesting to note how the guideline group framed the evidence for Meglitinides: the focus was placed on the fact that they were not superior to Sulfonylurea. The guideline group concluded that there was no reason to change the recommendation in an earlier guideline. This shows that not all evidence is equal, if a drug has already been recommended before like Sulfonylurea, then superiority of the alternative drug presented in one way or another has to be shown before recommendations are changed.

---

18 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 66.

19 For fear that the therapy adherence by patients would be low. And frequent visits to the GP would be necessary to adjust the doses of the medication.

A third class of drugs of which evidence was compared were the Thiazolidinediones. There are two drugs in this class: Rosiglitazon and Pioglitazon. The discussion of these two drugs was heavily predicated with evaluation of the emerging adverse effects emerging from ongoing trials and the consequential worry about the safety of these drugs. This mainly involved evaluating the potential increased risk of myocardial infarction and death from cardiovascular causes. The discussion concentrated on a meta-analysis[20] that showed a 43% increased relative risk of myocardial infarction, which in absolute terms meant that 14 more people had a myocardial infarction compared to the control group.[21] The same study concluded that there was a 'borderline significant'[22] increased risk of death from cardiovascular causes. For Pioglitazone the main risk discussed was that of serious heart failure.[23]

This data was emerging at the time the guideline group was deliberating. The studies were not seen as robust enough as to warrant a complete rejection of the drugs but the eventual recommendation suggested only a narrow range for the drugs to be used: it was only to be used when other drugs were unacceptable, as in the case of insulin, or not

---

20 Steven E Nissen and Kathy Wolski, 'Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes' (2007) 356 New England Journal of Medicine 2457.

21 86/14371 in the pioglitazone group and 72/11634 in the control group.

22 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 93.

23 Lincoff A and others, 'Pioglitazone and Risk of Cardiovascular Events in Patients with Type 2 Diabetes Mellitus: A Meta-Analysis of Randomized Trials' (2007) 298 JAMA 1180.

tolerated in the case of Metformin and Sulfonylurea.[24] Here the group was again working towards establishing sub-groups of people that might benefit from these drugs. The therapeutic room here, however, seems particularly small as the drugs cannot be prescribed in people who have evidence of heart failure or are of increased risk for fracture.[25] These drugs provide an interesting case as they show clearly the difficult balancing act needed from guideline group members. The emerging safety evidence had to be weighed in deciding whether to advice against medications that were used frequently in clinical practice. In this case the aspect of side-effects was the preponderant consideration in evaluating the correct position of these drugs in the treatment pathway.

The final drug considered as a second-step treatment was Acarbose. Acarbose is not favoured by the guideline group. The main reason cited is that it has a high percentage of side-effects. The most important one being flatulence. It also notes its higher cost and the fact it lowers blood sugar slightly less compared to Metformin and Sulfonylurea. Here again the guideline group stated that it would not change its recommendation in an earlier

---

24 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 123.

25 ibid.

guideline. Even though Acarbose as a mono-therapy has the advantage of not causing hypoglycaemia, it is not preferred over any other drug.[26]

The way the various diabetic drugs are treated therefore differs depending on the availability of evidence, their clinical and safety profile, and their historical role in treating diabetes. The compromise that has to been struck between the various elements of cost, side effects, efficacy and clinical role is the same for all of them, which element is decisive differs. In this way a guideline is built up through several layers of decisions that in themselves are compromises between the various interests. In advising to use Sulfonylurea the guideline group did not make a blanket recommendation, as it carefully crafted alternative clinical scenarios where other drugs could be used, thereby retaining clinical autonomy and choice for patients.

If the second-step treatment 'fails' and the blood sugars are not yet sufficiently controlled it is possible for a third blood glucose lowering drug to be added to the existing two. Three classes of medication could be considered at this stage DPP-4, GLP-1 and the Thiazolidinediones. The 2009 guideline update mainly concerned the position of the DPP-4 and GLP-1 drugs in the treatment of diabetes.

---

[26] ibid 86.

DPP-4 inhibitors were a new (in 2008/2009) class of drugs. The 2008 guideline group did not consider these drugs in their guideline as it was decided the drugs would be covered in a 'rapid update' provided by NICE, which became the 2009 short guideline.[27] These new drugs were thus evaluated separately from the other drugs considered. The guideline group that drafted the 2008 guideline was unhappy with this arrangement as it prevented them from considering all the relevant drugs in one guideline.[28] The 2009 guideline is based on a health technology assessment report reproduced in the annex of the guideline.[29] The Cochrane systematic review of RCTs used as the main evidence to make decisions on the DPP-4 drug, only 4 RCTs turned out to be relevant. They fit with the licenced use of the drugs, are all sponsored by the manufacturers, and are short studies: one year or less in duration.[30]

The 2009 guideline group recommended the use of DPP-4 as second line therapy if either Metformin or Sulfonylurea was not suitable for this patient, or as third line therapy as an alternative for Pio- or Rosiglitazone. An interesting addition to these recommendations is the formulation of a 'stop rule' where the use of a DPP-4 inhibitor should only be continued after a beneficial metabolic response. This rule is interesting for

---

27 NICE (n 5).

28 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) Foreword.

29 NICE (n 5) Annex 2.

30 ibid Annex 2 87.

various reasons, it is clear that this rule is not based on evidence, but is a result of consensus within the group. A substantial decrease of blood sugar was seen as clinically meaningful, it would make sure that people did not stay on medication that was not beneficial, and would ensure blood sugar did not remain uncontrolled for too long.[31] In this way the guideline group fulfilled two goals. They gave sufficient heed to the fact that there was no longer-term safety data for these drugs, as they only recently entered the market. They also, and this is a continuation of the 2008 guideline group, strive to define smaller sub-groups of patients who benefit from, in this case, DPP-4.

Glucagon-Like Peptide-1 mimetics were the newest class of drugs available at the time and their appropriate place, in the pathway of care was part of the discussion in both the 2008 as well as the 2009 guideline groups. The only drug in this class licensed at the time was Exenatide. Exenatide was considered by the 2008 guideline group as a drug that might help prolong the period a patient with diabetes could postpone using insulin. Exenatide led to weight loss for patients and a lowering of blood sugar. The lowering of blood sugar was comparable to other drugs (either insulin of Glitazones), and the main debate was around the fact that the drug caused weight loss. It is important to note that glucose lowering drugs are not prescribed for their weight loss properties, although if the

[31] ibid 64.

drugs have this effect this could be seen as positive. This is because weight loss helps in maintaining blood sugar control.

The question facing the guideline group was how to weigh this evidence. Did the potential for weight loss justify the extra costs associated with this new treatment? I will deal with the detailed cost-benefit model used to answer this question when dealing with costs in the construction of these guidelines: however, both guideline groups suggested the benefits outweighed the costs. Exenatide had clinical benefits as a guideline group member discussed:

> This was a dilemma because we were back at cost, (…) we were now at 69 pounds a month, what we didn't want [is] people using GLP-1 it is not appropriate for, the more obese you are the more insulin resistance is the issue instead of beta-cell failure. The large amount of insulin (…) can cause weight gain. We didn't want it [GLP-1] to be a weight-loss treatment, but the GLP-1 could help people change their diet and so reduce their weight.[32]

For people who are obese, specifically for those with a BMI higher than 35, Exenatide could be an option[33] Here again, the guideline group defines a specific sub-group of diabetes patients who they expect to have the most benefit from this treatment. It is interesting to note that both guideline groups went to great lengths to interpret the limited

---

32 English Diabetes guideline group member 3 (Leeds, 20th of May 2015)

33 Or in the 2009 guideline also for those with a BMI <35 if 'therapy with insulin would have significant occupational implications or weight loss would benefit other significant obesity-related comorbidities.'

evidence[34] in such a way that Exenatide could gain a place in the pathway. The main basis for its desirability as a treatment was not blood sugar reduction but was the potential for weight loss, combined with the high costs of very large doses of insulin for people who are obese. Only in this scenario could Exenatide be prescribed. The reason that both the 2008 as well as the 2009 guideline group were keen to afford Exenatide a place can partly be explained by their willingness to accommodate those with anxieties about their weight and to prescribe medications that have a beneficial effect on weight: because this would benefit the whole of diabetes care in an integrated way (even if cost-effectiveness could not be established by looking at the effect of the drug in a more isolated fashion). In this case the limited evidence was used as a support for the clinical reasons to favour Exenatide, it provided an extra option for clinicians to treat people with diabetes. This clinical need provided sufficient reason to include the drug, while the 'stopping rule' invented by the group was aimed at assuaging fears that the cost of the drug would be too great. As one guideline group member summarised the issue succinctly: 'the justification for Exenatide is that it should lead to weight loss.'[35]

If we compare the new diabetic drugs to the 'old' ones there does not seem to be a fundamental difference in how they are treated in the guideline development process. The necessity to strike a balance between the various elements of costs, side effects, efficacy

---

34 The 2008 group relied on 8 RCT's and the 2009 group on 5.

35 English Diabetes guideline group member 2 (Cambridge, 6th of May 2015)

and clinical role remains. For the new diabetic drugs, however, the costs are higher; efficacy is not yet clearly established so the focus of the discussion is placed on the proposed clinical role, and the judgement of the group is focused on whether the drugs are likely to cause fewer side-effects. DPP-4 was included because it provided an alternative to the Thiazolidinediones class of drugs, who suffered from safety issues. And Exenatide was welcomed because it could prolong the period a patient could stay without insulin and allow further tailoring of treatment. The uncertainty surrounding long-term safety of the drugs is addressed by the guideline groups through consensus-based 'stop rules' which enabled them to strike a balance between the lack of evidence on long-term safety issues and the clinical wish to employ these drugs.

*Therapeutic Target for Blood-Sugar Lowering*

In addition to formulating the best therapeutic pathway for a patient: defining first, second and third line therapies the guideline also provides guidance on blood glucose levels. This advice has two functions, it advises on a target level for blood glucose for people with diabetes in general, and it provides thresholds above which intensification of treatment should commence. These targets are based on an analysis of the UKPDS study.[36] The achievement of these targets levels by patients should lead to a decrease in all

[36] Stratton and others (n 12).

microvascular and microvascular complications of diabetes and mortality correlated with a decrease in blood sugar. The guideline describes this as follows:

> The risk of each of the microvascular and macrovascular complications of Type 2 diabetes (…) was strongly associated with hyperglycaemia as measured by updated mean HbA1c [blood sugar level]. There was no indication of a threshold for any complication below which risk no longer decreased, nor a level above which risk no longer increased.[37]

Given this data the blood sugar target level should be set as stringently as possible. The group balanced the need for tight blood sugar control with the cost of very intensive treatment and decided to maintain the level of 6.5% HbA1c [blood sugar level] set in an earlier guideline:

> However, expensive therapies or very intensive interventions are required to achieve glucose control in the normal range in most people with diabetes. Consequently a population target should not be any tighter than the HbA1c [blood sugar level] of 6.5 % previously chosen for those at macrovascular risk. Nearly all people with Type 2 diabetes are of high CV [cardiovascular] risk, usually in association with insulin insensitivity, but if not with age. Additionally there has been very recent concern (no evidence yet to review) about pursuing very intensive glucose control (target <6.0 %) in people with higher CV [cardiovascular] risk and longer duration of diabetes, mostly on multiple insulin injection therapy.[38]

---

37 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 42.

38 ibid 45.

The very recent concern the group refers to is the ACCORD[39] trial, which was not yet published at the time this guideline was drafted but had to be stopped early as it later transpired because people in the intensive blood sugar lowering group had a 22% relative and 1% absolute increased mortality after a 3.5 year follow up period.[40] This established for the first time that a glucose target that is too low can also do harm through increasing death from cardiovascular causes. A member of the guideline group remarked on this choice of target level: 'It's all clinical judgement, the clinical view around the table. All the other guidelines seemed to have 6.5% at that moment, I can tell you the 6.5 [target] came out of thin air. There is the whole issue of goal-setting, which is necessarily an evidence-free area.'[41] What this guideline group member is explaining is that all the studies the group could rely on to set these target levels, discussed people who have achieved certain levels, which does not necessarily translate into goals for patients as a population as a whole. So it was the clinical consensus of the group that drove this recommendation.

The English diabetes guideline group went through each step in the treatment of diabetes in a systematic fashion. Every treatment/drug was considered separately drawing upon existing medical studies, clinical knowledge, and cost-effectiveness studies. The way for the group to find a balance between these various types of knowledge was to define

---

39 The Action to Control Cardiovascular Risk in Diabetes Study Group, 'Effects of Intensive Glucose Lowering in Type 2 Diabetes' (2008) 358 New England Journal of Medicine 2545.

40 ibid 2553.

41 English Diabetes guideline group member 1 (Warwick, 26th of May 2015)

sub-groups of patients for which a particular drug was suitable. In this way costs could be presented as being justified, clinical choice could be preserved and differences in opinion could be bridged.

## 5.3 Achieving Consensus: Internal Workings of the English Guideline Group

The English guideline group consists of a wider variety of medical professionals: GPs, practice nurses, endocrinologists and academic medical professionals with a specialisation in diabetes care. The impetus for producing this guideline was the regular process of updating the existing diabetes guidelines NICE had produced.

In analysing how the English guideline group achieved consensus we will take their discussion of evidence as a starting point because it illuminates the process followed to achieve consensus. Their discussions about evidence shows how the guideline group worked their way through the various medications and achieved a series of compromises to achieve a result which everyone could be happy with. This process differed depending on the type of medication that was under discussion. For the older and more established medications, the question was not whether they should have a role in the treatment of diabetes, but only to what extent their existing position should be modified. For the newer drugs the question was more fundamental: should their use be recommended? Only after it

was decided that newer drugs could be recommended, then their place in the therapeutic pathway could be established.

The discussion therefore varied with each drug, depending on which the aspect of the drug's profile as a treatment was the most controversial. An example would be the discussion about the inclusion of Exenatide, 'the new diabetic drug', in the guideline. This decision was a difficult one for the guideline group in the sense that it pitted several forms of knowledge against each other, and a resolution of this conflict required making a definite choice. Exenatide was considered to be too expensive for the NHS to afford, the original cost-effectiveness modelling performed for the drug showed that it had an incremental cost-effectiveness ratio far above £40,000 per year. A guideline group member discusses this cost-effectiveness model in the following terms:

> We were very challenging about the cost-effectiveness studies. (…) they did not include enough on the cost of hypoglycaemia. So the cost-effectiveness [study] was redone. (…) The general consensus was that we weren't quite sure we come up with a good enough number [of quality adjust life-years] but we have done the best we could.[42]

Subsequently the modelling was changed to more closely reflect the situation of a patient who the guideline group thought could profit from Exenatide. This clearly showed how the clinical wish to employ Exenatide was a motivating force to make sure patients would

---

42 English Diabetes guideline group member 3 (Leeds, 20th of May 2015)

benefit from this new drug and that considerations of cost were of secondary importance. This same guideline group member expresses clearly the direction of thinking in the guideline group: 'What we did (…) Exenatide came out as we cannot recommend it, the clinicians around the table said: you cannot say that, this is a new clinical area, a new drug, we must not allow the NICE guidelines to say a blanket no. So we wrote this slightly strange recommendation (…) those additions were written by the clinicians (…) as a sop to the health economics data, you cannot say no (…) because it is new.'[43]

The choice to recommend Exenatide was thus a combination of several factors that worked together to create a space for Exenatide in the guideline. The potential for weight loss was seen as an important factor and something older medication could not provide (as it would lead to weight gain). Especially in a group of obese diabetes patients, to postpone the decision to start insulin therapy was positive for patients. On the other hand, arguments against Exenatide were the high costs and unknown safety profile.

In this case the guideline group thought the arguments for inclusion outweighed the arguments against. They approached the issue of the cost of the medication by casting doubt upon the cost-effectiveness data, the relevant section of the guideline clearly showed this approach:

---

[43] English Diabetes guideline group member 1 (Warwick, 26th of May 2015)

> Exenatide therapy is expensive, and the health economic modelling suggested it was not cost-effective for an unselected population as compared to commencing human insulin therapy. However, the GDG did not consider comparison with an unselected population to be applicable (…)The GDG was uncertain that these agents (including Exenatide) would be found to be not cost-effective if the model fully reflected the negative quality of life issues of insulin, (…) Furthermore, the more obese require much higher insulin doses, such that insulin costs alone can easily exceed those of Exenatide (…) In these circumstances a confident judgment of costs and benefits to be gained from HbA1c and weight change, and side effects, could not be made. However the GDG's judgment was that costs of insulin and Exenatide by the end of the first year would be equivalent on average for people with a starting BMI (before these medications) of approximately >33 kg/m2, while in this obese group the small metabolic advantage to insulin on HbA1c would easily be outweighed by the metabolic advantage of 4 kg weight loss on Exenatide. In this restricted circumstance, and particularly at higher BMI's, the cost-effectiveness of Exenatide would then be at least as good as that of insulin.'[44]

It is clear the guideline group took aim at the cost-effectiveness data from various angles. The first step was to create a sub-set of diabetes patients, corresponding to the clinical views of the group of patients for which the group thinks this treatment would be beneficial. The next step was to identify several factors that are not captured on the economic model to underpin the conclusion that the fact that drug is not cost-effective cannot be supported. A key point is the fact that insulin is seen to be linked to negative quality of life issues, and that the weight loss related to the use of Exenatide is a way of addressing these issues in contrast to insulin which would lead to an increase in weight for most patients. When framed in this way the argument carried a smaller weight because the drug could be shown

---

[44] National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 122.

to be as cost-effective as insulin if one would adopt a broader view of costs (and benefits) than the model provides for. At the same time, as we observed earlier, the guideline group did not disregard cost entirely and devised a 'stop-rule' as a device to limit the use of Exenatide to a specific sub-group:

> 'Consider exenatide as an option only if all the following apply for the individual: a body mass index over 35.0 kg/m2 in those of European descent, with appropriate adjustment in tailoring this advice for other ethnic groups [,] specific problems of a psychological, biochemical or physical nature arising from high body weight inadequate blood glucose control (HbA1c ≥7.5 %) with conventional oral agents after a trial of metformin and sulfonylurea [,] other high-cost medication, such as a thiazolidinedione or insulin injection therapy, would otherwise be started. Continue exenatide therapy only if a beneficial metabolic response (at least 1.0 % HbA1c reduction in 6 months and a weight loss of at least 5% at 1 year) occurs and is maintained.'[45]

Interesting to note is that the BMI cut-off point in this 'stop-rule' is 35, while the narrative text cited before mentions 33. One of the members of the guideline group members confirmed that this was an editing mistake, and the number in the narrative should have been changed to 35. He said: 'You make these changes to quite a tight schedule; it should have been updated [in the evidence statement].'[46] While this might seem a trivial error, on closer inspection it shows that the guideline was extensively debated in the guideline group, and that several proposals for cut-off points were considered before that of BMI greater

---

45 ibid 124.

46 English Diabetes guideline group member 2 (Cambridge, 6th of May 2015)

than 35 was adopted. Several guideline group members describe how this cut-off point was debated, as well as whether patients should fulfil both the weight loss and blood sugar level decrease requirement or only one of them, as one guideline group member recalls: 'The weight cut-off was a major major bone of contention, (…) we weren't giving it [Exenatide] as a weight loss treatment, if someone has significant improvement in their $HbA_{1c}$ even if they had no weight loss (…) this is what we want.'[47] The safety concerns around Exenatide are no different for that drug then for any of the new medication, and so did not feature prominently in the reasoning of the guideline group.

Exenatide provides us with an example of the consideration of a new drug that changes the pathway of care, and how the reasoning of the guideline group is an interplay between evidence, cost, clinical knowledge, patient wishes and institutional constraints. The option to consider Exenatide was based on evidence, searched for through a process of clinical questions and careful assessment of the weight and strength of evidence fully in keeping with evidence-based medicine. But the appraisal of what this evidence means for clinical practice, and whether the drug should be seen as a positive addition to the arsenal of options available for treating diabetes, is one informed by clinical knowledge from people who have the experience of using the drug. These considerations were then pitted against cost-effectiveness data that has a different purpose, to find the optimal medication

[47] English Diabetes guideline group member 3 (Leeds, 20th of May 2015)

within given cost constraints, to enable a compromise that could satisfy the conflicting demands of those two types of knowledge.

Not all medication is treated in exactly the same way. Exenatide was particularly controversial and required extensive debate. Other medications were less controversial: Metformin is an example. The fact that Metformin is cheap, safe and effective and well tested in practice meant that the guideline group did not have to debate the position of Metformin in the guideline as extensively. This underscores the fact that the guideline drafters do not start with a tabula rasa but instead build upon earlier guidelines in a layered way, in which the only aspects that were changed were ones where new evidence, or other reasons, forced a change. An example of the other reasons forcing a change is the consideration of Rosi- and Pioglitazone in this guideline.

Rosiglitazone was highlighted in guideline discussions, not in relation to its effectiveness or costs, but in relation to its safety. The safety data which we have discussed in the evidence section of this case study provided the guideline group with a difficult dilemma. The crucial decision to be discussed was the precise weight that should be attached to the emerging findings about negative side-effects of especially Rosiglitazone. Would this mean the drug should no longer be prescribed at all? The group decided that the safety data did not warrant this far reaching conclusion. One of the members described

their reasoning thus: 'we weren't clear at the time whether the signal on heart attacks was genuine or not.'[48]

While another put it in this way:

> The Glitazone water was seriously muddied in the middle of doing this guideline (…) the Glitazone I think are a group of drugs that have split the diabetes fraternity (…) some people already have concerns in the group (…) I think as a result of the meta-analysis people wanted to be more conservative (…) I thought it was a really poor study, (…) the fact that studies had been excluded (…) that didn't have any cardiac events didn't make sense.[49]

Another guideline group member says:

> 'We put in some sort of weasel phrase (…) be aware of current safety concerns. (…) There was no evidence to do that [not recommend the drug] if you actually look at Stephens analysis (…) the absolute numbers where 72 versus 88 events, which was an increased risk of 44%, the relative risk looks huge, the reason why it was significant, was that Rosiglitazone was introduced on the basis (…) that we deal with insulin resistance, therefore we are bound to reduce cardiovascular events [heart attacks] as soon as (…) the fact that Rosiglitazone (…) increases [cardiovascular events] you have [a] ship pulled below the water line.'[50]

---

48 English Diabetes guideline group member 2 (Cambridge, 6th of May 2015).

49 English Diabetes guideline group member 3 (Leeds, 20th of May 2015).

50 English Diabetes guideline group member 1 (Warwick, 26th of May 2015).

This shows how the guideline group has weighed this evidence, which suggested risks and reasons to be vigilant, but this did not present enough risks to advice against using the medication. With the reference to the advice of the European Medicines Agency and the Medicines and Healthcare Products Regulatory Agency[51], the guideline group underscored the need for clinicians to stay up-to-date on the developments regarding Pio- and Rosiglitazone.[52] The example of Rosiglitazone, illustrates how the guideline group dealt with the question of the potential non-inclusion of a treatment in the guideline. Just like the inclusion of a treatment, a potential exclusion of a treatment from the guideline requires that a certain threshold be met before the group decides to include or exclude a drug. In this case, the evidence did not warrant an exclusion of Rosiglitazone, in the eyes of the guideline group.

The decisions taken in the guideline group are tailored attempts at reaching a consensus between a varied group of healthcare professionals. While discussions were always presented as ones about evidence they were complemented by knowledge of clinical practice, and acceptable compromises were reached between the members of the group on how clinical practice could deal with the choices on diabetes treatment.

---

51 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 123.

52 The market authorization of rosiglitazone was suspended (which means the drugs had to be taken off the market) in March 2010 by the European Medicines Agency because of the increased risk of myocardial infarction. EMA, 'European Medicines Agency Recommends Suspension of Avandia, Avandamet and Avaglim. Anti-Diabetes Medication to Be Taken off the Market.' <http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2010/09/WC500096996.pdf> accessed 1 September 2016.

Sometimes this pitted economic knowledge on cost-effectiveness against clinical knowledge of use of medication, other times safety concerns were prominent. An important tool for the group in reaching a common position after a long process of deliberation was the possibility of separating out distinct sub-groups of diabetes patients all with a separate treatment regime, this allowed for all the disparate views to be reconciled.

## 5.4 External Responses: Confrontations over Clinical Choices

The response to the guideline was muted. There was an article in response to the 2008 guideline that outlined the need for an update of the guideline because of the new safety data for Pioglitazone and the emerging data for the new diabetic drugs.[53] The new guidelines from NICE did not seem to cause much of a stir. So in this case any external responses to the guideline did not cause any changes to the text of the guideline, but it might have influenced the decision to quickly update the 2008 diabetes guideline.

This was very different when the draft of what became the 2015 NICE diabetes type 2 guideline was published.[54] This was because in the draft 2015 diabetes guideline Repaglinide was given a prominent place in treating diabetes, many healthcare professionals thought this was the wrong move. Among the criticism levelled was the fact

---

53 Stephen Atkin and Christopher Walton, 'Commentary: Controversies in NICE Guidance on Management of Type 2 Diabetes' (2008) 336 BMJ 1308.

473 'Type 2 Diabetes in Adults: Management | Guidance and Guidelines | NICE' <https://www.nice.org.uk/guidance/ng28> accessed 17 October 2016.

that the guideline group had only considered the cost of the medication, ignoring the fact that the drug was not often prescribed in clinical practice, had no licensed combination other than with Metformin so would have to be stopped if another drug needed to be added, and required frequent visits to GP practices to deal with adjusting of the dosage of the drug because of the variable effects the drug had. [55] The final guideline dropped Repaglinide, only mentioning it in a footnote, after a second consultation on the guideline was held.

This episode shows that in some cases, where a guideline directly contradicts clinical practice, NICE can face difficulty in only providing cost-effectiveness arguments as justification and was often forced to weaken or change it recommendations to accommodate criticism by clinicians.

---

55 Anthony H Barnett, 'NICE Draft Type 2 Diabetes Guidelines: A Cause for Concern' [2015] The Lancet Diabetes & Endocrinology <http://linkinghub.elsevier.com/retrieve/pii/S2213858715000534> accessed 5 May 2015; Jo Stephenson, 'Nurses Claim NICE Diabetes Guidance Will Put Patients at Risk' (*Nursing Times*, 15 September 2015) <https://www.nursingtimes.net/clinical-archive/diabetes/nurses-claim-nice-diabetes-guidance-will-put-patients-at-risk/5090329.article> accessed 17 October 2016; Caroline Price, 'NICE Risks Making Itself a "Laughing Stock" over Guidance on Metformin Alternatives, Say Experts' (*Pulse Today*) <http://www.pulsetoday.co.uk/clinical/diabetes/nice-risks-making-itself-a-laughing-stock-over-guidance-on-metformin-alternatives-say-experts/20009139.article> accessed 5 May 2015.

## 5.5 Conflicts of Interest: Potential Influences

Several members of the guideline group reported a potential conflict of interest: Annex E of the guideline provides details of these.[56] This includes being paid for lectures, sitting on advisory boards of pharmaceutical companies, and involvement in educational and charitable organisations that are sponsored by pharmaceutical companies. The question is whether these involvements have had any influence on the guideline.

In all the interviews conducted guideline drafters were asked about whether they felt that they had a conflict of interest and whether this influenced their recommendations. Perhaps unsurprisingly none of the members felt that the conflict of interest had any influence on their recommendations.

Some members of the guideline group stated that they feel a conflict of interest did not influence their recommendations in the guideline group because they work with all diabetes related pharmaceutical companies, and so do not favour one over the other. One of the guideline group members formulated this argument as follows: ‘you work for

---

[56] National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) Annex E.

everyone therefore you are beholden to no one.'[57] While another guideline group members echoed this: 'Everybody had every company, as a potential conflict of interest.'[58]

This argument is not wholly convincing as it neglects the potential bias that an association with pharmaceutical companies brings to the assessment of medications that are off-patent and where there is no campaign to promote these drugs by pharmaceutical companies. So even if guideline makers feel in this way they favour one company over another, there still might be a selection bias when it comes to deciding which drugs should be prescribed for the treatment of diabetes.

Despite this potential for selection bias, it is difficult to establish a causal link between the fact that a certain treatment has been recommended and the conflict of interest of one or more of the guideline group members. There is, however, a link between the consideration of the evidence on Exenatide in the guideline group, and conflicts of interest that provide some support for the argument that the group was (in)directly influenced by conflicts of interest.

The stress the group places on the weight loss gained by patients on Exenatide, and the reduced risks of hypoglycaemia for the newer diabetic drugs could be driven by the campaigning of the pharmaceutical industry to promote these aspects of the new drugs.

---

57 English Diabetes guideline group member 1 (Warwick, 26th of May 2015)

58 English Diabetes guideline group member 3 (Leeds, 20th of May 2015)

The problem, however, is that we do not know whether the group favours those drugs because they believe in the superiority of those drugs, based on their own independent reasoning, or because they are influenced by the conflicts of interest they have. And it is quite likely that the practitioners themselves do not know this either. We rarely ever question the source of our beliefs.

So, this study is restricted to arguing that it is plausible that the existence of conflicts of interests made the group look more favourable upon the newer diabetic medication. Existing literature in this field also points in this direction, showing a correlation between conflicts of interests and recommendations for drugs in guidelines, without being able to discern any direct causal relationship.[59]

## 5.6 Institutional Factors impacting on Guideline Drafting

In the development process of these two diabetes guidelines the institutional structure and processes of NICE played an important role. The drafting of two diabetes guidelines, the 2008 guideline and the 2009 short update guideline was the result of a decision by NICE to cut-up the assessment of diabetes care in two guidelines: one guideline to assess the older medications and the first licensed GLP-1: Exenatide and a second guideline to include

---

[59] James N George, Sara K Vesely and Steven H Woolf, 'Conflicts of Interest and Clinical Recommendations Comparison of Two Concurrent Clinical Practice Guidelines for Primary Immune Thrombocytopenia Developed by Different Methods' (2014) 29 American Journal of Medical Quality 53.

DPP-4 and GLP-1. Originally, NICE had intended to perform a health technology assessment to include the newer diabetic drugs but later decided to 'upgrade' the treatment to a short guideline. As one guideline group member explained:

> NICE did not want to include this [the new medication] in the scope because they thought it would subject to a separate technology appraisal. (…) In hindsight we should have a broader scope that included those drugs.[60]

This disjointed treatment of the guideline process led to some frustration for the members of the guideline group, which was expressed in the foreword to the 2008 guideline.[61] The result of this institutional set-up was that the 2008 guideline would be out of date, immediately after it was published. Already during the drafting process of the 2008 guideline it became clear there would be an update, which duly followed in 2009. Furthermore, the extensive drafting of the 2008 guideline took around two years to complete and only appraised one of the new drugs, while the short guideline was produced in four two-day meetings and covered more of the new drugs. This disparate treatment of these two guidelines were a direct result of the particular institutional choices made at the time and might plausibly have been related to the fact that NICE only has limited resources

---

60 English Diabetes guideline group member 2 (Cambridge, 6th of May 2015).

61 National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) Foreword.

and puts a limit on the amount of clinical questions that can be set within the scope of one guideline, for the 2008 guideline 32 clinical questions were identified.[62]

This shows one of the inherent difficulties in producing guidelines that is not particular to this specific guideline: they take time to develop by which the medical field might have changed substantially.

## 5.7 Costs: Cost-Effectiveness Modelling and Clinical Choices

The cost-effectiveness of treatments for diabetes is an integral part of the NICE guideline. For all the treatments provided, the guideline group searches for economic evidence in the form of a cost-effectiveness analysis of the treatment (compared to one or more other treatments). If existing literature is available this is used to estimate the ICER (incremental cost-effectiveness ratio) per QALY (quality-adjusted life-year) gained. Examples of this approach are the cost-effectiveness of: Metformin, Sulfonylurea and Rosiglitazone as second line therapy. The guideline group also commissioned a new cost-effectiveness study done to compare various third-line treatments for diabetes.[63] This new analysis is especially relevant for our study as this analysis informed the choice of the guideline group to give Exenatide a place in the guideline together with Rosi- and Pioglitazone and NPH-

62 ibid Annex A.

63 ibid Appendix C.

Insulin. The following results, given in table C20 in the cost-effectiveness report, provide the starting point for their discussions.

*Table 2: Costs and QALYs of Diabetes Drugs*[64]

| | **Total QALYs** | **Total costs (£)** | **ICER** |
|---|---|---|---|
| *Human insulin* | 8.05 | 19,527 | [First choice] |
| *Pioglitazone* | 8.03 | 19,828 | dominated |
| *Glargine* | 8.00 | 21,467 | dominated |
| *Biphasic analogue* | 8.05 | 23,033 | extended dominance |
| *Rosiglitazone* | 7.94 | 24,978 | dominated |
| *Exenatide* | 8.08 | 29,001 | £280,495 |

From this table it becomes clear the least costly option is Human Insulin, and the option with the highest quality-adjusted life years is Exenatide. Pio- and Rosiglitazone as well as Glargine are less effective than human insulin, have lower quality-adjusted life-years, and thus are 'dominated' by human insulin in the terminology of the model. Biphasic analogues are as effective as human insulin (both 'score' 8.05) but the costs are higher, in the terminology of this model termed 'extended dominance'. The ICER (incremental cost-effectiveness ratio) is a way of expressing what the cost would be if the NHS would pay for one quality-adjusted life year. In this case because an extra 0.03 quality-adjusted life

64 ibid Appendix C table C20.

year on Exenatide cost around £9,000 pounds to 'buy' a whole quality-adjusted life year would cost £280,495. This is far above the £30,000 threshold NICE has defined as their willingness to pay to gain one quality-adjusted life year. These results stay the same in several sensitivity analyses[65] and are consistent throughout. Human insulin is thus presented as the most cost-effective option for third-line treatment of diabetes. There is, however, one scenario in the model where Exenatide can be considered cost-effective, with an incremental cost-effectiveness ratio of £29,865 just below the threshold defined by NICE.[66] This is when Exenatide is compared with the other drugs under specific circumstances detailed in table C37 of the cost-effectiveness report.[67] In this analysis, it is assumed that Exenatide is successful in achieving a 5% weight loss over two years, while the other treatments cause weight gain. To express this difference the Exenatide patients are given a BMI of 27 in the model compared to a BMI of 33 for the other patients. It is further assumed Exenatide does not cause nausea. In these limited circumstances Exenatide can be seen as cost-effective, mainly because weight loss is assigned an added utility gain. The utilities in this scenario were based on a poster presentation by the (former) manufacturer of Exenatide.[68] It is this particular scenario, weight loss of 5% versus weight gain on insulin, which plays a big role in the clinical consideration of this economic

---

65 Table C23-C31 in Appendix C.

66 Table C37.

67 ibid Appendix C table C37.

68 'Contributed Poster Presentations' (2006) 9 Value in Health A24, A42.

evidence, as we have seen in the analysis of the discussions of the guideline group, particularly in the following section from the guideline:

> However the GDG's judgment was that costs of insulin and Exenatide by the end of the first year would be equivalent on average for people with a starting BMI (before these medications) of approximately >33 kg/m2, while in this obese group the small metabolic advantage to insulin on HbA1c would easily be outweighed by the metabolic advantage of 4 kg weight loss on Exenatide. In this restricted circumstance, and particularly at higher BMI's, the cost-effectiveness of Exenatide would then be at least as good as that of insulin.[69]

A striking fact is that all the treatments when assessed are so close to each other, with differences in QALY in the range of only 0.05 or slightly more in some of the sensitivity analysis. With such small differences, which all fall well within the margin of error for a model of this complexity, some of the underlying assumptions of the model require further scrutiny. And this is clearly what the guideline group members have done as the above quote illustrates.

One of the main reasons behind the very similar results for all the drugs is that the model used to calculate the cost-utilities only models the relationship between decreasing blood sugar and the complications of diabetes. And because the drugs are very similar in the extent to which they lower blood sugar the model provides these results. This also explains why cost-effectiveness modelling could not provide a definite answer in the eyes

---

[69] National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London (n 4) 122.

of the guideline group, because the modelling did not include weight loss, the fear of and cost of hypoglycaemia, and the costs of continuous blood glucose monitoring. The last two points could be addressed by adding these to the model, but the effect of weight loss could only be indirectly assessed, through modification of the parameters to mimic the effect of the drug. In this way the guideline group could square the circle by both accommodating the cost-effectiveness considerations, through the careful construction of a clinical scenario in which the drug could meet the formal requirements of cost-effectiveness and clinical need in having an opportunity to use the drug. Eventually with the passage of time the restrictions on the use of this treatment could be further relaxed, once the initial decision to give it a place in the treatment regime for diabetes patients had been taken.

The way the guideline group framed the cost-effectiveness evidence is very revealing, it shows how the guideline group was able to reconcile to conflicting sources of knowledge, and different interests. On the one hand the cost-effectiveness data pointed towards the conclusion the drug Exenatide could not be recommended for the treatment of diabetes. But because one specific clinical circumstance could be found where the drug, within the parameters of NICE policies, was cost-effective the drug could be recommended, trough tailoring the guideline.

## 5.8 The Dutch Guideline on Diabetes Type II

The Dutch guideline for diabetes was drafted by a guideline group supported by a member of the NHG and completed in 2013.[70] This was the third revision of the original diabetes guideline, the first guideline drafted by the NHG in 1988.[71] The guideline groups consisted of GPs with a specialisation in diabetes care, endocrinologists, and a professor of diabetes treatment in primary care settings.

## 5.9 Evidence-base for Dutch Treatments for Diabetes Type II

The treatment options presented in the Dutch guideline on Diabetes type II are wide-ranging. The first step consists of education on the disease and life-style advice. Losing weight and stopping smoking are important in controlling the blood glucose levels of diabetes patients. This aspect does not require any 'medication', however, is nonetheless considered by doctors to be a very important element of the treatment of diabetes. One of the GPs in the guideline group even provides running classes for his diabetic patients where he takes them running through the local village every fortnight. It is important to draw attention to this element of diabetes care, even though it does not consist of providing medication because it precedes all the further steps to be taken. As diabetes is a progressive

---

70 GEHM Rutten and others, 'NHG-Standaard Diabetes Mellitus Type 2 (Derde Herziening)' (2013) 56 Huisarts Wet 512.

71 'NHG Standaard Diabetes Mellitus Type II' <https://www.henw.org/archief/pdf/id10142-diabetes-mellitus-type-ii-nhg-standaard.html> accessed 6 July 2016.

disease most people will at some point require blood-glucose lowering medication. The various types of blood-glucose medication and the choices surrounding this treatment will be our primary focus.

*Blood-glucose Lowering Medication*

The Dutch guideline presents as the first medical treatment option, when blood-glucose lowering medication is needed, the drug Metformin. If this does not reduce the blood-glucose levels sufficiently one of the class of Sulfonylurea can be added with a preference for Gliclazide as first choice. If both of those medications are insufficient the patient should be given insulin with a preference for HPV insulin. This list of options seems to be relatively straightforward; there are three types of medication to be given consecutively. The list is noteworthy, however, for that what it omits. In the list, the 'new' diabetes drugs that have been developed in the last 10 years don't have a place.[72] The guideline does mention those drugs but it also states that they should only be used after the preferential drugs are used and have found to be not tolerated or ineffective.[73] This implicit choice against the new diabetic drugs warrants attention and explanation, to understand the main drivers and arguments for this choice.

---

72 This includes DPP-4, GLPT-1, SLDG-2 and newer forms of insulin.

73 Rutten and others (n 70) 2.

The evidence-base used in this decision is derived from a systematic review published in the British Medical Journal clinical evidence section which was co-authored by some members of the guideline group.[74] This systematic review was adopted by the members of the guideline group as the basis of their decisions on drug treatments for diabetes. Several group members explained that it was not necessary to do additional searching for evidence[75] on this topic as the systematic review was very extensive and doing additional or different searches would only unnecessarily duplicate results.

The fact that the systematic review was adopted was not surprising given the overlap between the members of the guideline group and the authors of the review. This, however, did not mean that the systematic review in itself could determine in which order to prescribe the medication, or which medication to prioritise over others. Guideline group members offered several different reasons for making the choice for the older diabetic drugs. One group of arguments revolved around concerns of safety and effectiveness. Put in the following terms:

> 'And we had an opinion about DPP-4 and GLP-1 because there was and is little literature on [it]. But we did have reviews of the total effect of those drugs, and that was very modest. We agree with each other very quickly if

---

74 Kees J Gorter and others, 'Diabetes: Glycaemic Control in Type 2 (Drug Treatments)' (2012) 2012 BMJ Clinical Evidence <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3462437/> accessed 3 February 2016.

75 With the exception of a small update for papers between the end of the search conducted in the systematic review (Feb 10) and the start of the deliberations of the guideline group.

> you don't know long-term effects and safety than we have to adopt a very conservative stance.'[76]

A selected and related reason brought forward is the cost of the new medication. This was stated in more general terms such as the drugs are relatively expensive and the guideline group members feel they don't know enough about their effectiveness and safety profile so we didn't choose them. While another guideline group member highlighted that the main reason was in the nature of the algorithm the guideline group wanted to provide to GPs:

> We talked about the nature of the algorithm we wanted to use. Would it be one where we set out extensively if this happens you do that, but we said it would become so complex, a horns nest, but we didn't want that. We did want nuanced targets for blood glucose (…) So we arrived at metformin, sulfonylurea and Insulin, and if you want another medication use your own judgement.[77]

The guideline also provides a list of 'other medication' that can be given if the preferred medication is not suitable. The list is in order of preference; the medication mentioned first is preferred over that mentioned second and so on. The second drug[78] on this list is Acarbose, which in contrast to the guideline in England is seen as a good and safe and cheap alternative to Gliclazide. The choice for Acarbose is interesting, the gastro-intestinal side-effects (flatulence) are not seen as so heavy as to warrant a rejection of this drug. One

---

76 Dutch Diabetes guideline group member 4 (Amsterdam, 11th of June 2015)

77 Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

78 After Repaglinide.

of the guideline group members who did have a PhD on this drug (under supervision of one of the other members of the guideline group), when asked about the inclusion of Acarbose, explained the choice as follows:

> I have not prescribed a single patient with Acarbose [in my career], I have nothing special with this drug. But that might happen now, because if I look at the evidence calmly and the safety, it is the least bad. (…) It is interesting to see that the small chance someone will die from a drug [with Pioglitazon] will be taken more easily [by prescribing GPs] then the small change of flatulence and diarrhoea, very remarkable, people think it is an embarrassing drug or something. But we know if you start the drug gently and up the dose carefully this is not much of a concern, and you can always stop giving it. (…) Especially if you compare it with the new drugs, the long-term effectiveness is unknown (…), as a patient I would never take it.[79]

While another member of the guideline group when asked about Acarbose remarked:

> It did not feature in the previous guideline. It also said do not prescribe. What has changed, it is a very funny finding: Acarbose was never really advertised in the Netherlands, so there isn't a group of doctors that is enthusiastic about Acarbose. (…) But after the new drugs arrived we asked ourselves what to do with Acarbose? (…) When we analysed it [it looked good].[80]

---

79 Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

80 Dutch Diabetes guideline group member 1 (de Bilt, 9th of June 2015)

This shows a certain pragmatic attitude to this drug: while none of the interviewed members of the guideline group were particularly enthusiastic about the drug, they did see it as a cheap and safe alternative for a small group of patients that would not be able to use Metformin. Whereas they felt that side-effects could be managed through careful incremental increasing of the dose of the medication.

*'Other' Diabetic Drugs*

Only if none of these drugs establish a satisfactory control of blood glucose levels the guideline suggests other drugs can be prescribed. The additional drugs that are mentioned in the guideline are Pioglitazone, DPP-4, and GLP-1.

*Pioglitazone*

The decision to list Pioglitazone as the last option deserves further scrutiny as it contrasts with the decision by the English group to give Pioglitazone a place as second or third line treatment for diabetes. Pioglitazone is evaluated very unfavourably by the group, some of the members even wanted to exclude Pioglitazone from the guideline, and to state explicitly that it was not recommended for the treatment of diabetes. A guideline group member took a strong view and said:

> I felt we should throw TZDs [ Pioglitazone] out, but X [different guideline group member] remarked I do not know why, (…) I thought he was the least reliable, I found it difficult to work with him, he said TZDs should be in there, but I disagreed. We don't even know how they [TZDs] work, the guy

> who invented them spoke on a conference, in a back room, (…) to explain it was actually (…) a different receptor (…) all together. This is a drug where we do not know how it works, and it has caused harm![81]

Other members of the group thought that Pioglitazone could have a role and therefore wanted it included in the guideline. The eventual compromise struck in the group was that if Pioglitazone would be mentioned, then all drugs with a market authorization in the Netherlands should be included in the guideline as one of the guideline group members explains: 'It is sold in the Netherlands so we should give advice on it. But then we said we should also say something about Acarbose, and Repaglinide.'[82] The guideline group had a negative view of Pioglitazone because it was not superior to other drugs such as Metformin and Sulfonylurea, and because for large numbers of patients the drug was not suitable. The therapeutic range of the drug was so small that they felt the drug should be a last resort.

The conclusion of the group was that DPP-4 and GLP-1 should only have a minor role in the treatment of diabetes; a combination of the lack of long-term safety data, no clear evidence of superiority and high costs led to this decision. One of the guideline group members was very sceptical about the pharmaceutical industry claiming superiority of the

---

81 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

82 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

drug because of a lower incidence of hypoglycaemia: 'The TECOS[83] trial came out, no extra cardiovascular deaths, yeah that's great that's the minimal you can expect that you don't die from the drug. But where is the good news? Hypoglycaemia? But 90% of patients don't have hypoglycaemia, what are we talking about.'[84] In contrast to the English group the Dutch group did not define smaller sub-groups of patients for which these drugs would be suitable, it also did not spend a lot of time and resources on establishing whether a DPP-4 was superior to a GLP-1, by putting all drugs in one table they shelved that problem.

The discussions in the Dutch guideline group were interesting in the approach to 'evidence' they developed. While discussions started from a close reading of medical studies relevant to the particular drug that was being considered, evidence was seen as a broad category that should include considerations of side-effects as well as safety.

*Target Levels of Blood Sugar*

The second large topic of discussion in the guideline group was the appropriate target levels for blood-glucose. Put simply, this is the decision in cooperation between patient and health care professional to decide on the levels of blood glucose that are to be seen as 'normal' for the patient concerned. This decision is important as a low level of blood-glucose

---

83 Darren K McGuire and others, 'Association Between Sitagliptin Use and Heart Failure Hospitalization and Related Outcomes in Type 2 Diabetes Mellitus: Secondary Analysis of a Randomized Clinical Trial' (2016) 1 JAMA Cardiology 126.

84 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

reduces the risk of micro- and probably also macro-vascular complications of diabetes and so increases quality of life, but if the level is too low mortality might increase. A balance has to be struck between these two options to approach an optimal solution for a patient. Ideally the blood sugar targets of every patient would be personal tailored optimally to his or her personal circumstances. In reality, however, this does not always happen, and regardless a guideline could never exhaustively describe all the circumstances that should be taken into account. So the Dutch guideline chooses to give target levels for blood sugar.

In the most recent guideline the guideline group has chosen to differentiate the target levels for the first time. The background for these changes are three new trials the ACCORD, VAADT and ADVANCE trials.[85] One of the members explains why they decided to differentiate the blood sugar targets: 'For older people it has advantages to have slightly higher target levels, it prevents harm, whilst for newly diagnosed diabetics you go for a strict target, and that's reflected in the guideline.'[86] The scientific evidence was robust enough to recommend the change. But it was not that the group followed the change in evidence and therefore changed the guideline on this point: the process took place in reverse. The clinical conviction among the members of the group, that glucose targets for

---

85 The ACCORD Study Group, 'Long-Term Effects of Intensive Glucose Lowering on Cardiovascular Outcomes' (2011) 364 New England Journal of Medicine 818; William Duckworth and others, 'Glucose Control and Vascular Complications in Veterans with Type 2 Diabetes' (2009) 360 New England Journal of Medicine 129; The ADVANCE Collaborative Group, 'Intensive Blood Glucose Control and Vascular Outcomes in Patients with Type 2 Diabetes' (2008) 358 New England Journal of Medicine 2560.

86 Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

the elderly have a different meaning compared to those for newly diagnosed people with diabetes, led them to search for evidence to back up this conviction, as this quote reveals:

> 'three or even four meetings, so much discussion about this. We searched a lot of literature and why: nobody excepted, everyone was deeply convinced that a blood sugar treatment of someone who is older than seventy has a completely different meaning than for younger people. And (…) we had some data from one of those big intervention trials that had just come out. And from the ZODIAC study (…) and we had some data from our own cohort. We had an idea that it [the evidence] is not so hard but so many indications that it increases mortality in the elderly, we have to do something with this fact. (…) This [the target levels] is not so evidence-based but this is the best we have. (…) The guideline is a compromise based on what we knew from the literature.'[87]

This shows how the idealised depiction of searching for evidence leading to changes in guidelines has to be nuanced. Guidelines changed without the evidence changing, as with Acarbose, or clinical convictions changed and evidence is presented as justification for this change after the fact.

The guideline group members quickly agreed on the need to differentiate the target levels but talked for a long time about the best way to present this new information in the guideline. One guideline group member noted: 'We went through more than five different flowcharts because we wanted to give the clinician an aid [to help them decide].'[88] And

---

87 Dutch Diabetes guideline group member 4 (Amsterdam, 11th of June 2015)

88 Dutch Diabetes guideline group member 1 (de Bilt, 9th of June 2015)

another remarked: 'We knew that if this flow-chart wouldn't work, people wouldn't understand it, they would throw the whole guideline away, so we had to get it right.'[89] The discussion centred on the most accessible way of presenting the new options, and several diagrams were trialled, to help the group decide what the best way of conveying the information to GPs and diabetic nurses was. This is an illustration of the fact that the guideline group was actively engaging with the implementation of their guidance, and this was facilitated by the lack of conflict or uncertainty about the assessment of the evidence. Particularly relevant in this regard is that this guideline was developed in conjunction with other 'products' by the NHG that facilitate 'protocol-driven' care by diabetic nurses, this was a part of a movement to organise diabetic care in so called 'care-chains'[90], to aid (further) substitution of hospital care (second-line care) by GP-led care (first-line care). Eventually the group was happy with the ordering of this part of the guideline, and endorsed the diversification of blood sugar targets.

A detailed consideration of two key themes that emerged in the guideline groups use of evidence shows how the Dutch guideline group defined their task. Their approach to drafting a guideline was steered towards producing a clear-cut guideline for use by GPs. The decision on the target levels for blood glucose show how evidence is used as a justification for a clinical belief, were both approaches are complementing each other in

---

89 Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

90 Dutch: Ketenzorg.

this case; where on blood glucose lowering medication prior assumptions on the importance of safety leads to a specific interpretation of existing medical studies.

## 5.10 Achieving Consensus: Internal Workings of the Dutch Guideline Group

The Dutch guideline group consists of several GPs with expertise in diabetes care, and several specialists who treat diabetes patients in hospitals. The group knew each other well as they worked together on academic articles, some members also had academic and professional relationships with each other, and others were involved in drafting earlier guidelines by the NHG on diabetes. All in all, one can say that this group was clearly socially close-knit, given that diabetes experts among GPs are thin on the ground in a small country like the Netherlands this is not wholly surprising. The closed character of the group is relevant, however, for our consideration of how the group reached consensus on key elements of the guideline. The members of the group shared a view of diabetes medication that was informed by a one vital assumption: that the long-term safety of diabetic medication is very important. One of the guideline group members put it in these words: 'You know the NHG has the principle that new has to be better than old, and safe in the long-term. We want to prescribe things that are safe, so we are very conservative in recommending new treatments if we don't know whether they are safe.'[91] This sentiment

[91] Dutch Diabetes guideline group member 1 (de Bilt, 9th of June 2015)

was shared by the majority of the group which steered their consensus on certain issues. This meant that in the treatment for the evidence of the new diabetic drugs, the absence of long-term safety data was presented as a critical factor. Another guideline group member reinforced this view:

> 'The critical points were safety and price, and you know it is not difficult to change gear, it is just like driving a car, you start in gear one if you reach a certain amount of revolutions of the motor you move up a gear. You start with the drugs that are cheap and safe and work, if they don't work you move up a gear. That's the philosophy (…) I still stand behind it.'[92]

The consensus on this issue shaped the decision to only recommend the old diabetic drugs: 'I think is a guideline group we were quite rigid, we don't want to give them [the pharmaceutical companies] that because it is too expensive for the healthcare system.'[93]

A further concern in their discussions was the potential for the pharmaceutical industry to use the guideline to say that the new drugs were now recommended by the NHG: 'We were very afraid that there would be too much room for the new drugs, but we also didn't want to deny GPs the room to deviate and give new drugs.'[94] and another said: 'I know from my colleagues that they were approached by the pharmaceutical industry, which they discussed in the guideline group.' (…) We were afraid the pharmaceutical

---

92 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

93 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

94 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

industry would [mis]use the standard, which did happen.'[95] This misuse of the standard that this guideline group member referred to was an advertisement by a pharmaceutical company that claimed their drug was in the guideline, which was a misleading statement because the drug was mentioned in the guideline, but not in their recommended algorithm which is what the advertisement suggested. The company retracted the advertisement after a complaint by the NHG.[96]

In addition, the group recommended Gliclazide, which was driven by the earlier systematic review that was conducted by the group, which showed a positive safety profile for the drug and its cardio-protective properties. A guideline group member remarked on this systematic review and the choice for Gliclazide:

> If you analyse the RCT that have been done that compare DPP-4 with a su[lfonylurea], that look at the risk of hypoglycaemia you see that those studies are mainly done with su[lfonylurea] that give a lot of hypo[glycaemia] Then you can observe an advantage for DPP-4. But if you take Gliclazide, that gives the least hypo[glycimae], there are only two studies! In those studies you observe a difference that is hardly clinically relevant.[97]

---

95 Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

96 'Klacht NHG Tegen MSD Gegrond Verklaard | NHG' <https://www.nhg.org/actueel/nieuws/klacht-nhg-tegen-msd-gegrond-verklaard> accessed 14 October 2016.

97 Dutch Diabetes guideline group member 1 (de Bilt, 9th of June 2015)

The group stood behind the conclusion that was drawn in this systematic review and adopted Gliclazide as the preferred Sulfonylurea. There was not a lot of doubt about the nature and robustness of the evidence, but this decision was extensively debated because it would mean quite a substantial change in clinical practice - as this guideline group member discusses:

> we talked much longer, not that we didn't agree with each other, but because it was quite a decision, because we now put in the guideline: you should use Gliclazide. That was quite a decision, because you say other drugs are less good or you shouldn't use them at all. (…) Glimipiride you could use, but we would rather you do not. That was quite something.' We didn't really talk about the evidence [but about the clinical consequences].[98]

This decision was (relatively) unique, because most other guidelines treat Sulfonylurea as a class, implying there are no (clinically) relevant differences between the drugs in this class. There was a sense of pride among the members of the group in making this choice; they felt this continued a (valued) tradition of cautious prescribing by GPs, with an eye on minimizing side-effects of drugs. One of the guideline group members placed this decision in a wider 'tradition' in NHG diabetic guidelines to critically assess Sulfonylurea for safety concerns, and how this squared with the logic of the hierarchy of evidence in evidence-based medicine:

---

98 Dutch Diabetes guideline group member 4 (Amsterdam, 11th of June 2015)

> As a guideline group you are looking at the best of what the pharmaceutical industry wants us to see. The waste [negative studies[99]] we don't see. (…) And those studies are designed to achieve that [positive] results. (…) So if you want to look at safety, which is important to us, (…) you have to turn to population studies, where the drugs has been used long-term in clinical practice, where you have a yearly follow-up to see what happens with those people that takes those drugs. But strangely in the evidence-based hierarchy this [those studies] have a lower value, even though if you look at safety it has much more value.[100]

The fact that the group decided to single out Gliclazide was therefore based on evidence of its superiority from the perspective of the lower incidence of hypoglycaemia and better cardiovascular profile. From the perspective of the hierarchy of evidence it is interesting to note that, as the guideline group members indicated, cohort studies and other lower rated studies play an important role in the evidence base in this guideline, moving beyond a focus on pure RCTs.

A third change the group reached consensus on was the differentiation of targets for blood glucose lowering. This debate in the group centred on the exact relevance of the various large RCTs conducted in this area and best way to interpret the findings of the ADVANCE, VADT and ACCORD RCTs, an English retrospective cohort study[101], as well

---

99 Meaning here studies showing results where drugs are not effective.

100 Dutch Diabetes guideline group member 1 (de Bilt, 9th of June 2015)

101 Craig J Currie and others, 'Survival as a Function of HbA1c in People with Type 2 Diabetes: A Retrospective Cohort Study' (2010) 375 The Lancet 481.

as the Dutch[102] and American[103] observational studies. The RCTs provided the main impetus for the conclusion that very intensive targets can be harmful. It is the observational data that provides the knowledge about the optimal window for blood sugar levels. In this instance we can see how a combination of trial and observational evidence provides a solid foundation for a differentiation of targets. In this instance, such combinations of evidence were the guiding force behind the decision to differentiate targets for blood sugar levels. The group debated the differentiation of the targets but was convinced by the need for a higher target for the elderly and those who have diabetes for over 10 years. The UKPDS data, based on a younger cohort of people who had newly diagnosed diabetes, supported the positive effect of a stricter target in that stage of the disease because this would be achievable without many side effects. In this way the evidence allowed the identification of several sub-groups of diabetes patients and the guideline was further tailored towards those groups of patients than had happened in the past.

The best way to communicate this decision was the subject of intense debate in the group. This element was unrelated to the medical evidence but was more a form of 'psychology' where the group tried to imagine the user of the guideline and the easiest way for him or her to read the guideline. In ordering the diagram in a clear way the group aimed

---

102 Gijs W Landman and others, 'The Relationship between Glycaemic Control and Mortality in Patients with Type 2 Diabetes in General Practice (ZODIAC-11)' (2010) 60 The British Journal of General Practice 172.

103 Elbert S Huang and others, 'Glycemic Control, Complications, and Death in Older Diabetic Patients: The Diabetes and Aging Study' (2011) 34 Diabetes Care 1329.

to make the choices as intuitive as possible. In this sense the drafting of the actual diagram (as well as the words in a guideline) extends beyond the 'evidence' into an effective tool for communication of complex knowledge.

The use of consensus regarding the three key decisions in this guideline show the interrelated and complicated way 'evidence', clinical knowledge, and preferences interact to produce answers. The differentiation of blood-glucose targets came about through a close analysis of the specificity of the various clinical trials and the reliability of observational data - and in this way comes closest to a pure application of evidence. But in this case it is still clinical expertise that drives the question whether the reported results are robust enough to warrant a change in practice and recommendation.

## 5.11 External Responses: Acceptance and Resistance

The response to the guideline was muted. Several organisations welcomed the guideline, such as Farma Magazine, that provides information to pharmacists[104], as well as on the website of Medicijnbalans an organisation subsidised by the Ministry of Health to provided objective information about drugs, where a video appeared outlining the various changes

[104] 'Herziene NHG-Standaard Diabetes Mellitus Type 2 | Farma-Magazine.Nl Onafhankelijk Platform Voor Apothekers' <http://farma-magazine.nl/nhg-standaard-diabetes-mellitus-type-2/> accessed 22 April 2015.

in the new guideline.[105] In general these comments showed several parties appreciated the simpler algorithm to follow, and were supportive of the choice to limit the use of the new diabetic drugs. Among GPs the guideline was well received which, was shown by the fact the drugs that the guideline recommends (Gliclazide, Acarbose, Repaglinide) show an increase in prescriptions, and the new diabetic drugs (GLP-1, DPP-4) show a decrease.[106]

As noted, one of the pharmaceutical companies used the guideline to claim their drug was now mentioned in the guideline, which is true but misleading as the drug was not recommended: they have later retracted this claim.[107] In addition one of the pharmaceutical companies published a 'contra-guideline' consisting of a folder detailing the NHG guideline, and the position statement by the EASD[108], suggesting these two were of equal weight. While in reality the EASD is financed by the pharmaceutical industry, and their position statement cannot be seen as a guideline, but is rather a statement of all the different available medication that does not provide guidance in the way a guideline does.[109]

---

105 'Medicijnbalans - NHG-Standaard Diabetes Mellitus Type 2, December 2013' <http://medicijnbalans.nl/video/40243> accessed 3 May 2015.

106 'Databank | GIPdatabank' <https://www.gipdatabank.nl/databank.asp?tabel=85_dia&geg=ddd&item=bijlage&label=00-totaal> accessed 12 October 2016.

107 'Klacht NHG Tegen MSD Gegrond Verklaard | NHG' (n 96).

108 European Association for the study of Diabetes.

109 'Hoe de Farmaceutische Marketingmachine Medische Richtlijnen Manipuleert' (*De Correspondent*) <https://decorrespondent.nl/467/Hoe-de-farmaceutische-marketingmachine-medische-richtlijnen-manipuleert/14363052-1f180d9f> accessed 22 April 2015.

There are several reasons that can explain why this guideline was well received and followed by Dutch GPs. The recommendations that the guideline make do not ask GPs to do something extra, such as providing a new treatment, but instead ask them not to do something: prescribe the new diabetes medication. Such advice is much easier to follow, it doesn't require any change of practice, or the investment of many resources. In accordance with the study of Burgers and colleagues[110] these are the circumstances that are conducive for guidelines effectiveness.

In conclusion external comments on the guideline did not alter its content, and the guideline seems to work well in practice.

## 5.12 Conflicts of Interest: An Anti-Industry Bias?

Three of the seven members of the guideline group reported a conflict of interest. Two had received research funding from pharmaceutical companies, one had joined several advisory boards of pharmaceutical companies. On first sight one would expect that this might steer the group into favouring newer diabetes medication and the newer generation of insulin drugs. The group also consisted of several people who had more of an anti-industry bias, who were very active in providing diabetes information to GPs without sponsorship by industry. It was these people that were dominant in the discussions in the group, and the

---

110 Jako S Burgers and others, 'Characteristics of Effective Clinical Guidelines for General Practice.' (2003) 53 The British Journal of General Practice 15.

recommendations closely reflect their view of the evidence. Several members were approached by the NHG to take part in the guideline group, because they had participated in guideline making for the NHG before. These individuals then suggested further names and approached other people they knew to take part in the guideline group. In this selection process people who were seen as 'independent' (from industry) were favoured. One of the guideline group members explains this in detail: 'there was a fixed core and several other people where asked every time to take part.'[111] Another commented: 'I was surprised, because I am no opinion leader (…) I asked my boss have you given this to me? But it turned out [guideline group member X] had suggested me, because he thought I would be independent.'[112] While another said: 'yes, I got a call from the chair, and I suggested some further names (…) it's very inner circle but that happens in this country.'[113]

The result of this selection process was that a majority of the members of the group shared common assumptions about the role of the GP in diabetes care, the value of new diabetic drugs and the relevance of blood glucose targets, which were rooted in a critical view of the pharmaceutical industry and a strong insistence on safety data. A guideline

111 Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

112 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

113 Dutch Diabetes guideline group member 4 (Amsterdam, 11th of June 2015)

group member explains that in the run-up to making the guideline he and another guideline group member became convinced they should break ties with industry completely:

> 'around the time the guideline was made X and I had the idea, we should have no ties with the pharmaceutical industry, I feel we were the most outspoken on this point. I had 2.5 year ago contact with X who said, if you are a guideline committee you should break with industry. And I replied but then I would amputate a part of my research? And he replied: that's making choices. And he had done it, not an unimportant man, gosh! I decided to break with them [industry] and spent the last half year sending letters to all of them, telling them I would no longer do research with them. At that time the guideline was ready, but I think it helped the discussions.[114]

The members of the guideline group realise that the way the group was composed favoured certain outcomes, given the fact that the Netherlands is so small all of them know of colleagues with different views on the new diabetes medication. A guideline group member put it thus: 'I certainly know colleagues who give talks where they are very positive about new drugs. (…) Whether they are just more into new drugs than I am or [they are biased].'[115] The guideline group members talked extensively about the various ways in which outsiders tried to influence them:

> I felt that the endocrinologist that were asked [to take part] were asked for their independent view on the pharmaceutical industry. We also get commentary from other specialists, were we could clearly see, oh this person always tells this story. Those people are used by the pharmaceutical industry to push their drugs. There was an acute awareness about how to

---

114 Dutch Diabetes guideline group member 4 (Amsterdam, 11th of June 2015)

115 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

> judge these comments. Those are people who go completely for the newest drugs and are always able to find the best available evidence for it.[116]

While another expresses his surprise at the fact that his participation in the group had leaked to the pharmaceutical industry: 'I was asked [by the NHG] to sign a confidentiality agreement, (…) I do so on Friday and on Tuesday I was speaking to someone who works for [pharmaceutical company X] who knew [expletive] I was part of the guideline group! (…) One of the other guideline group members must have done that…' And I told the group X already knew it.[117] And he goes further to detail how the company then went to try to 'help' him not to miss any trials: 'so I found a whole package of trials on my doorstep, so at least I wouldn't miss anything. With me it only happened with company X.'[118] He decided that because of a potential conflict of interest, he would not write part of the guideline that made recommendations on the drugs he was familiar with.

There was also a strong institutional dimension to the way the guideline group dealt with conflicts of interest, namely the view that the NHG valued independence very highly: 'I think if we had given a more prominent place to drugs were there was a conflict of interest that in that case there would have been more discussions [with the NHG].'[119] The guideline group members took this into account when they had their discussions. This was

---

116 Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

117 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

118 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

119 Dutch Diabetes guideline group member 1 (de Bilt, 9th of June 2015)

an institutional element of their conception of conflict of interest that influenced everyone in the group. Another interviewee remarked on the potential tension between the wish to employ the foremost experts and also wanting to have independent experts:

> I don't buy the argument that there are no independent people. It is true that people in the front of their field (…) will more often have conflicts of interest. We have tried to address [at hospital X] this through giving everything we get for talking to the pharmaceutical industry to a foundation we set up, and we use this to fund our own research.[120]

The effects of the industry through medical journals on the prescribing of doctors was seen as a problem by another member of the group: 'the latest issue of Diabetes Care was filled only with articles about diabetic drug treatments, DPP-4 and SGL-2, you read it, and you think this looks suspect, that's my worry, that's industry. (…) You trigger people and doctors are going to read it, and the average doctor will not read this critically, the industry clearly plays a role.'[121]

The discussions about conflict of interest with members of the guideline group were surprisingly candid: they openly talked about attempts at being influenced by the pharmaceutical industry, and for some of them, the decision to really make a radical break and no longer conduct any research funded by industry, even to the detriment of the research groups they were a part of. In many of their stories scepticism shone through about

---

120 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

121 Dutch Diabetes guideline group member 4 (Amsterdam, 11th of June 2015)

the efforts of the pharmaceutical industry to promote their new drugs. The vehicle to express this scepticism was a strong focus on safety, and safety evidence. In almost every question put to them about the new diabetic drugs, the answer included a reference to the lack of knowledge of the long-term effects and missing safety data.

The way the guideline group members dealt with the potential conflict of interest in this case shows more of an anti-industry bias than a pro-industry stance. This outlook influenced strongly how they read the evidence, by consistently insisting that safety was the overriding concern they came to different conclusions than the guideline group in England did. It also led them to not follow the hierarchy of evidence strictly, as proposed by evidence-based medicine, where RCTs trump other evidence but instead rely on cohort studies, and anecdotal evidence when it came to the safety of the drugs.

A further important element is the institutional environment the guideline drafting was conducted in. Several guideline group members spoke spontaneously about the critical view of the pharmaceutical industry within the NHG, and their expectation that they had to live up to those ideas, otherwise their guideline might be needed to be revised when they presented it to the authorisation committee. The committee is comprised of advisers that who have to accept a guideline, before it can be published as an official NHG guideline. This culture of a critical stance towards industry, which the guideline group members felt

a part of, was of course conducive for them to reach a consensus on not recommending the new diabetic drugs.

The combination of a small homogenous group, and an institutional environment which was favourable, worked together to steer the guideline towards a conservative course when it came to prescribing blood-sugar lowering medication.

## 5.13 Institutional Factors: Too Much Independence?

The guideline on diabetes is a slightly atypical guideline for the NHG. It was the first one made by the organisation, and the most developed. It went through several revisions in such a way that a new guideline group only changed those elements that warrant it. In this sense it is very different from guidelines that are designed from scratch in areas where there was no guideline before. A guideline group member explains the process:

> 'It is guideline in which much has already been done I decided to do the systematic search myself and write the justification. The NHG was in a process of making the process of selecting and analysis more systematic. I already had several search strings that I had used, that I could employ.'[122]

[122] Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

Another remark: ‘We did proposals for chapters, based on our research. (…) I felt I had a complete picture.’[123]

The way the Dutch guideline group approached the guideline process was materially different to the approach taken in England. Instead of a research question being set, literature being searched systematically and presented by the NHG, the Dutch group took a different approach. They divided the different sections of the guideline between themselves, based on their own expertise on those aspects of diabetes care. They conducted a literature search based on search string they made themselves, and produced a summary for the group to discuss. In this way the group independently searched for and appraised the evidence, with limited added assistance from the NHG. The fact that searches might be less systematic, as every group member employed their own way of searching for evidence, was accepted. The fact that there were already several systematic reviews, written by members of the guideline group that could underpin parts of the guideline might have played a role in this approach to the organisation of the group: there was less of a need for assistance by the NHG in developing research questions and search strategies.

This process provides a nuance to the more mechanical view of how guidelines ought to be produced: as this study shows that these processes do not always follow a

123 Dutch Diabetes guideline group member 4 (Amsterdam, 11th of June 2015)

regimented course but depend on the specific institutional history and group specific approach to evidence-based medicine.

## 5.14 Costs: Inclusion by Omission

The Dutch guideline does not include any cost or cost-effectiveness modelling. The prices of the drugs are not mentioned in the guideline nor is a formal assessment of costs and benefits of the choices in the guideline performed. Guideline group members when asked were consistent in stating that cost did not influence the choices they made, it also did not lead them to change their preferences as the following quote illustrates: 'yes cost played a role, but a small one. Yes, we were aware that we have a duty to make sure healthcare is delivered efficiently, and that includes cost-efficiency. But it was never a reason to make a choice.'[124] While another remarked on the absence of cost-effectiveness data: 'We haven't looked for cost-effectiveness research, this is not part of the NHG remit, so we didn't consider QALY's.'[125] But that does not mean costs were completely neglected in the guideline but they were used in a different way. The approach the guideline group took to those two new drug classes was summed up very well in the following anecdote a guideline group member told about advice he had given to other clinicians:

> During a case report evening, we evaluated a patient [where a GP had asked questions about the best strategy for the patient] (…) and if we feel that this

[124] Dutch Diabetes guideline group member 4 (Amsterdam, 11th of June 2015)

[125] Dutch Diabetes guideline group member 3 (Nijmegen, 17th of June 2015)

> a patient that should get a DPP-4 or GLP-1 he gets it. I feel this is how it should be, everyone who really needs it gets the new drugs, but no one who doesn't need it get the best drug. My data shows that less than 10% gets hypo[glycaemia] so 90% you can treat for sixty cents and for the 10% you spend €400.[126]

Guideline group members referred to cost for the healthcare systems as a whole, so in that sense had an incentive to avoid drugs that are too expensive, but this was secondary to safety concerns about the new drugs. The way costs were expressed was simply through calling drugs cheap or expensive. Because the new diabetic drugs were expensive and the 'old' ones cheap, this could count as an extra reason to be restrictive in prescribing them. A recent report by the Institute for Rational Medicine Use calculated that the decision to be restrictive with prescribing new diabetic drugs had saved the healthcare insurers and patients €335 million over a period of four years, compared to healthcare policies and prescribing decisions in Germany.[127]

The reason cost did not feature as prominently in the guideline has to do with the fact that decisions about treatments which belong in the package of care in the Netherlands, and which treatments should be funded by health insurers through their health insurance

---

126 Dutch Diabetes guideline group member 2 (Amsterdam, 10th of June 2015)

127 'Huisarts Boycot Dure Diabetespil, Farmaceuten Lopen Geld Mis | Economie | de Volkskrant' <http://www.volkskrant.nl/economie/huisarts-boycot-dure-diabetespil-farmaceuten-lopen-geld-mis~a4351320/> accessed 4 August 2016; Insitituut voor Verantwoord Medicijngebruik, 'Kosten Bloedglucoseverlagende Middelen Nederland versus Duitsland' <http://www.medicijnbalans.nl/content/userfiles/files/Kosten%20bloedglucoseverlagende%20middelen%20Nederland%20versus%20Duitsland(1).pdf> accessed 4 August 2016.

packages is determined by the Zorginstituut. There was therefore no expectation that if care was described in the guideline this would also be funded under the care package healthcare insurers have to provide, as this decision rests with the government.

This marginal role of cost in the guideline could be seen as positive from the perspective of the NHG. It is more likely that these guidelines by the NHG will be seen as made *by* healthcare professional *for* healthcare professionals, and not as instruments of healthcare rationing. The fact that the prescriptions of Gliclazide and Acarbose increased, and the new diabetes medication decreased, after the guideline was published makes it very plausible the guideline was successful in its aim.[128]

## 5.15 Comparative Analysis: Differences Dominate

When comparing the English and Dutch guidelines a picture emerges where differences are predominant. The reasons for this are multi-layered and go beyond (mere) differences in evidence. In analysing these differences and similarities, the complex nature of the factors that drive them will become apparent.

128 'Databank | GIPdatabank' (n 106).

*The Evidence-Base: The Importance of Framing for the Construction of Evidence*

Diabetes guidelines and the evidence underpinning them throw up some interesting questions. Despite the greater convergence in number of studies the two guidelines have in common, there is greater divergence in recommendations than in the case of the lower back pain guideline. Why this is the case it is complex, several factors explain this when taken together, including: the year in which the guideline was drafted; the process of selection of the evidence; the frame through which the evidence was viewed; and the role of the guideline in the healthcare system studied. While all these factors affect each other as well, and are linked to other aspects of the guideline process, nonetheless here they are presented as analytically distinct to show how they work more clearly.

The English guidelines were published in 2008 and 2009, and their drafting process started in 2006. The Dutch guideline was published in 2013, while drafting started in 2010. While this few years difference might seem small, in this case they are significant for some aspects of diabetes care. Between 2009 and 2011 the evidence on target levels for blood glucose shifted with the publication of important new trials, Rosiglitazone was taken off the market, and more safety data emerged on Pioglitazone and the GLP-1 and DPP-4 drugs. The different targets for blood glucose lowering could thus potentially be explained by differences in the evidence-base. The fact that the new 2015 NICE guideline[129]

[129] 'Type 2 Diabetes in Adults: Management | Guidance and Guidelines | NICE' (n 54).

differentiates between two targets (for those with only diet or Metformin and those with more oral medication) in contrast to three targets described in the NHG guideline that was published in 2013, shows that this is an area where clinical interpretation and preferences is an important factor, in addition to evidence.

A further important structural difference between the two guideline groups lies in the process of selecting evidence. The English guideline group follows a strictly prescribed formal searching of the literature, and has an extensive process of eliminating lower quality evidence, in such a way that RCTs are most important in their evidence. The Dutch group is less systematic in their searching of the evidence, but more importantly they see cohort studies as more important and valuable than RCTs for some of the questions they faced in the guideline. This resulted that the English recommendations reflecting closely the RCT data, while the Dutch group was steered in a different direction through the reliance on cohort studies.

Following from these two points is the fact that the two guideline groups use a different frame for analysing their evidence; and this is potentially a decisive factor in explaining the observed differences. This frame can be described as a set of principles that were important for the guideline group in evaluating the evidence, which steered the extent to which they led evidence influence their recommendations. The Dutch group cited safety as an overriding concern, in reading the evidence for the new diabetic drugs the lack of safety data was therefore decisive. The English group framed their discussion through the

concepts of patient-centred care and choice for clinicians, and viewed the new drugs more favourably as a result.

A further factor that clearly influenced the guideline groups in different ways was the position of the guideline in the wider healthcare system. The Dutch guideline made by GPs for GPs, and is a clear expression of the desire to provide practical advice for use in clinical practice, while at the same time providing a counterweight against the pharmaceutical industry. The English guideline is multidisciplinary, and has to pay heed to issues of cost and cost-effectiveness, and is therefore more likely to be seen as a tool for healthcare rationing.

*Consensus: Homogeneity is Key?*

When analysing the way consensus is reached in the guideline groups the homogeneity of the Dutch group is an interesting factor, as this seems to help the group to reach consensus quickly on important parts of the evidence, while focusing attention on communication of the various messages in the guideline. It is important to note that the Dutch group used consensus not just to supplement the evidence, as is envisaged in normative models of evidence-based guideline making, rather they sometimes used evidence as a rationalization after the fact. A good example of this is the new targets for blood sugar levels, where the clinical wish to change practice, mainly for the elderly, led to a search for evidence to help support that position. The relative homogeneity of the Dutch guideline group also helps to explain that the positions they took are more radical and clear cut, they have made fewer

compromises. This is because they shared a body of knowledge and shared common assumptions about safety, cost, and efficacy of diabetes treatment.

The English group was more diverse, and less close-knit: more compromises were necessary for a guideline group that had to work with less evidence, and there was more pressure to keep the cost of medication under control. They made compromises that are visible in the text of the guideline, through the method of defining smaller sub-groups of patients for which they advise a certain treatment, thereby avoiding excluding treatments all together while keeping the costs within agreed limits.

*Professional Pressures to Conform: The Role of External Actors*

Neither guideline drew much criticism or commentary after publication. But that does not mean that external actors do not play a role in influencing guideline making groups. In this case we can infer indirectly that guideline group members are influenced by their perception of 'normal clinical practice.' When a guideline steps outside this norm, as happened with the draft of the 2015 NICE guideline, pressure for various groups will make sure this is 'corrected.' It seems that only with very strong and overwhelming evidence is it possible for guideline groups to recommend radical changes, otherwise they will be confronted with resistance in various forms.

When we analyse the various conflict of interest that guideline group members report it becomes clear that in both groups a large minority of participants have one or more conflicts to declare. These written declarations only tell a story on paper: through interviews with various participants it was possible to establish that these conflicts play out in very different ways.

The Dutch group employed an anti-industry bias in its reading of the evidence. Instrumental was the fact that several members of the group worked for a foundation that had a mission to provide objective diabetes information to GPs, without sponsorship by industry. Some members of the guideline group who declared conflict of interest decided (before they joined, or during the process), to end all ties with the pharmaceutical industry. This shared position shaped their reading of the evidence, and influenced the way in which the group reached consensus.

A further aspect, that insofar has been neglected in the literature on this topic, is the strong institutional dimension of how the group felt they should approach the evidence and deal with conflicts of interest. There was a strong expectation by the NHG, that new drugs

should be evaluated critically, and so the group felt supported in taking the stance that the new drugs were not recommended for the treatment of diabetes.

The English group leant more towards favouring the pharmaceutical industry in their reading of the evidence. It does not necessarily follow that this was because they had conflicts of interest, but it is plausible this partly explains their enthusiasm for Exenatide, and the successful attempt to keep this drug in the guideline, despite cost-effectiveness evidence which showed an unfavourable picture.

*Institutional Factors: Producing Evidence (In)Formally*

The fact that one set of guidelines was produced under NICE, and another by the NHG, obviously made a difference. Their separate and distinct institutional histories and culture influenced the guideline groups in several ways.

The English guideline group had to follow the NICE manual for producing guidelines[130], which required the consideration of economic evidence on cost, set out a strict procedure for the definition of research questions, selecting evidence, and appraisal. The group was confronted by this process through the decision by NICE to keep some of the new diabetic drugs that were licensed at the time out of the scope of the guideline. This

130 'Developing NICE Guidelines: The Manual' <https://www.nice.org.uk/process/pmg20/> accessed 15 August 2016.

decision caused some disquiet in the group because they feared – and were proven right – that this would necessitate a quick update of their guideline, as it would be out of date just after publication.

A further consequence of this institutional regime is exposed in the selection of evidence by the guideline group. The evidence the group relies on is primarily (but not exclusively) RCTs, with a focus on the data of the United Kingdom Prospective Diabetes Study (UKPDS).

The Dutch diabetes guideline group functioned in a more informal institutional environment, where the guideline group members had more control over the searching of the relevant literature, and where no cost-effectiveness data was included in the deliberations. This made the process of selecting literature less systematic than was the case in the England. However, importantly the guideline group relied on more observational data from Dutch, Danish, English, and American studies to underpin their choices, especially with regards to the safety of blood sugar lowering medication.

*Cost and Cost-Effectiveness: Economic Evidence Creating Different Pressures*

Cost-effectiveness modelling had an effect on the discussions of the English guideline group. They considered economic evidence, especially when it came to considering the appropriate third-line therapy for diabetes. The negative cost-effectiveness of Exenatide forced the guideline group to devise restrictive conditions for its use, to square the

conflicting pressures of costs with the clinical desire to employ the drug. The Dutch guideline group did not employ any cost modelling. They did, however, formulate a more general rule: that their choices should not be costly for the health service as a whole.

This led to the paradoxical result that the English guideline recommends drugs which are costly in more cases than in the Dutch guideline. This suggests that the threshold defined by NICE to conceptualise cost-effectiveness had the unintended result the group recommended more expensive medication, because they found an instance where this drug was (just) cost-effective.

*Conclusions*

The Dutch and English diabetes guidelines were produced to provide guidance in a difficult and quickly moving area of clinical practice: diabetes care. In several critical areas, such as oral blood glucose lowering medications and target levels for blood sugar, the guidelines differ from each other. This chapter has focused on explaining the reasons behind these differences. This has presented a multi-layered picture of various interlocking factors that have steered the two groups towards adopting recommendations that differ.

In conclusion diabetes guidelines show how guidelines are a focal point of many different forces. Patients, medical professionals, and government all exert pressures to achieve the results they would like to see. The guideline groups deal with those particular pressures in their own way, which explains why their guidelines differ in important

respects. The significance of these findings for the theory and practice of evidence-based medicine will be discussed in the next chapter.

# 6 GUIDELINES IN A COMPARATIVE SENSE

Chapters 4 and 5 showed that while there are some similarities between guidelines, there are also some important areas of difference between medical practice guidelines in the Netherlands and England. This chapter draws together the findings from the case studies presented in those chapters and considers the implications drawing upon all four different cases. This chapter outlines six key areas for analysis that have emerged from the case studies: evidence, consensus, institutions, external factors, conflicts of interest, and cost-effectiveness.

## 6.1 Evidence: Social and Institutional Construction of Objectivity

A central aim of this study has been to understand how guideline developers select, use, appraise and understand medical evidence in constructing guidelines that purport to be evidence-based. This study has endeavoured to add to a more nuanced understanding of evidence, conceptualising it as a social and institutional construction used in a collaborative and communicative process aimed at creating 'objective facts.' This approach rejects other theories that offer either normative description of how evidence should be used, or see evidence as the mechanical application of knowledge to specific medical cases. This

section examines the results of the study as they relate to social and institutional choices in evidence selection, appraisal and use in the two case studies and across all four guidelines.

The case studies provide a varied image in how the different groups deal with evidence. The first factor that explains some of the difference between the Netherlands and England is the level of formality employed in searching for medical studies in medical databases. The institutional environment of NICE provides a more formal way of searching evidence, as they have a dedicated team of technical support staff and specialists on searching the medical literature. This results in a more formal selection of evidence in the guidelines in English case studies. A more formal way of searching for evidence in theory should produce more varied evidence to support decision-making. However, in England the opposite was observed as guidelines were actually based on fewer studies and fewer types of studies than the Dutch guidelines. This was caused by the fact that the 'search strings', which are the words that are used to search these medical databases, were themselves based on specific clinical questions that had been formulated. The search strings copied very closely the precise words employed in the specific clinical questions posed by the guideline group. As the search strings were defined narrowly to very closely mimic the clinical question formulated, they turned up fewer results that met the inclusion criteria. This led to the rather paradoxical result that following evidence-based methods, employing a systematic searching methodology, led to less, and less varied, evidence to consider. The specific ways in which in these clinical questions were formulated was heavily influenced by the methodology prescribed in the NICE guideline manual. As will

be discussed, this manual, the clinical questions, and search strings were all institutionally shaped and socially constructed: meaning that the selection of evidence used in English case studies was directly impacted by institutional and social choices.

The Dutch guideline makers had less institutional support, and methods for searching for evidence were employed in a more informal way. The drafters of the lower back pain guideline read and appraised every study themselves which forced them to consider only RCTs and systematic reviews that could be managed by the group. They employed an informal process among group members checking with each other whether anyone knew of studies that had not been mentioned, in addition to searches that were done by the NHG. The Dutch diabetes group took a different position: they divided the task of searching for literature up among themselves, with everyone designing their own search strings, while also using an existing systematic review members of the group had written for the British Medical Journal.[1] This difference in approach meant that the Dutch guideline groups used more varied evidence, included more observational studies, and were open to other types of evidence - especially if studies were conducted in the Netherlands.

As a result, the Dutch groups did not fully follow the evidence-based approach to searching for medical studies as was done in England – but their guidelines can be said to

---

[1] Kees J Gorter and others, 'Diabetes: Glycaemic Control in Type 2 (Drug Treatments)' (2012) 2012 BMJ Clinical Evidence <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3462437/> accessed 3 February 2016.

be based on more varied evidence, than their UK counterparts. This shows how in the selection of evidence more factors play a role than a purely mechanical searching of medical databases. Because this searching is a result of a social and institutional process of selection, and the outcomes of the searches are determined by the questions already set by the group and influenced by the institutional structure surrounding the guideline makers. This gives the various guideline groups different starting points right at the beginning of the guideline process.

As well as selection of evidence, the differences and similarities in recommendations are the result of a more fundamental difference in *framing* between different guideline groups. The various guideline groups all employ different frames in assessing or appraising the evidence, so even if all the evidence they viewed had been exactly the same, they would have still drawn different conclusions from it. This is best illustrated through a comparison of the appraisal of evidence by the diabetes guideline groups in England and the Netherlands. Appraising the evidence happens after evidence has been selected, and involves the process of attributing meaning to the results drawn from the medical studies selected from the searches. For example, the Dutch guideline group recommended Gliclazide, while the English group did not differentiate between the various Sulphonylurea family of drugs.[2] One could argue this might be due to the fact the Dutch

[2] Neither group recommended Glibenclamide, so this drug is not relevant here.

guideline group had different evidence to review, as they base their recommendations on a study conducted in 2012 whereas the English group could have had less evidence to review as their guideline was published in 2008. But this misses the point, as they both refer to studies available before 2008 and make this particular choice based on those studies. The Dutch guideline groups actually made their choice based on their assessment that Gliclazide had a positive safety profile. The evidence in this case was guided by the framing that prioritised the safety of diabetes medication over other considerations.

This specific clinical approach, which stressed safety issues, lead them to adopt a frame where the appraisal of evidence was geared towards satisfying them that the drug chosen was the safest available. So they framed the clinical question in this way: Which of the Sulphonylurea drugs is the safest? This led them to recommend Gliclazide. This does not mean the English guideline group was wrong not to further differentiate between the different drugs in this example. This normative statement is not supported by this example rather this example is used to outline a mechanism behind different choices: a different frame leads to a different appraisal of evidence, and therefore results in different recommendations. All guideline groups make the claim that the recommendations they make are directly based on evidence, while this set of examples has shown that there is a social and subjective process of framing that mediates the process of evidence choices.

In this way this thesis further extends earlier work by Cambrosio et al. on 'coordination regimes'[3] by offering further insight on how this process of coordination takes place within a guideline group, by offering that a process of framing makes this coordination activity possible. Furthermore, a different question left open by Cambrosio et al. is how the process of creating (socially constructed) facts can be seen as objective by outsiders. Later on in this chapter the argument will be advanced that the perceived objectivity of these facts lies in the rhetorical power of evidence-based medicine as an approach to generating scientific knowledge.

The process of negotiating a guideline consists of a complicated set of trade-offs and necessary compromises in an attempt to satisfy all the different audiences of guidelines. Guideline makers have to position themselves in a field of diverging powers: government interested in healthcare rationing, patients looking for the newest treatments, pharmaceutical companies trying to get their product recommended, and clinicians who look for usable advice in clinical practice.

The selection and appraisal of evidence is intimately linked to the negotiation and process by which evidence choices are made. A real understanding of the role of 'evidence' starts with the realisation that evidence is always used for a purpose. This purpose is

---

[3] Alberto Cambrosio and others, 'Regulatory Objectivity and the Generation and Management of Evidence in Medicine' (2006) 63 Social Science & Medicine 189.

discussed and established during the process of guideline creation. 'Evidence' can therefore not be seen separately from the process of negotiating in a guideline group. Negotiating of evidence is highly dependent on specific cultural and institutional contexts, and is embedded in a multi-layered frame of understanding that is created in every guideline group. So it can be said that it is not the evidence that drives recommendations, but the complicated interrelation between social and institutional forces.

Another example that further illustrates this argument is the contrasting conclusions the English and Dutch guideline groups have drawn from looking at exactly the same evidence on educational materials for lower back pain. While the English guideline group focussed on the fact that the results of the medical studies of educational materials were not statistically significant, the Dutch group focussed on the clinical relevance of the results. They adopted different frames, as what they wanted the purpose of the evidence to be differed: the English group was looking for robust knowledge that was methodologically sound, while the Dutch were more interested in practical knowledge that was usable. In this case, these differences in approach and frame actually led both groups to the same recommendation. The both recommended that educational materials should have a role in treating lower back pain. The English group based this recommendation on consensus between the guideline group members rather than the medical studies. The Dutch group did base their decision on the same medical studies, because they found value in the practical use of the knowledge the studies provided. It is in the confrontation between

evidence and a frame that a reading of the medical studies leads to the construction of social facts.

Examples drawn from all four guidelines show that for all guideline group evidence is of central concern. But how exactly the meaning of evidence is constructed is a process of framing that is different and unique for every guideline group. The primary reason for that difference is that evidence can only be understood and applied in a guideline once a guideline group has collectively trashed out what the primary purpose of the guideline is going to be. In other words, every guideline groups creates its own frame to view the evidence through that is unique, while also exhibiting common features.

The relationship between a frame and the evidence is not a static one, as these examples demonstrate. It could be conceptualised as a continuum with on the one hand the Dutch diabetes case where an overarching frame had a large influence on all aspects of evidence selection and interpretation to the English back pain case were participants who were far more neutral about what they wanted and did not formulate any particular direction so their framing had a less decisive effect on the evidence.

This thesis argues that evidence is a tool used by guideline groups in making clinical choices, which are also normative and political choices. This section has shown that these choices are social in nature and are the result of institutional shaping and social construction. The claim by decision-makers that guidelines are evidence-based is not

meaningless because this claim allows guideline groups to suggest that the results of their deliberations are objective. But they are not entirely evidence-driven choices, in that sense they are subjective. Guidelines are intensively subjective, necessarily compromised by a multitude of conflicting pressures, and encapsulated in institutional frames. These frames are multi-layered in the sense that they incorporate concerns that go beyond evidence alone. Multi-layered framing includes influences from institutional environments, conflicts of interest, the composition of the guideline group and cost considerations: these aspects will be further explored in the next sections.

## 6.2 Consensus: Different Mechanisms throughout the Process

To understand how guidelines are made we need to appreciate the various elements that influence how guideline groups reach consensus. Consequently, how this consensus feeds into the various recommendations in a guideline. For the purpose of this thesis, the term consensus goes beyond the social negotiations and agreements and understands the role of consensus in directing policy-making and the wider considerations that made the guideline groups move from the medical evidence base to making recommendations. This section will analyse how the Dutch and English guidelines are influenced by the aim of the guidelines, understandings of disease, and considerations of cost, clinical practice, and institutional framings surrounding the guidelines. This section will then make the argument

that these factors all influence the process of reaching consensus and shape the resulting guidelines.

The consensus between various participants in guideline groups is essential for its success. And the only way for guideline groups to achieve consensus is to make compromises. This study has detailed various examples of compromises made in various guideline groups and we can draw from these examples to establish two mechanisms that drive guidelines on the same conditions apart, and two mechanisms that work to make guidelines on the same disease similar. This study focuses on the importance of the outcome of consensus-based decisions on the text of the guidelines, and how this is informed by the purpose of guidelines as a form of professional self-regulation.

*Purposeful guidelines*

The way the members Dutch and English guideline groups thought about the purpose of the guideline they were drafting influenced the goals and objectives they adopted during the negotiations and process of creating a guideline. Guidelines will differ depending on the aim of the guideline in the wider healthcare system. The NHG guidelines are drafted by one specific profession (GPs), and are exclusively aimed at providing guidance for the daily clinical practice of GPs. Furthermore, these guidelines provide an important element of training for new GPs, and provide a way for the profession to secure it status as a specific and specialist profession among the other medical professionals. For example, for the

drafting of the Dutch guideline on diabetes the focus of the recommendations was specifically on the clinical role of GPs in managing diabetes care and did not consider other diabetes related healthcare concerns. Importantly, the group drafting the guideline were predominately GPs with clinical experience. As a result, the consensus they reached was informed by their implicit agreement with the consensus already formed on the goals of the guideline.

Guidelines made by UK NICE have a different purpose, and these guidelines are also perceived differently by medical professionals. NICE guidelines are always multidisciplinary; several different groups of medical professionals were represented in the guideline group. In the case of the NICE guideline the goals of the group were less clearly focused as they had to provide recommendations that would cover the whole span of the delivery of care of diabetes patients. This meant the NICE guideline was not the statement of a specific professional group, as in the case of the Dutch GPs, and therefore medical professionals might not feel the same allegiance to them. Due to the necessity to satisfy various different medical groups, this guideline was characterised by far more compromises to help secure an agreement on thorny issues. The implication of this social outcome is that guidelines might differ because there was a need to accommodate diverging views on the purpose of the guideline, rather than them being dictated by medical evidence or clinical necessity.

*Cost motivations*

Convergence between English guidelines is driven by cost motivations that play a role in the deliberations in guideline groups in England. NICE guidelines are therefore seen to be instruments of healthcare rationing. This aim of healthcare rationing is served through the inclusion of cost-effectiveness as a criterion in the discussions about recommending treatments. An identifiable threshold for cost-effectiveness also provides a yardstick to help parties decide whether they can recommend a treatment or not. The fact that in practice guideline groups sometimes deviate from this target of £20,000-£30,000 per quality-adjusted life year gained, as the results of the diabetes case study show, does not negate the fact that this requirement gives the guidelines a different position. This translates into different recommendations by guideline groups, as we have seen paradoxically in the case of the diabetes guideline the English group recommended more expensive treatment than the Dutch group. While it has been shown that issues of cost-effectiveness directly impacted on the recommendations in the English case, as it feature heavily in their discussions which resulted in compromises that restricted access to expensive medication.

The Dutch guidelines do not formally include aspects of cost-effectiveness as these are dealt with by the Zorginstituut, a body separate from professional associations like the NHG. This is significant because this changes the way the Dutch guideline group reaches consensus, aspects of cost play a far more marginal role in their discussions. It also reinforces the view by other professionals that the advice given in the guideline is only

driven by clinical and evidence driven considerations, and so the NHG avoids being seen as providing guidelines that are 'imposed from above'.

*Organisation of Care*

Another aspect that helps explain how consensus can drive differences between guidelines is the discretion that guideline groups have to decide on the organisation of care. In drafting a guideline, decisions have to be made about which healthcare professionals can perform specific elements of the treatments of conditions: this is referred to as the organisation of care. NICE guidelines specifically exclude saying anything specific about organisation of care, restricting itself to indirectly influencing the influence of care through the prescription of training for healthcare professionals.[4] NHG guidelines do take an explicit line on the organisation of care, which is easier as their guidance only concerns GPs, and it can only indirectly affect other professionals if the guideline recommends not to refer a patient to a group of medical professionals.[5]

This aspect had some limited influence on the development of the guidelines, where the Dutch diabetes and back pain guidelines were more clearly concerned with (re)organising the organisation of care delivered by GPs, in contrast to the English

---

4 'Developing NICE Guidelines: The Manual' <https://www.nice.org.uk/process/pmg20/> accessed 15 August 2016.

5 As NHG guidelines are only written for GPs, they cannot influence the guidelines of other professionals directly.

guidelines that covered a larger range of medical specialities. Those aspects were however relatively minor in the overall text of the guideline.

*Clinical Practice*

Clinical practice is one key area that informs consensus among guideline group members. A key finding in this study is that guidelines are more similar if the consensus in the group is directed towards a replication of current practice. On the surface this might look like a statement of the obvious, if guidelines replicate current practice, and this practice is similar in the two countries, consequently the guidelines will be similar too. The defining factor, however, is that first a guideline group reaches consensus on the desirability of achieving a result that is close to current practice in the field, and in this light the evidence is analysed. So the formal process of evidence-based medicine is reversed. It is not the painstaking search for evidence that then informs the decision to choose to stick with current practice, rather as the evidence does not provide answers to which therapy is superior so it is wise to maintain the status quo. Rather than this being about clinical practice and evidence, the decisions made by the group were based on maintaining status quo in their recommendations and a prevailing attitude of consensus and agreement to comprise.

The Dutch lower back pain guideline provides a good illustration of this process. Early on in the process of negotiating this guideline several members left the group, mainly because they could not agree which of the possible treatment for lower back pain was the

best. As a result, the group adopted a pragmatic view, where treatments were adopted as equal, even if some differences could be seen in the evidence underpinning them. As a result of the adoption of this principle to regard treatments as equal, consensus was achieved and as a result the Dutch guideline changed clinical practice in delivering lower back pain in only minor ways. This demonstrates that social consensus processes are linked to the outcomes of recommendations of guidelines.

This influence of clinical practice is not invariably directed towards replicating the status quo. A diametrically opposing case could be made when analysing the framing adopted by the Dutch diabetes group. While the Dutch diabetes guideline was conservative in only recommending a limited range of treatment this was driven by a desire to *change* clinical practice. The change in this case consisted of *reversing* the existing trend to prescribe newer diabetes medication.

It is not always possible to predict whether a guideline group will adopt a stance towards replicating current practice, or whether they seek a progressive guideline. Our case studies throw up somewhat surprising results. Both diabetes guidelines are progressive; they both seek to change existing practice, something that correlates with the quick change in this field of medicine. But while the English try to include new medication as much as possible, stretching the definition of cost-effectiveness in the process, the Dutch group seeks progress in reversing towards more classical diabetic treatment regimes. For the two

back pain guidelines the relationship is murkier still, neither guideline group has access to an abundance of evidence to work with, so both end up largely replicating existing practice.

The existing guidelines in the relevant medical field, in combination with existing clinical practice does plausibly exert influence on the development of guidelines because it demarcates a starting position, the direction of this influence is, however, highly contextual. Conceptually this could be seen as only part of the wider frame in which guidelines groups evaluate evidence, in this sense these frames are necessarily multi-layered.

*Conceptions of Disease*

Lying behind ideals of clinical practice sit the ideas and conceptions of disease by guideline makers. Guidelines will differ if guideline groups differ in their conception of illness or their conceptions of (the model) patient. The guidelines on lower back pain show that there is a slight difference in the way the different guideline groups conceptualise lower back pain as a condition. The English group offer a more mechanical model were a potential (if unknown) physical cause of pain predominates. The Dutch put the stress on the physiological treatment of lower back pain, and emphasise the need to deal with chronic pain in a holistic way. As a result, the recommendations of the guidelines are underpinned

by different ideas, and the consensus around those ideas, held by members of the drafting groups.

Another example of this is provided by the guidelines in diabetes. The different diabetes guideline groups do not differ in their conception of the disease, but they do differ in their conceptions of the (model of) patients with diabetes. The English guideline group driven by the need to reach consensus in the difficult area of glucose lowering drugs does this successfully through defining the diabetes patient as consisting of a myriad of sub-groups. This allows all parties to come together behind a decision-tree that is full of carefully crafted exceptions. The Dutch group sees diabetes patients as a uniform group, when it comes to prescribing glucose lowering drugs. They see no need to differentiate between those of different sex, ethnicity or weight and believe that the diabetes population as a whole would benefit from a highly standardised approach, where exceptions (although they do occur) are rare. Different ideas behind the consensus making in the guideline groups resulted in real differences in the recommendations of the guidelines.

These different disease conceptions are one of the factors that can bring guideline groups to change recommendations even if the medical studies they base those on stay the same, the fuel a constant reinterpretation of scientific knowledge in the construction of guidelines, as from one guideline to another the underlying ideas of disease might shift.

*Perceptions of Change in Evidence*

Perceptions of evidence and the role this plays in the consensus reached by the guideline groups is an important function in the social construction of guidelines. There are two possibilities for guideline groups, they can either recommend a change or reaffirm existing practice. In this case, members of the guideline groups reaching consensus chose the status quo more often than they did not. This was because of a perception that the evidence underpinning their recommendation had not substantially changed. Evidence that is 'stable' is a prerequisite for the guideline group to endorse recommendations that reflect the status quo. This quality of stability or change of evidence is not related to the amount of studies being published or any other numerical value. Rather it is related to the perception by the guideline group of whether there is a *qualitative* change in the evidence. This change has to be substantial, and for this to be the case a threshold has to be reached before a guideline group concludes the result of new studies warrants a change in recommendation. This *threshold* is dependent on the particular guideline group, and the medical field the guideline is part of.

The English lower back pain guideline showed an example of a situation where a guideline group is satisfied that a change in evidence warrants a change in recommendation. They recommended acupuncture based on a RCT conducted in England,

which showed a beneficial effect for patients.[6] Here the threshold is quite low, one small RCT is enough to overturn an earlier recommendation not to recommend acupuncture. This contrasts with the decision of the Dutch diabetes guideline group not to recommend the new glucose lowering diabetic drugs, mainly because of lack of long-term data and safety concerns. In their case we can point to a host of new studies on these drugs, none of which were sufficient for the group to classify this as a change of evidence that required a change in recommendations.

What is important here, is that it not just evidence that drives choices it is the perception of qualitative change which forms the pivotal reason behind changing recommendations. This thesis suggests that these perceptions in choices are sometimes more important than the evidence that underpins them. This suggests that a numerical view of evidence (or the picture of an evidence-base as a collection of studies) does not capture fully how guideline group members work.

---

6 Kate J Thomas and others, *Longer Term Clinical and Economic Benefits of Offering Acupuncture Care to Patients with Chronic Low Back Pain* (Gray 2005) <http://www.carolinashealthcare.org/documents/NorthEastInternalIntegrative/research-library/BackAndNeckPain/LongerTermClinicalEconomicBenefit.pdf> accessed 8 October 2014; P Savigny and others, 'Low Back Pain: Early Management of Persistent Non-Specific Low Back Pain' (London: National Collaborating Centre for Primary Care and Royal College of General Practitioners 2009) 156–158 <http://www.nice.org.uk/guidance/cg88/evidence/cg88-low-back-pain-full-guideline3> accessed 13 August 2015.

These findings are broadly in line with those of Pagliari[7] and Hopthrow[8] and Berg[9], in that participants of groups are not all equal, some contributing more than others. Hopthrows processes of information sharing, systematic processing, group formation and process loss also map on the guideline drafting processes. Those insights are not the most relevant findings to come out of this study as this earlier literature fails to fully understand the intimate connection between pre-existing knowledge the participants have before entering the group, and how these understanding filter into the process that occurs within the group. This study would argue that these pre-existing notions, or frames are vital. In this way differentiating from approaches that adopt a neat internal and external divide between the guideline group and 'the outside'.

*The Construction of Frames*

Moreira, in his discussion of repertoires of evaluation, places the focus on identifying several different repertoires that were all present in the deliberations of a guideline group.[10] This study shows that these 'repertoires' are not separate applications of knowledge as

---

7 C Pagliari, J Grimshaw and M Eccles, 'The Potential Influence of Small Group Processes on Guideline Development' (2001) 7 Journal of evaluation in clinical practice 165.

8 Tim Hopthrow, Gene Feder and Susan Michie, 'The Role of Group Decision Making Processes in the Creation of Clinical Guidelines' (2011) 23 International Review of Psychiatry 358.

9 Marc Berg, Ruud ter Meulen and Masja Van den Burg, 'Guidelines for Appropriate Care: The Importance of Empirical Normative Analysis' (2001) 9 Health Care Analysis 77.

10 Tiago Moreira, 'Diversity in Clinical Guidelines: The Role of Repertoires of Evaluation' (2005) 60 Social Science & Medicine 1975.

social action, but it is a mix of several of these 'repertoires' or frames that colour the appraisal of evidence and structure the reaching of consensus in a guideline group. Moreira suggest that further research is needed into how these repertoires are 'blurred, (…) coordinated and crystallised in the CPG [clinical practice guideline] document.'[11] This study suggests that we can understand this process by restating it as one of an interaction between a contextual mental frame of reference with medical knowledge, called evidence, leading to the pragmatic construction of 'facts' that form the basis for recommendations. How these frames are constructed can be illustrated by the following examples.

The analysis of the Dutch diabetes guideline shows how the participants put two principles up front: long-term effectiveness and safety. They adopted this frame *before* considering any of the evidence. Their outlook was informed by the culture of the institution they were a part of, the composition of the group, and negative experiences with Rosiglitazone, one of the diabetes drugs that had to be taken off the market. They consistently applied this frame to the evidence they reviewed, which made them resistant to change (in evidence) that did not accord with those principles. These contingent factors are vital to grasping the specific interpretation of medical knowledge a guideline group employs, which varies with each set of individuals.

---

11 ibid 1984.

Consensus can therefore be a reason for differences between guidelines as well as steering convergence in guidelines. In conclusion, there are several elements that influence how consensus is reached by guideline makers – and these elements have limited but specific impacts on the outcomes of the text of the guidelines. More importantly, the way in which reaching consensus operates depends on the multi-layered framing employed by the guideline group members. This frame depends on the institutional context in which the guideline is drafted, the composition of the guideline group, and the position of the guideline in the wider healthcare system. However, these frames go beyond the social and consensus processes covered in this section and are also shaped by institutions and cultures of institutional practice.

## 6.3 Institutions: Cultures of Institutional Practice

Analysing the drafting of guidelines cannot be complete without considering the institutions that are responsible for publishing and endorsing the guidelines that guideline groups draft. The institutional environment of guideline production is important, but is not addressed in the literature in this particular field. This study provides some insight into the myriad of ways guideline making institutions shape guideline groups and by extension inform the guidelines they produce. Institutional factors provide part of the reason why guidelines differ, as the organisational set-up of guidelines influences their drafting. However, the influence of institutions on guidelines goes further than this: it affects the role conflicts of interests have to play, it shapes the framing of the evidence through the

perception of a specific cultural history by guideline makers, it influences the composition of the guideline group, and it imparts specific directions on what guidelines are for and so what their proper role is in the relevant healthcare systems. These institutional influences inform the whole process of developing guidelines, and they touch upon many relevant aspects of guideline drafting. These aspects will be further discussed in the following sections, before drawing out conclusions about how institutional factors influence the drafting of guidelines. This thesis will suggest that organizations form cultures of institutional practice which directly inform medical guidelines.

Institutions shape part of the framing of evidence by guideline group members. This framing influences the purpose and aims of the guidelines of the guideline making process. How guideline group members view evidence is partly due to how they view this institutional context and how that operates during the process of guideline drafting. This is partly informed by their understanding of earlier guidelines that have been produced within the organisation. A sense of 'how we do things here' prevails throughout guideline making. A difficulty in ascertaining this influence is that the text of a guideline does not necessarily reflect this kind of institutional influence. Therefore a reliance on interviews is necessary to reconstruct potential mechanisms of influence.

*Audience*

For example, the NHG guidelines are made by GPs for use by GPs. This starting point looks obvious but has implications for the framing of evidence in guideline groups. Both the lower back pain and diabetes guideline group stress the clinical relevance for GPs of their recommendations. This is driven by the mandate that they are given by the NHG to deliver answers to practical problems that have been put forward by other GPs. The lower back pain group prioritises the clinical relevance of studies over their statistical information.[12] While the Dutch diabetes group puzzles endlessly on the formulation of the treatment algorithm for the differentiation of blood sugar targets, so as to maximise the change GPs will understand and follow the visual information.[13] These observations provide some support for the fact that the overriding cultural influence on guideline groups that are part of the NHG is the production of practical knowledge for GPs. That guidelines are evidence-based is important, and efforts are made to further formalise the searching and analysing of evidence, but this is less important than *useable knowledge*. This is achievable for the Dutch guideline groups always include several practising GPs who have intimate knowledge about the clinical environment the guideline is going to be used in. In

12 As detailed in their discussion on educational materials.

13 Guideline group members claim they had 4 or 5 meetings just about this one issue.

the Dutch case, institutional pressures focus on the practicality and usability of guidelines rather than other motivations.

The NICE guidelines are part of an institution that has a different outlook. Guidelines are produced for the whole NHS, not one specific profession. NICE also has a duty to consider the cost-effectiveness of treatments advised in guidelines. This leads to an approach to guideline making that stresses the objectivity and robustness of formalised procedural measures, to help achieve consensus and weigh different interests. Evidence-based medicine is important, through the use of GRADE and other formalised measures for establishing the technical robustness of evidence as well as economic studies, NICE tries to create conditions for a multidisciplinary group of medical professionals to reach consensus on recommendations. The robustness of the procedures that NICE employs are the main weapons it has to secure the legitimacy of the guidelines it produces. This means that clinical relevance of the medical studies underpinning the guideline is relevant, but from an institutional perspective the guideline group is incentivised to accord more weight to statistical results of medical studies. These differences in institutional context drive – albeit indirectly – the evidence and the recommendations in the various guidelines.

*Composition of the group*

The composition of guideline groups is a further element where a specific institutional context matters. Institutional rules govern the composition of guideline groups and choice

of participants in the group. While at the NHG selection of guideline group members is an informal process: with a preference for GPs with a doctorate in combination with GPs with an academic background, it always results in a guideline group where the majority of the members are GPs. As a result, the members will share a certain professional outlook, and are orientated towards the interests of Dutch GPs. In forming the diabetes guideline group, certain guideline group members could also influence the choice of other guideline group members by 'putting forward names' to the NHG. The process at NICE seems to have changed over the years. While the members of the lower back pain guideline are drawn from the various professional associations that treat lower back pain, the diabetes guideline selection process is more of a mix of people that are tied to professional associations and those that are not. The NICE groups feature a greater mix of professionals from various different specialisms, compared to the Dutch groups. Participants in these guideline groups therefore have to strive to reach recommendations that are supported within and across several professional specialisms. It could be said institutional bodies directly dictate the composition of guideline groups, and through this influence direct the resulting negotiations which in turn produces particular recommendations. These factors are all unrelated to considerations of evidence: but can be said to be critically important in practice. Thus suggesting that some of the lead factors influencing the drafting of guidelines are not directly tied to evidence and are, in fact, determined by institutional configurations and specific socio-institutional cultural contexts.

*Conflicts of interest*

One additional way institutional power is demonstrated is through the institutional control exercised over the reporting and policing of conflicts of interest in guideline drafting. The reporting of conflicts of interest is mandated by the manuals produced by guideline making organisations such as the NHG and NICE.[14] Both organisations require the reporting of financial and non-financial conflicts of interest. There are some slight differences in procedure between the two institutions: for example, the time that has to elapse after which a conflict of interest is long enough in the past to be deemed irrelevant. More relevant for this study, however, is the institutional culture that shapes how conflicts of interest are dealt with in practice.

The NHG has adopted a strict line, and is very focal about wanting to avoid conflicts of interest as much as possible. The fact that the diabetes guideline group consisted of several people with conflicts of interest showed that the NHG is not always successful in this aim. Nonetheless members of both guideline groups interviewed, who conducted their work between 2003-2013, all stated that the NHG takes a strict line on conflicts of interest.

---

14 NHG,'Procedureboek' <https://www.nhg.org/sites/default/files/content/nhg_org/uploads/procedureboek_versie_090810.pdf> accessed 17 November 2014; 'Developing NICE Guidelines: The Manual' (n 4).

So regardless of the reality, there is at least a strong perception, among all guideline group members, that conflicts of interest should be avoided.

NICE policy on conflicts of interest has grown stricter over the years. While members of the 2008 guideline group on lower back pain where drawn from the relevant professional associations, in an informal way without a public vacancy, for recent guidelines all position are formally advertised. The interpretation of conflict of interest has also become tighter. The guideline group that produced the 2016 guideline on diabetes has among them far fewer people with a conflict of interest than the 2008 and 2009 guidelines.[15] The main difference with the NHG is that the assessment of conflicts of interest relies on a formal assessment of the ties the potential guideline group members report, in contrast to a practice at the NHG where in selecting guideline group members social networks play a bigger role, and an assessment of how independent someone is more depended on social relations.

The results of the case studies suggest that conflicts of interest are policed by institutional structures, but the true impact of this on the text of the guidelines remains unclear and difficult to trace. This study cautiously suggests that the NHG was more successful in promoting a culture of guideline drafting that was critical of the

---

15 'Type 2 Diabetes in Adults: Management | Guidance and Guidelines | NICE' s 4 <https://www.nice.org.uk/guidance/ng28> accessed 17 October 2016.

pharmaceutical industry, and its potential influence on guidelines. Once again this demonstrates clearly that institutional choices shape guideline drafting, rather than choices being made in a vacuum and solely based on evidence.

*Institutional mission*

The final element that stands out if we analyse the results of the case studies is the perceptions guideline group members have of the purpose of institution. This relates to how the guideline drafters see the institution they are working for, its position in both healthcare system and wider society, and the ultimate purpose of making a guideline in the first place. This is reflected most clearly in the existing publications and products produced by the Dutch and English institutions, and other institutional documents that reflect the mission of the organisation.

The NHG guidelines are a part of a larger range of 'products' produced by the NHG aimed at Dutch GPs and other medical professionals working with GPs in primary care. The NHG produces protocols, handbooks and factsheets all aimed at improving several aspects of care by GPs. Guidelines are only one of the forms the NHG uses to communicate with GPs, and it is clear that guidelines are seen as an educational and advisory tool for use by GPs and practice nurses and other health professionals. A further, but more longer term, effect of the guidelines has been that the standing of the profession, being a GP, has considerably increased over the years. Guidelines are now an integral part of training new

GPs. Guideline group members are aware of these efforts, all of them have worked with the NHG earlier, or attended on or more of its activities. The diabetes guidelines produced by the NHG shows the influence of this institutional background. It deals explicitly with the question whether NHG guidelines should dictate that organisation of diabetes care, or whether local guidelines should be followed, and recommends to follow the NHG organisational model of diabetes care delivery.[16] This shows how NHG guidelines are integrated in a package that influences the organisation of primary care in the Netherlands. In contrast, in England a focus on advancing the position of a specific medical profession is understandably absent. The NICE mandate requires guidelines that include all professions and so does not result in strong institutional steering of specific products to further the interests of one profession. The English guidelines present a mixed bag of various documents influenced by the needs of the healthcare system as whole, and represent a range of institutional priorities.

*Institutional Framing*

Important institutional differences emerge when considering the drafting of a guideline as a process. These institutional impacts do not only occur *within* a guideline group, but are also intimately connected to institutional factors *outside* the group. For example, the Dutch

16 GEHM Rutten and others, 'NHG-Standaard Diabetes Mellitus Type 2 (Derde Herziening)' (2013) 56 Huisarts Wet 512, Footnote 101.

guidelines were produced with a view to strengthening the GPs as a profession, or at the very least hold their current position in the medical field, and as an educational tool for GPs in training. Whereas NICE guidelines are more closely aligned to government goals of healthcare rationing and approach the NHS as a whole: these guidelines aim to improve care within explicit constraint of cost-effectiveness. This institutional context results in markedly different guidelines notwithstanding the fact that they consider the same medical condition.

This description points to the importance that institutional scripts have in driving participants in guideline groups to take certain decisions. These findings fit very well with Meyer's concept of 'agentic actors', our guideline group makers indeed describe themselves as driving by ideals of medicine/clinical practice or other cultural ideas or goals, while also having a self-interest. As some of the interviews also provide evidence for the motives of participants to take part, some put themselves forward out of curiosity, others for career advancement while there are those that like the status associated with drafting a guideline. These references are, however, more muted and subtle, while outwardly towards the institution they are a part of they only communicate in the language of evidence-based improvement of quality of care.

In conclusion, this thesis suggests that institutional factors specifically and concretely influence the drafting of guidelines. Institutions achieve this through promulgating specific cultures of institutional practice which directly inform medical

guidelines. This opens up a fresh perspective to consider professional self-regulation, such as guidelines, in a way that heeds the particular institutional embeddedness of the processes of standard setting and information gathering. On the other hand, this thesis confirms existing scholarship on sociological institutionalism, and its relevance for understanding the actions of guideline drafters.

## 6.4 External Factors: Limited Influence

This study found limited impacts of external factors on the final guideline texts. This omission can partly be explained by the fact that after guidelines are published commentary or criticism hardly has an effect. And the comments that are made by commentators before a guideline is published are not always public, as was the case in the Netherlands. Nonetheless discussing external factors is valuable because it reveals some of the path-dependencies inherent in guideline development, and provides contrasting empirical examples of its occurrence.

Guidelines are written to impact the delivery of care, and to help improve the standard of care. Crucially they can also affect the organisation of care and the autonomy of healthcare professionals. To understand the mechanisms behind choices in the guideline drafting process looking at how a guideline is received after it is published is relevant.

The results of the case studies provide an interesting contrasting picture of guidelines that run into criticism and those that hardly received any response. By analysing these results in detail it will become clear that path-dependency plays an important role in understanding the responses to guidelines, and the choices made in guideline drafting. Path-dependency is defined here in a general way: 'most generally, path dependence means that where we go next depends not only on where we are now, but also upon where we have been.'[17] The responses to guidelines seem to have been driven by a resistance against change, when historical choices are (dramatically) changed by dropping or adding treatment for a condition or disease. This resistance then sometimes leads to the reinstatement of historical choices, despite the earlier contention that those choices are *inferior*.

A few examples for the case studies will help to explain this argument. Several of the guideline analysed in this study showed no external responses of disapproval of its contents by professional associations. The Dutch back pain and diabetes guideline were met with silence and support respectively. The English guidelines on diabetes were supported, although the pharmaceutical industry asked for a wider role for some

[17] Stan J Liebowitz and Stephen E Margolis, 'Path Dependence' [1999] Encyclopedia of law and economics 981 <http://ecsocman.hse.ru/data/018/784/1216/0770book.pdf> accessed 2 November 2016.

medication[18], medical professionals were supportive. The exception was formed by the English guidelines on back pain. The resistance against dropping injections in the back as a treatment were strong. The guideline group came under pressure to change the guideline, but resisted this pressure and the guideline remained unchanged on this point. This shows how when a guideline group wants to deviate from a path of (in this case) treatments chosen in the past this will cause resistance that sometimes leads to a return of the status quo.

The recent 2016 guideline on diabetes by NICE[19] that provided the update to the 2008 and 2009 guidelines analysed in this study, provides a further interesting example of resistance to change, and the wish to preserve historical choices. This example makes clear that historical choices and their results *make it harder to change those choices in the future*. The 2016 draft guideline favoured Repaglinide, a decision that caused uproar.[20] This resistance against changing the second step in blood sugar lowering therapy from Sulfonylurea to Repaglinide was predicated (mainly) on the fact that it was prescribed in

---

18 As could be gleaned from the consultation documents. 'Diabetes-Type-2-Update-Consultation-Table2 Cg-66.Pdf' <https://www.nice.org.uk/guidance/cg66/documents/diabetes-type-2-update-consultation-table2> accessed 14 May 2015.

19 'Type 2 Diabetes in Adults: Management | Guidance and Guidelines | NICE' (n 15).

20 Anthony H Barnett, 'NICE Draft Type 2 Diabetes Guidelines: A Cause for Concern' [2015] The Lancet Diabetes & Endocrinology <http://linkinghub.elsevier.com/retrieve/pii/S2213858715000534> accessed 5 May 2015; Caroline Price, 'NICE Risks Making Itself a "Laughing Stock" over Guidance on Metformin Alternatives, Say Experts' (*Pulse Today*) <http://www.pulsetoday.co.uk/clinical/diabetes/nice-risks-making-itself-a-laughing-stock-over-guidance-on-metformin-alternatives-say-experts/20009139.article> accessed 5 May 2015.

less than 1% of cases of diabetes[21], coupled with the fact that it was seen as clinically inferior. Without giving any judgement on the evidence-base the guideline group used to make this recommendation, the fact that the drug was unpopular made it unfeasible that it could become a mainstream medication in treating diabetes. The historic choices not to favour Repaglinide made it very unpractical to change this choice later on. The guideline group completely reversed their earlier choice for Repaglinide and relegated the drug to a small footnote in the algorithm they eventually agreed upon.

Repaglinide was seen as an inferior drug, so the cost-effectiveness arguments that led to its adoption as preferred medication could not be sustained in the light of its perceived clinical inferiority. The 2014 Osteoarthritis guideline provides a further good example of this process. In the draft of this guideline the guideline group recommends not to offer paracetamol routinely in the treatment of osteoarthritis based on a combination of safety issues and the fact that it was less effective than no treatment in the economic

---

[21] '6.1 Drugs Used in Diabetes : British National Formulary' <https://www.evidence.nhs.uk/formulary/bnf/current/6-endocrine-system/61-drugs-used-in-diabetes> accessed 21 July 2016.

assessment.[22] This – as with the diabetes guideline - led to a storm of critique[23] and the decision was reversed.[24]

The main reason was that the alternatives to paracetamol the Non-Steroidal Anti-Inflammatory Drugs and opioids also had side-effects, so patients switching to these medications would be even worse off than when in a situation that they would use paracetamol. Here the optimal choice could not be made: removing the choice of paracetamol was not a path that could be taken, because providing no treatment where historically treatment with paracetamol was provided was considered unacceptable. Coupled with the fact that paracetamol was (and is) a medication that is available over the counter, so that restrictions on its prescription might have consequences beyond this guideline, made this choice in the given circumstances impossible. External forces can have an influence when the choices suggested by the guideline makers are so radical that sufficiently large number of people might reject the choice. Or the choices are perceived to have a future substantial impact, and so guideline makers self-limit their choices. If a substantial number of people refused to accept the recommendation of a guideline, this

---

22 NICE, 'Osteoarthritis (Update): Draft Guideline' (2013) 410–416 <https://www.nice.org.uk/guidance/CG177/documents/osteoarthritis-update-draft-guideline2> accessed 2 November 2016.

23 NICE, 'Osteoarthritis (Update): Consultation Comments Table' (2013) <https://www.nice.org.uk/guidance/CG177/documents/osteoarthritis-update-consultation-comments-table2> accessed 2 November 2016.

24 'Osteoarthritis | About-This-Guideline | Guidance and Guidelines | NICE' <http://www.nice.org.uk/guidance/cg177/chapter/about-this-guideline> accessed 9 October 2014.

would lead to a loss of legitimacy, and so influences the guideline drafters during the and after the process of writing the guideline. Therefore external forces are a relevant factor, but only in limited circumstances where guidelines push the boat out.

The various different experiences with critique on guidelines suggest that when guideline makers want to deviate from the status quo they can expect in some cases to be resisted. This is something guideline group members are well aware of, and affects the way they draft guidelines. In this way I would propose to extend the repertoire of politics as offered by Moreira[25] to not only include the government and pharmaceutical industry and further include in this category the politics of the medical professionals for whose benefit the guideline is written. Not only do medical professionals resist change if this affects the organisation of care, such as injections for the back, but they are equally poised to reject a limiting of choice in treatments without compelling justifications. Guideline group members are of course part of these professions, and are aware of what the standard of care is in the medical field on question. They therefore frame their decisions with reference to the expected response by their fellow medical professionals.

Differences between guidelines could therefore partly be explained by external factors, however, this has not been borne out by this study. In these case studies, none of the guidelines were changed as result of 'outside feedback.' This research has demonstrated

---

[25] Moreira (n 10).

that guideline drafters did react to the *potential* of feedback, and therefore in most cases avoided it. Guidelines are shaped by the specific configuration of medical professionals in the relevant medical field, their perceptions of external factors and institutional structures, and how vocal they are in defending their (clinical) positions.

## 6.5 Conflicts of Interest: The Difficulties of Attribution

The role of conflicts of interest in guidelines is a nuanced one. The presence of members of a guideline group that have a conflict of interest does not necessarily mean that this has influenced the drafting of the guideline. Nonetheless this study has set to establish to what extent a conflict of interest has had a discernible influence on the development of guidelines. In some instances the pharmaceutical industry has tried to influence guideline makers, but it remains unclear whether this ultimately affected to final guideline text.

There exists a sharp divide between guidelines on lower back pain and guidelines on diabetes. While conflicts of interest seem largely absent in the case of lower back pain, in diabetes they do play a role. Furthermore, there are important differences in how conflicts of interest have influenced guideline groups, and it remains unclear whether a causal relationship between these conflicts and the recommendations in a guideline can be made.

The analysis of the results of the case studies on the back-pain guidelines shows that some members of the guideline groups report conflicts of interest. There is no evidence that these have had any influence on the recommendations in the guideline. Some evidence cited in the guideline is co-authored by members of the English guideline group, while in the Dutch group studies are cited that are authored by a member of the group. This might lead to a 'cognitive bias' where studies produced by one of the members of the group is given a higher weight then might be warranted on the basis of the study itself. There is no evidence that this bias has occurred, there is evidence that guideline group favour studies conducted in their own country over others, but this has not resulted in recommendations that differ between the lower back pain guidelines.

The picture for the diabetes guidelines is rather different. In the Dutch guideline group 3 of the 7 members of the group reported conflicts of interest. The central study cited in support of recommendations on blood glucose lowering medication was a study co-written by 5 of the 7 members of the guideline group. The ingredients for a group that suffered from a cognitive bias as well as conflicts of interest were therefore present. The group did develop a bias but this was more plausibly an anti-industry bias. The study conducted by the 5 of the group members did receive significant weight, this led to a strong recommendation for the 'old' diabetic drugs and a very limited position for the drugs that came on the market more recently. The Dutch guideline on diabetes therefore is an

international outlier and one of the most restrictive when it comes to the position of the newer diabetic drugs.

Equally interesting and revealing is what happened during the process of forming the guideline group that was going to write the Dutch diabetes guideline. One of the guideline group members discussed in great detail how a member of a pharmaceutical company he had contact with was already aware that he was going to be a member of the guideline group just two days after he was approached, before this information was public. He went on to describe how this company provided him with 'relevant trails' to make it 'easier' for him to do his work, and as he explained to make sure the research on their drug was featured in the evidence selected for the guideline. This experience led him to conclude that he should further distance himself from the pharmaceutical industry. It shows that concerns about conflict of interest are not only theoretical but attempts at influencing guideline makers take place in practice.

The English guideline group has several members that report a conflict of interest. There is a very limited overlap between the studies cited in the guideline and any members of the guideline, so the presence of a 'cognitive bias' seems unlikely. There is, however, some evidence that points to the guideline group weighing aspects of the new diabetic drugs, the potential for weight loss and avoidance of hypoglycaemia, as important, leading them to favour these drugs and give them a place in the treatment of diabetes. This does not mean that this was caused by a conflict of interest. It is just as likely that this was the

result of a genuine consensus that these drugs were clinically superior in some specific clinical circumstances and therefore were given a place in the treatment of diabetes.

This suggests that if we limit ourselves to the text of the guidelines conflicts of interest in these case studies are no more than a potential factor. It explains some of the differences in diabetes guidelines, specifically the different choices regarding blood glucose lowering medication in the diabetes guideline. In the other guidelines, and for the lower back pain case studies, conflicts of interest had a very limited role. This might be due to the difficulties in establishing a causal relationship between the disclosure of a conflict of interest and the final recommendations of a guideline.

Using a broader lens and looking at the development of guidelines as a whole, it is clear from the example of the Dutch diabetes guideline that attempts at influencing guideline makers are made, even if they are not always successful.

## 6.6 Cost-effectiveness: Considerable Tensions between Ideals and Practice

Drafting guidelines often involves evaluating evidence on cost. This evidence takes the form of medical studies that provide knowledge about the cost-effectiveness of treatments. A further category of evidence is economic evidence. This evidence often takes the form of cost-benefit or cost-effectiveness analysis where the cost of medication or treatment is

contrasted with their effectiveness in improving the health of patients. The way cost-effectiveness played a role in guidelines varied in the Netherlands and England and equally varied between guidelines. Cost-effectiveness consideration did not form a formal part of the drafting of guidelines in the Netherlands. In contrast cost-effectiveness modelling was a mandatory aspect in the drafting of the English guidelines.

This study explicitly considers the relationship between economic evidence considered by guideline groups and other evidence considerations. Analysing the role of economic evidence further reveals the consequences this might have for differences between guidelines, given the large differences in the Netherlands and England in dealing with issues of cost-effectiveness. First analysing the Dutch guidelines in detail will show that an informal role for cost issues does not necessarily lead to decisions that are 'costly.' A review of the English guidelines in contrast to the Dutch experiences shows some potential unintended consequences of the formulation of thresholds in dealing with costs. But even between guidelines in one country differences occur in how cost evidence is handled, and how it shapes the development of a guideline.

When analysing the results of the case study on the NHG lower back pain guideline the issue of cost was not mentioned at all in relation to this guideline. The case study on this guideline therefore does not include an assessment of costs or cost-effectiveness. As

noted earlier the absence of cost considerations is due to the fact that the decisions on which treatments are part of the package of insured care lies with the Zorginstituut.

The Dutch diabetes guideline shows a different picture. The institutional situation has not changed but in this guideline reference is made to costs. Not by using a formal model but in a simple qualitative sense: through labelling drugs as cheap or expensive. The interviews with members of the guideline group show that they feel a general obligation to pay heed to the cost of the healthcare system as a whole of the choices they make in the guideline.[26] I would suggest, however, that cost played a larger role than the interviews would suggest, but in an indirect fashion. The Dutch guideline group justified their choice for the 'old' diabetic drugs with a reference to safety data and lack of long term studies into the efficacy of the drugs. These considerations would plausibly be different if these newer drugs were not 30-50 times more expensive than the old ones. In this sense 'the responsibility for the cost of the healthcare system' did figure in their deliberations, as the guideline reflects in the judgement that these drugs are 'expensive.'

The English guidelines provide a different picture, both guidelines are full of references to economic studies, contain modelling of treatment choices based on a formal assessment of healthcare states. A specific technical jargon is employed which includes QALY's, quality-adjusted life years, ICER, incremental cost-effectiveness ratio, and

---

[26] Paragraph 5.14.

EuroQol health-related utility values, which is a standardised measuring tool used in surveys to measure how patients score their future health state when confronted with symptoms of a disease.[27] The aim of this modelling is to quantify the cost to the NHS of providing a treatment that will give an average patient the average of 1 year of health.[28] NICE has set a threshold value, which reflects the willingness to pay of the NHS to provide this 'average' patient with 'health', this threshold is £20,000-£30,000 per year of health. Knowing this background, we can evaluate the judgement made in the two NICE guidelines studied. The back-pain guideline includes economic studies in their evaluation of evidence, whenever one is available.[29] All these studies show the costs of the treatments recommended is below the threshold set by NICE. Only for one treatment, spinal fusion the treatment cost more than 20,000 pounds.[30] The group responded to this result by limiting access to the treatment to only those who followed all other steps/treatments in the guideline within a year, they could be referred for an *opinion* on getting spinal fusion. As members of the guideline group explained in the interviews, they expected the group off people who would complete the guideline within 1 year to be so small that the cost of the treatment would not be an issue. Because the treatments for lower back pain were all

---

27 An example is the EuroQol for diabetes: Adrian Bagust and Sophie Beale, 'Modelling EuroQol Health-Related Utility Values for Diabetic Complications from CODE-2 Data' (2005) 14 Health Economics 217.

28 Health is 'quality-adjusted' to reflect the fact that the person has a disease, it is an expression of 'relative health' versus the condition without the treatment.

29 Thomas and others (n 6). Is one example.

30 ` Savigny and others (n 6) Appendix E.

relatively cheap these considerations did not force the guideline group to change their recommendations. The inclusion of economic evidence in the guideline did however make more explicit what the additional costs would be of providing a course of acupuncture on the NHS, something which was used in the external response to the guideline after it was published.

They were asked to draft a guideline for anyone with back pain between 6 weeks and one year.[31] This excluded from their consideration people with chronic pain (defined as back pain for longer than a year). This was thus a guideline that was constructed for a very particular clinical condition, which allowed the guideline group to push difficult questions of what the position of spinal surgery should be in *chronic back pain* to the background. This experience underscores the need to see the drafting of a guideline within a specific institutional context and a particular medical field: the back-pain guideline was the result of specific institutional choices by NICE to define the scope in a (limited) way. This example shows how cost-effectiveness considerations can be neutralised in this case through declaring a specific treatment outside of the scope of the guideline and allowing the treatment to be recommended through other mechanisms.

The experience of including cost-effectiveness in guidelines is not uniform among all NICE guidelines. The diabetes guidelines were influenced by economic evidence to a

---

31 ibid Appendix A.

greater extent than the back-pain guidelines. This can be partly explained by the nature of the medical condition: treating diabetes is complex. The guideline group therefore had to make more complex decisions, and faced various trade-offs in deciding about aspects of diabetes care. The group was helped by the fact that there were more economic studies that provided evidence of the cost-effectiveness of various diabetes treatments, as well as a model based on the UK Prospective Diabetes Study (UKPDS) to help calculate the cost and effectiveness of several blood sugar lowering drugs.

In this case study the analysis was focussed on the decision the English group made on the use of Exenatide, this was not the only drug where economic evidence played a role but it was chosen because it provides a clear example of the mechanisms behind the conflict between economic and clinical knowledge in drafting guidelines. Exenatide was shown to be not cost-effective in the original cost-effectiveness model.[32] The guideline group responded to this result by attacking the assumptions underlying the modelling. They argued that the model did not sufficiently allow for the positive effects of the drug (reduction in hypoglycaemia and weight loss). In the revised modelling one instance was found where the drug was cost-effective. The resulting compromise tried to accommodate the economic evidence (which suggested that the drug would only be cost-effective in

---

[32] National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London, *Type 2 Diabetes National Clinical Guideline for Management in Primary and Secondary Care (Update)* (Royal College of Physicians 2008) <http://www.ncbi.nlm.nih.gov/books/NBK53885/> accessed 14 April 2015.

limited circumstances) with the clinical wish to be able to employ the drug to increase patient's quality of life. This resulted in a carefully phrased compromise where the drug could only be prescribed to people we were obese and responded positively to the drug with a decrease in blood sugar level as well as weight loss.[33]

This aspect of guideline drafting neatly encapsulates several elements that make NICE guidelines on these aspects fundamentally different when compared to NHG guidelines, while at the same time revealing the unintended consequences of trying to square healthcare rationing with producing clinical guidelines. In line with Goldberg this study suggests the use of economic modelling is the ideal location to see how guidelines are places where political questions are transformed into technical ones. [34] The operation of this mechanism can be shown through analysis of the underlying motivations that were given in the debate surrounding Exenatide in the guideline. For the English group, the conflict between the assumptions underlying the (original) economic modelling that led to the conclusion that Exenatide was not cost-effective clashed with the assumptions the guideline group members held about what they considered to be good care. The guideline group members operationalised diabetes care as a continues process, where Exenatide would be an option to delay or even prevent the starting of insulin therapy. They thought

---

33 ibid 124.

34 Maya J Goldenberg, 'Iconoclast or Creed?: Objectivism, Pragmatism, and the Hierarchy of Evidence' (2009) 52 Perspectives in Biology and Medicine 168.

of costs in a holistic way, all the cost of insulin therapy, including assistance with injection and the monitoring of blood glucose levels, needed to be considered. The result of these deliberations was that the clinical availability of Exenatide as a treatment option was of greater weight then considerations of cost.

The debate surrounding this issue appears to be a technical debate about the definition of costs and the calculation of benefits, but it is also a debate about *values*. By defining good care implicitly as a continuous process, as part of a management of a chronic and lifelong condition, providing drugs that are not strictly cost-effective is a perfectly rational choice and is in conformity with values about patient care. Due to the inclusion of economic evidence, the English diabetes group is forced to reveal some of these values more clearly, than in a group who did not include this evidence in their debates.

The use of economic has a potential unintended consequence, however, as we can glean from the debate surrounding Exenatide. The explicit definition of a threshold invites the guideline group to find a situation where the medication is (just) below the threshold so that this can be seen as meriting the label 'cost-effective.' As we have seen a conception of diabetes care where Exenatide was a valuable drug led to the inclusion of the drug in the treatment algorithm with conditions attached to its use. At the same time this meant that the option not to advise the use of the drug at all, did not appear on the table because the drug was cheaper than the NICE threshold. This led to the paradoxical result that a 'cost-effective' NICE guideline recommends more expensive medication than the Dutch

guideline that does not concern itself with cost (at least formally). The realisation that cost-effectiveness in the sense of not reaching the NICE threshold is not necessarily always the most cost-effective option is an important one. This result should be interpreted cautiously, however, because this one example might not be generalizable to all NICE guidelines and be tied to the particular circumstances in the drafting of the diabetes guideline.

The mixture in ways cost-effectiveness forms part of the drafting of a guideline does not allow for very general conclusions, but some indications of underlying mechanisms can be established. The role of economic evidence forces guideline groups to give explicit consideration to the two different conceptions of guidelines: a guideline as an instrument of healthcare rationing and the guideline as an instrument for the highest possible quality of care. The English diabetes guideline group successfully resolves this tension by redefining the economic evidence in such a way that a compromise can be found that satisfies clinical and economic considerations. In contrast the Dutch guideline on back pain includes recommendations for treatments (physiotherapy) that are not covered under the basic health insurance package in the Netherlands, so the cost need to be borne by patients (through additional insurance or cash payments). So neither the cost to patients nor the cost for the healthcare system are formally considered. In this way the Dutch back pain guideline avoids the conflict described here: it has as its sole purpose to advise on the best quality of care.

Economic evidence, therefore, has the potential to counteract clinical evidence. If drugs are established to be too expensive in relation to their effectiveness they might not be recommended. The case studies provide no evidence that economic evidence forced guideline groups to actively change their recommendations to cheaper drugs.[35] There is evidence, however, that the prescribing of Exenatide was curtailed to accommodate the fact that the economic evidence showed that the drug was not cost effective if prescribed to all patients. Therefore, while cost evidence did not lead to cheaper drugs being recommended, there is some evidence from these case studies that it did lead to more expensive drugs not being the first choice of recommendation. While the effect of economic evidence on differences in guidelines is small, the fact that the English guidelines are perceived as healthcare rationing instruments might influence their implementation.[36]

This study finds that the practice of considering economic evidence does not yet fully live up to the ideal of choosing treatments that are both clinically effective and cost-effective. This is because scientific evidence is never perfect or complete, and even if it was, it would not always be able to satisfy the joint requirements of treatments being both the cheapest and the best option. In the end it is always the human factor of the clinician armed with practical knowledge about the true value of the treatment that can successfully

---

35 The draft 2016 guideline on diabetes however was heavily influenced by cost-effectiveness considerations, as discussed in the controversy surrounding Repaglinide.

36 A full appraisal of these effects falls outside the scope of this thesis.

mediate between these conflicting pressures. This helps us uncover one of the strengths of guidelines as a regulatory tool, in that they provide a clear consultative, iterative, and social dynamic process to thrash out complicated and thorny scientific questions. Especially in the case of cost this is important, because otherwise normative choices would be made arbitrarily by a spreadsheet based on the lowest cost option. This study proposes that it is more fruitful to understand the consideration of cost not as a technical exercise, but as an invitation to consider the relative weight of various values.

Evidence on cost-effectiveness, its use in the medical guideline drafting process, and analysis of that process through socio-legal methods, can provide more clarity on the practices underpinning medical regulation. The way in which doctors and guideline drafters mediate cost pressures while maintaining appropriate clinical choices, helps to prioritise patient care and wellbeing. This blunts some of the neo-liberal drivers in this sector, which are characterised by an oversimplification of medical decisions to be a simple judgment of the lowest cost. This discussion of ideals and political motivations in medical healthcare regulation is a particularly salient and topically debate in current healthcare policy and is deserving of future research.

In conclusion, cost-effectiveness considerations are a very important element of the discussions in English guideline groups. However, they play a limited role in the Dutch guidelines to date because of the current institutional context of the Dutch healthcare regulation system. As evidence on costs are likely to gain prominence in the Dutch system

for future guidelines, this thesis opens up possibilities for future research in this area. Posing questions on the validity and value judgements implicitly made in cost choices early in the policy-process, would enable decision-makers to be self-reflexive about their evidence choices in healthcare guidelines.

# 7 CONCLUSION

This concluding chapter is aimed at addressing what factors and mechanisms explain the differences between the development of evidence-based guidelines. The conclusions drawn from this study will be organised around the themes identified in the sub-questions of this thesis: What role does evidence play in creating differences and similarities in guidelines, and what types of 'evidence' is used? How do institutional framings impact on guideline development? What is the role of group decision-making in the development of guidelines? What is the role of evidence on cost and cost-effectiveness in the formation of guidelines? How do external factors determine choices made in guideline development? And finally, how do these factors and mechanisms impact on guidelines as a form of professional self-regulation?

## 7.1 Evidence

Evidence has been of central concern in this thesis. To study empirically how guideline groups grapple with the fact that guidelines have to be evidence-based has been a key theme in understanding the development of medical practice guidelines.

It turns out guidelines are not based on evidence in a straightforward sense, rather they are the result of a complex negotiation of social, economic, political and normative

mechanisms that both select and shape evidence in various ways. The notion of cognitive framing turns out to be a powerful conceptual tool to describe this black box in detail.

A major conclusion of this thesis is that evidence-based guidelines are not as evidence-based as they claim to be: rather the term evidence-based guideline is used to justify normative and political subjective choices in professional self-regulation. This opens up a new understanding of how scientists construct and produce knowledge and how this is consequently presented as empirical truth. The analysis in this thesis suggests that guidelines are an expression of a temporal closure of debate on the optimal medical choices regarding a specific condition.

This process of selecting evidence and evaluating the methodological soundness of evidence is relatively well understood. This study has shown how guideline groups make different decisions. Some only include evidence of a certain quality (RCTs), others rely on summaries of studies (Cochrane Reviews). One could stop there and consider the fact that guideline groups differ in their process of selecting evidence as proof that this explains the differences in the development of guidelines. This is where existing scholarship is lacking as it omits from consideration how this evidence is then used in a social and constructive process to make recommendations in a guideline. It is better to speak of 'evidence-informed' guidelines, where medical studies can inform the understanding of the likelihood of certain treatments being effective but do not determine conclusively which recommendation should be drawn from that information and therefore which action should

be taken. So even if the process of selecting evidence was the same, and groups would consider exactly the same medical studies, different inferences could still be drawn from the medical knowledge presented to them.

This study has clearly shown how the consideration is evidence is a process where *meaning is attributed* to the medical knowledge contained in medical studies through a social, discursive and iterative discussion. This discussion is fed by a context dependent combination of principles of good care (that treatments should provide maximum of choice, with a minimum of harm to patients), deep knowledge of clinical practice, and political considerations (who gets to do what and when and how). Guideline recommendations have to fit within an existing *frame* of treatment of a given condition, and regardless of the strength of the evidence, consensus within the drafting groups, needs to be reached. It is only through the combination of these various factors that recommendations are reached and that evidence informed guidelines can be said to be based on evidence.

It is through a detailed understanding of the interaction between a clinical cognitive frame that the appraisal of evidence is conducted. Evidence is thus not a form of knowledge that is mechanically applied to a well understood clinical set of circumstances. Rather evidence is socially constructed in a discursive process that is context dependent. This explains why guidelines can be similar notwithstanding large differences in evidence cited in the guideline, and can be different although the evidence used is similar. It is the framing, through a process of debate, which creates recommendations. This does not mean that

guidelines can completely ignore evidence, medical studies do provide certain limitations on the arguments a group can make. However, it is through this argumentative process that evidence is imbued with meaning.

Guidelines exhibit a certain duality, a tension between their text and their drafting. Guidelines are attempts to mobilize practical knowledge to answer clinical questions, to provide *guidance*. Guidelines are part of a continuous dialogue between members of the medical profession(s) involved in the treatment of the medical condition of the guideline. This helps answer the question Moreira poses about guidelines[1]: how the text of a guideline can be so technical while the process of drafting is such an intensely messy and social enterprise? A possible answer to this question lies in the function a guideline performs in the dialogue with other professionals. To be able to be perceived as a convincing and *legitimate* guideline, the subjective elements of its construction must, as much as possible, remain hidden from view: calling a guideline evidence-based provides such a device.

This rhetoric move within the context of guideline drafting has two separate effects. Firstly, it changes the terms under which other 'participants' in the debate on the condition or disease can attack (recommendations in) a guideline. Only if people who criticise the

---

[1] Tiago Moreira, 'Diversity in Clinical Guidelines: The Role of Repertoires of Evaluation' (2005) 60 Social Science & Medicine 1975.

guideline can point to conflicting evidence (in the form of medical studies) will the criticism be accepted as legitimate. This conveniently deflects attention from some of the 'non-evidentiary justifications' guideline groups use to make decisions. Because these are also decisions captured by, and protected under, the 'evidence-based' framing they are partly immunised from legitimate criticism. This rhetoric move achieves a 'closing' of the debate on the optimal treatment of a given condition by delegitimising critique that cannot muster compelling medical studies that provide evidence that directly challenges the recommendations of the guidelines. This has the effect of shutting out some voices from the debate and restricting the boundaries of who can provide inputs to this process and the people who are not included.

The second effect of the rhetoric move towards objectivity is probably the most profound, if the hardest to substantiate, as it results in the transformation of a messy contextualised social process of negotiation into a guideline whose recommendations can make a claim to being objective. This is not objectivity in the form of a universal and objective truth, but it does enable a temporary claim that legitimate scientific knowledge is being used to address clinical questions. What makes this particular claim special is the fact that it is believed and supported by a methodology: evidence-based guideline making. It is because of a belief that this methodology will produce the most reliable scientific knowledge, that guidelines can claim their recommendations have legitimacy. What is peculiar is that it does not matter whether guidelines actually follow this evidence-based

model closely in practice; it is enough to claim that guidelines are evidence-based. The text of the guidelines, in varying degrees, provides the assurance that the correct procedures were followed. It is for this reason that, especially in the case of the NICE guidelines, extensive appendices are included at the end of the guideline text. The main function of these appendices is to reassure the reader and provide sources of information. If a person would choose to read the full guideline, the appendices would reassure them that the correct methodological approach has been undertaken. The Dutch guidelines are less ostentatious in displaying the underlying methodological choices, and the guidelines the NHG produces are far shorter. But nonetheless the Dutch guidelines make the same claim, and nominally employ the same methodology the English guideline groups used. The reason the Dutch guidelines are less explicit about methodology is that they have another important source of legitimacy, one that trumps considerations of evidence-based medicine, and that is the reputation of the NHG, which authorises the guidelines. The broad support and virtually universal membership of the NHG by GPs in the Netherlands guarantees the legitimacy of the guidelines to a large extent.

In conclusion, the analysis in this thesis suggests that medical communities have limited reliable medical knowledge, and that despite the claim of being evidence-based, guidelines are still the result of consensus between medical experts. This contradicts the founding claim of evidence-based guideline making that guideline making is no longer

based on the subjective knowledge of medical professionals, but on solid empirical evidence.

The criticism of the evidence hierarchy such as those voiced by Chin[2] is therefore well founded. Guideline group members might occasionally mention the evidence-hierarchy but interpret this is very different ways, in some cases excluding some types of cases from consideration in other cases preferring a different hierarchy all together. More importantly this study has shown how making evidence-based guidelines is a process of using complementary knowledges that go beyond a mere following of an evidence framework, but drawn upon various different frames.

In line with the work of Cambrosio et al[3] regulatory objectivity is a good term to describe the process of guideline making, a collective process where people (guideline makers) and objects (medical knowledge/evidence) are connected by specific coordination regimes. This study builds on the work of Cambrosio by providing empirical examples of how these coordination regimes might function. I would argue that every guideline group creates a distinct frame that draws upon various sources of knowledge such as clinical

---

2 Benjamin H Chin-Yee, 'Underdetermination in Evidence-Based Medicine' (2014) 20 Journal of Evaluation in Clinical Practice 921.

3 Alberto Cambrosio and others, 'Regulatory Objectivity and the Generation and Management of Evidence in Medicine' (2006) 63 Social Science & Medicine 189.

practice, institutional culture and medical knowledge to formulate what in this case the result will be the 'coordination regime', for a specific guideline.

The earlier work of Moreira on guidelines suggests two repertoires of evaluation: process and practice.[4] In his work he suggests guideline groups would want to consider both the technical aspects of the knowledge they gather, the quality of the medical studies and their methodological soundness, as well as the practical usability of this medical knowledge. This study supports Moreira's findings insofar as the fact that guideline groups do indeed discuss those aspects in their development of medical guidelines. Moreira, however, leaves out some important aspects, he does not acknowledge the importance of institutional factors, and does not consider the specific position the guideline has vis-a-vis other medical professionals. This study argues that these frames are equally important. A further limitation of Moreira's study is that he does not consider the preconceptions guideline makers bring to the negotiation table *before* the process of guideline drafting begins. This study has shown how making sense of evidence in a guideline group is closely connected to a frame of reference that is comprised of pre-existing knowledge and clinical expertise among other factors.

---

[4] Tiago Moreira, 'Diversity in Clinical Guidelines: The Role of Repertoires of Evaluation' (2005) 60 Social Science & Medicine 1975.

This study would therefore argue for calling guidelines evidence-informed while being mindful of the subjective and contextual process of constructing meaning and 'facts' from evidence.

## 7.2 Institutions

This study has highlighted the importance of addressing the institutional factors influencing guideline drafting. Institutional impacts are multi-layered and range from direct influence on the composition of guideline groups, to a steering of the management of conflicts of interest, to creating a specific institutional culture of guideline drafting.

Institutions are important in the process of publishing and endorsing the guidelines that guideline groups draft. The institutional environment of guideline production can shape and form the initial terms and scope of the guideline, the outputs of the searching for evidence, and how the guidelines are received by external audiences. This study provides some insight into how guideline making institutions shape guideline groups and inform the guidelines they produce. Institutional factors are a critical part of the reason why guidelines differ, as the organisational set-up of guidelines influences their drafting.

Critically, the influence of institutions on guidelines is at the centre of the guideline drafting process: it affects the role conflicts of interests have to play, it shapes the framing of the evidence through the perception of a specific cultural history by guideline makers,

it influences the composition of the guideline group, and it imparts specific directions on what guidelines are for and so what their proper role is in the relevant healthcare systems.

Institutions also shape the framing of evidence by guideline group members. This framing influences the purpose and aims of the guidelines and of the guideline making process. How guideline group members view evidence is partly due to how they view the institutional context they have to operate in during the process of guideline drafting. Evidence and institutional factors are here presented as analytically separate categories, but in reality they are tied very closely together. A frame that a guideline group employs is, at least partly, informed by an understanding of what the institutions 'demands' what its processes and resources are, and how it conceptualises guidelines as part of an institutional mission.

The position of a guideline within a healthcare system in a wider sense is steered by a perception of an institutional mission. Who are we making guidelines for? Just one profession, or the whole healthcare system? These considerations will shape the approach a guideline group will take in the development of a guideline.

Institutions exert control through deciding who can join guideline groups and who cannot. These decisions matter, as the presence of guideline group members with specific expertise can be linked to the inclusion of treatments they are experts in.

All these aspects taken together have the effect of creating cultures of institutional practice that place demands on medical professionals while they are drafting guidelines. Institutions achieve this through promulgating rules and specific cultures of institutional practice, which directly inform the medical guidelines produced. This research opens up a fresh perspective to consider guidelines, in a way that heeds the particular institutional embeddedness of the processes of the development of guidelines, and other instruments developed by professionals.

This study has shown that institutional cultures are more important for guidelines than has been realised before. Making a guideline is something that happens within an institution that provides at least some of the 'cultural scripts'[5], that steer choices in the development of guidelines.

---

[5] John W Meyer and Ronald L Jepperson, 'The "Actors" of Modern Society: The Cultural Construction of Social Agency' (2000) 18 Sociological Theory 100.

## 7.3 Group Decision-making

Recommendations in a guideline are the result of the consideration of evidence within a frame of understanding, embedded in the wider institutional context. The process of group decision-making within a guideline provides the 'glue' or the bridge between the wider institutional frame and the consideration of the evidence. This study presents several factors of the decision-making process that together form part of the frame employed by the guideline group. Some of these are present before the group start to develop the guideline, others are constructed anew by every different guideline group, and their uniqueness provides the key to explaining differences in the development of guidelines.

Four of these factors are prominent in the discussions of all the guideline groups, although their intensity varies: organisation of care, clinical practice, conceptions of disease, and perceptions of change in evidence.

A relatively minor factor are the discussions within a guideline group on the organisation of care. Deciding about recommendations does not only involve recommending a certain treatment, or advising against one, but can also mean that the guideline group makes clear who should provide the treatment. When a guideline group would like to change the organisation of care they will interpret the evidence with this goal in mind, stressing the factors that support their position.

Of greater importance is the specific backdrop of clinical practice in existence before a guideline is developed. This study has shown `how the wish to change clinical practice might cause resistance within a guideline group. If a guideline group wants to (radically) change current practice by including new treatments, or wants to drop existing treatments that are widely used a *qualitative* assessment of the evidence will be made to ascertain whether the group feels the threshold to recommend a change in practice has been reached. These discussions can be very heated, sometimes leading to a breakdown in a guideline group, as the organisation of a pathway of care is the defining feature of a guideline in the eyes of medical professionals.

Underlying the discussion on the best organisation of care or the wish to change clinical practice are ideas about the nature of the disease at hand. Subtler, and harder to show, but important are the subtle differences between guideline groups in the ideas they have about the model patient, or the exact nature of the disease. The diabetes guideline provides a pertinent illustration as the differences in ideas about the (model) patient drove the two guidelines in very different directions. A model of diabetes care that prioritises the tailoring of care will steer the group in the direction of a highly tailored guideline with many exceptions, as happened in England. The Dutch approach was heavily steered towards a uniform model of diabetes care, purposefully restricting choice in furtherance of this aim.

These different considerations play out while medical knowledge moves forward. As part of the work of creating a new guideline, guideline groups will refer to the previously published guideline. One aspect the group needs to reach consensus on is whether the evidence has changed sufficiently to warrant a change in a guideline. This assessment is qualitative in nature, and various guideline groups employ have very different ideas about what evidence is needed to change a recommendation.

The literature on group decision-making suggests that the composition of and power differences within a guideline group affect the process of decision-making in a guideline group.[6] This has found limited support for the arguments in this strand of literature. Power differences in groups have less of an influence in this study than might have been expected on the basis of this literature. The collective nature of guideline group decision- making provides a moderating influence that prevents differences in seniority between guideline group members from influencing the final recommendations in a guideline. There are some instances, however, that prove the exception to the rule: the presence of a specialist in acupuncture in the English back pain guideline group drove the recommendation for that treatment. Even in that case the composition of the group itself

---

6 C Pagliari, J Grimshaw and M Eccles, 'The Potential Influence of Small Group Processes on Guideline Development' (2001) 7 Journal of evaluation in clinical practice 165; Tim Hopthrow, Gene Feder and Susan Michie, 'The Role of Group Decision Making Processes in the Creation of Clinical Guidelines' (2011) 23 International Review of Psychiatry 358.

was only a minor factor, other aspects that helped create a multi-layered frame of meaning were more predominant.

The literature on medical politics and healthcare interrogates the link between power and knowledge in healthcare. Guidelines can be tools to convert political questions into technical questions.[7] This study has given empirical support to these claims. Calling guidelines evidence-based is a rhetorical device to transform normative questions into technical ones. Contentious issues and subjective decisions taken in guideline groups can be hidden from view through the employment of the language of evidence-based medicine. The case studies have also shown how criticism of guidelines can be defeated purely to reference to medical studies, claiming the guideline is thus evidence-based, ignoring critique that asks guideline groups to make more explicit the values (or frames) the have based their decisions on.

The institutional theory provided by Meyer and Jepperson on 'agentic actors'[8] turns out to fit very well with the position and behaviour of guideline makers. All of them described their work in terms of achieving societal goods: better quality care. The case

---

7 Gary S Belkin, 'The Technocratic Wish: Making Sense and Finding Power in the "Managed" Medical Marketplace' (1997) 22 Journal of Health Politics, Policy and Law 509; Maya J Goldenberg, 'On Evidence and Evidence-Based Medicine: Lessons from the Philosophy of Science' (2006) 62 Social Science & Medicine 2621; Marc A Rodwin, 'Commentary: The Politics of Evidence-Based Medicine' (2001) 26 Journal of Health Politics, Policy and Law 439.

8 John W Meyer and Ronald L Jepperson, 'The "Actors" of Modern Society: The Cultural Construction of Social Agency' (2000) 18 Sociological Theory 100.

studies provide ample evidence however that self-interest in pursuing a position in a guideline motivated most actors as well. This fits neatly with Meyer and Jepperson's thesis that actors need to describe their actions as driven by ideals to gain high(er) status in society while masking their self-interest.

These different strands of literature fail to acknowledge the specific collective nature of the construction of frames in guideline groups. While guideline makers might be self-interested they are forced to work together to create a collective idea of the task at hand, and systematically make decisions within those parameters.

## 7.4 Costs

There is hardly any literature that conceptualises cost and cost-effectiveness studies as an integral part of guideline development. The differences in institutional positioning of considerations of cost in the English and Dutch healthcare system have nonetheless influenced guidelines. This study has shown how evidence on cost matters is an extra aspect that drives guidelines groups and could be seen as a specific form of evidence.

The fact that cost considerations are integral to deliberations of the English guideline groups might plausibly mean that they would recommend cheaper treatments, limited by what they could do through specific institutional concerns. Dutch groups, not limited in that way would be free to recommend more expensive treatments. The real

picture was different from this simplified model, English guideline groups found ways around cost constraints and recommended more expensive treatments in some guidelines. The Dutch groups used cost as an informal argument in the discussion to reject some more expensive treatments. The example of Exenatide in the English Diabetes guideline group was particularly relevant to show how guideline groups can stretch the definition of cost-effective to accommodate clinical choices that satisfy their preferences.

Guidelines do not only consider medical evidence but deploy economic evidence as well, using various degrees of formality as the case studies show. Guideline makers have to make choices, and compromises, to be able to square ideas on the optimal clinical treatments with that of optimal cost-effectiveness. The cost of advising for drugs that are seen as not cost-effective is high, as guideline groups, especially in England, will be under institutional pressure to justify that choice very thoroughly. In response, guideline groups restrict patient access to some expensive treatments to be able to justify those treatments as cost-effective. Consequently, evidence from medical studies is not the only factor steering recommendations in guidelines, and its role is smaller when treatments are more expensive.

The role of costs in guidelines remains an underexplored area, this study has only been able to explore this issue in a few guidelines, it is likely cost will become more prominent in the future as healthcare resources remain scarce.

## 7.5 External Factors

While guidelines are nominally independent external factors might plausibly affect their development. In this study two particular sources of potential influence have been considered: external commentaries on guidelines and conflicts of interest.

The external commentary on guidelines has not lead to changes in the text of the guidelines. In some instances critique on a guideline might plausibly affect a future guideline, but not the ones that are being developed. One of the reasons for external factors having a lack of impact on the text of a guideline is the vested interest guideline making institutions have in retaining their independence. The institutions do not want to be seen to 'give in.' An important tool in this respect is the fact that institutions can refer to the evidence-based methodology they have employed in drafting the guideline, limiting the amount of 'valid criticism' that can be made of the guideline.

The work of Goldacre[9] has been influential in drawing attention to the potential corrupting influence of conflicts of interest on the development of guidelines. This study has indeed found examples of attempts by the pharmaceutical industry to influence guideline makers, showing the continued relevance of Goldacre's critique. The detailed descriptions of attempts to influence members of the Dutch diabetes group reflects this. Nonetheless within the limits of this study it has not been possible to adduce a causal relationship between the conflicts of interest guideline group members declared and the resulting text of a guideline. Only a correlation between the number of members who report a conflict of interest and the amount of expensive medication they recommend can be established, but this is not sufficient to say that conflicts of interest had an actual influence on any of the guidelines in this study.

## 7.6 Regulation

This study has suggested that guidelines could be conceptualised as being both tools of mandated self-regulation and at the same time an instrument of control for governments to control medical professionals. This inherent tension in the development of a guideline has indeed played out in the case studies discussed in this study. Guideline can be placed on a

---

9 Ben Goldacre, *Bad Pharma: How Drug Companies Mislead Doctors and Harm Patients* (Fourth Estate 2012).

continuum from guidelines that are 'purely self-regulation' of medical professionals to those that are 'exclusively instruments of control' for governments.

Transparency of professional practice for outside observers is one of the effects of developing guidelines. This can illustrate the argument that guidelines provide a double function, both providing increased autonomy for the professional, while at the same time permitting more control. Levay and Waks suggest that the transparency that guidelines bring can be a positive development for the professional.[10] The results of the case studies provide some support for this positive view of guidelines. One of the main reasons for professionals to engage in guideline making efforts is to remain in control of the criteria of evaluation of their work as professionals. In that sense all the guidelines studied were successful, the definitions provided by professionals of what 'counts' as good clinical practice remains the standard the profession is judged on, these aspects have not been displaced in either country by government defined initiatives.

Complicating that picture is the finding that in this study guidelines perform both functions at the same time, while one of these aspects might be seen as dominant. The English diabetes guidelines were perceived as being influenced by government mandated rationing, through the extensive use of cost-effectiveness data. The Dutch back pain

---

[10] Charlotta Levay and Caroline Waks, 'Professions and the Pursuit of Transparency in Healthcare: Two Cases of Soft Autonomy' (2009) 30 Organization Studies 509.

guideline was more closely linked to professional aims of self-regulation and organisation of care. But all guidelines had aspects of both dimensions, which leads to the conclusion that guidelines are indeed specific regulatory tools who embody this duality.

The work of Evetts[11] and especially Fournier[12] on professionalization is particularly relevant. Fournier's assessment of professionalization initiatives, which include guidelines as disciplinary tools, chimes with this research. Guidelines can indeed be seen as disciplinary tools, as the resistance against the English back pain guideline has demonstrated. It was suggested corporatist healthcare systems, such as the Dutch one would allow for less clinical change through guidelines. This assumption turns out to be untrue, the opposite has shown to be correct. In the Dutch system, a more dispersed system of guideline making creates a closer relationship between professional and guideline makers which allows for greater clinical change, because the legitimacy of the guideline is greater.

---

11 Julia Evetts, 'Sociological Analysis of Professionalism: Past, Present and Future' (2011) 10 Comparative Sociology 1.

12 Valérie Fournier, 'The Appeal to "Professionalism" as a Disciplinary Mechanism' (1999) 47 Sociological Review 280.

Conceptualising guidelines as regulation has not been a central plank of this thesis, nonetheless it holds potential for comparative study of forms of regulation by different experts in different domains of public policy making.

## 7.7 Concluding Remarks

Technical, dry, and frankly often boring documents, is how we can best describe guidelines. This façade of technical, uncontroversial, even boring guidance is, however, one of the central claims to success for guidelines. It allows for the transformation of political question into technical ones, and for these now technical issues to be solved by technical experts, the healthcare professionals themselves. It turned out that the more normal and uncontroversial the externally portrayed text of the guideline might look, the more controversial and heated were the discussions that went into producing them. The world of guidelines is fascinating, sometimes weird, but always interesting. The focus of this study might have started with the observance of differences in text despite similarities in circumstances but it has moved deeper to try to understand the social world behind these words on paper, the lived reality of creating regulation and what this means for the people involved.

The results of this study do not provide a uniform answer that can explain the similarities and differences in guidelines. The study argues that guidelines are a way for social and normative choices to be concealed and presented as technical choices behind the

label ‘evidence-based’. Guidelines provide specific forms, and formal symbolic practices, to be able to justify and legitimize the making of choices on optimal healthcare. The searching for evidence, the extensive modelling of economic choices, the institutional hurdles that have to be cleared, all work towards the goal of mystifying the fact that a group of men and women in a room have to use their own judgement to make choices that will affect other healthcare professionals, without being able to claim a position of authority over them. In the evidence-based literature it is often lamented that guidelines are ineffective, not implemented, etc. While this might be true, the truly remarkable thing about the process of drafting guidelines is that so many of them are successful.

Evidence-based guidelines are better characterised as evidence-informed guidelines. Not because the decisions by guideline groups are taken on a whim, without regard to scientific knowledge, far from it, rather because it is the social process and the cognitive frames that guide this process that provide the real impetus that results in guideline texts.

Guidelines are highly contextual creations, where replication of ‘results’ is virtually impossible, because in every new guideline group a new frame of understanding is created, in such a way that even if ‘evidence’ could be said to be completely the same[13] new and different understandings would be reached every time. The similarity between guidelines

[13] Ignoring for the moment how unlikely this is in practice.

is therefore even more striking than differences, and it is due to the fact that guidelines across countries are closely tied to the medical field they speak to, that similarities exist, despite sometimes large differences in selecting medical studies.

The resulting guideline text claims (or can be perceived to claim) objective results of robust and impartial assessment of all medical knowledge available, so as to enable practitioners to trust their judgements and apply them in clinical practice as if they were their own judgements.

Guidelines have shown to be remarkable documents; hiding behind their patient words a varied, complex, and interesting world of colliding interests, characters, values and principles. In conclusion this thesis has critically studied the development of evidence-based guidelines and argues that guidelines are context dependent, socially constructed, and negotiated regulatory documents.

# BIBLIOGRAPHY


'6.1 Drugs Used in Diabetes : British National Formulary' <https://www.evidence.nhs.uk/formulary/bnf/current/6-endocrine-system/61-drugs-used-in-diabetes> accessed 21 July 2016

Alonso-Coello P and others, 'The Quality of Clinical Practice Guidelines over the Last Two Decades: A Systematic Review of Guideline Appraisal Studies' (2010) 19 Quality and Safety in Health Care 1

Armstrong D, 'Clinical Autonomy, Individual and Collective: The Problem of Changing Doctors' Behaviour' (2002) 55 Social Science & Medicine 1771

Armstrong D and others, 'Development of Clinical Guidelines in a Health District: An Attempt to Find Consensus.' (1992) 1 Quality in Health Care 241

Asch SE, 'Opinions and Social Pressure' (1955) 193 Scientific American 31

Atkin S and Walton C, 'Commentary: Controversies in NICE Guidance on Management of Type 2 Diabetes' (2008) 336 BMJ 1308

Atkins L and others, 'The Process of Developing Evidence-Based Guidance in Medicine and Public Health: A Qualitative Study of Views from the Inside' (2013) 8 Implementation Science 101

Bagust A and Beale S, 'Modelling EuroQol Health-Related Utility Values for Diabetic Complications from CODE-2 Data' (2005) 14 Health Economics 217

Banta D, 'The Development of Health Technology Assessment' (2003) 63 Health Policy 121

Barnett AH, 'NICE Draft Type 2 Diabetes Guidelines: A Cause for Concern' [2015] The Lancet Diabetes & Endocrinology <http://linkinghub.elsevier.com/retrieve/pii/S2213858715000534> accessed 5 May 2015

Belkin GS, 'The Technocratic Wish: Making Sense and Finding Power in the "Managed" Medical Marketplace' (1997) 22 Journal of Health Politics, Policy and Law 509

Berg M, ter Meulen R and Van den Burg M, 'Guidelines for Appropriate Care: The Importance of Empirical Normative Analysis' (2001) 9 Health Care Analysis 77

Black J, 'Constitutionalising Self-Regulation' (1996) 59 The Modern Law Review 24

——, ‘Critical Reflections on Regulation’ (January 2002) <http://www.lse.ac.uk/CARR> accessed 14 April 2016

——, ‘Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes’ (2008) 2 Regulation & Governance 137

Black J and Rouch D, ‘The Development of the Global Markets as Rule-Makers: Engagement and Legitimacy’ (2008) 2 Law and Financial Markets Review 218

Black N and others, ‘Consensus Development Methods: A Review of Best Practice in Creating Clinical Guidelines’ (1999) 4 Journal of Health Services Research & Policy 236

Board HC, ‘Standing Group on Health Technology’ (1998) 2 Health Technology Assessment <https://www.researchgate.net/profile/David_Greatbatch/publication/13370611_Qualitative_Research_Methods_in_Health_Technology_Assessment_A_Review_of_the_Literature/links/0912f50b7847ab963e000000.pdf> accessed 20 May 2016

‘Boetereapport NMa’ (*Rapport of a fine by the Dutch Competition Authority*) <http://medischcontact.artsennet.nl/web/file?uuid=03be9098-ad0e-4555-a31e-d726f780a925&owner=86bbbdd9-5ff8-4303-ab88-f8888509486c> accessed 25 May 2014

Boffey D, ‘NHS Chief Admits to Dismay at Jeremy Hunt’s Meddling’ *The Guardian* (21 December 2013) <http://www.theguardian.com/society/2013/dec/21/nhs-service-jeremy-hunt-malcolm-grant> accessed 17 May 2014

Boltanski L and Thévenot L, ‘The Reality of Moral Expectations: A Sociology of Situated Judgement’ (2000) 3 Philosophical Explorations 208

Braithwaite J, ‘Enforced Self-Regulation: A New Strategy for Corporate Crime Control’ (1982) 80 Michigan Law Review 1466

——, ‘The New Regulatory State and the Transformation of Criminology’ (2000) 40 British Journal of Criminology 222

Brouwers MC and others, ‘AGREE II: Advancing Guideline Development, Reporting and Evaluation in Health Care’ (2010) 63 Journal of Clinical Epidemiology 1308

Burgers JS and others, ‘Inside Guidelines Comparative Analysis of Recommendations and Evidence in Diabetes Guidelines from 13 Countries’ (2002) 25 Diabetes Care 1933

Burgers JS and others, ‘Characteristics of Effective Clinical Guidelines for General Practice.’ (2003) 53 The British Journal of General Practice 15

Burgers JS and others, ‘Towards Evidence-Based Clinical Practice: An International Survey of 18 Clinical Guideline Programs’ (2003) 15 International Journal for Quality in Health Care 31

Burgers JS and Everdingen JJE van, ‘Evidence-based richtlijnontwikkeling in Nederland: het EBRO-platform’ (2004) 148 Nederlands Tijdschrift voor Geneeskunde 2057

Burton AK and others, ‘Information and Advice to Patients with Back Pain Can Have a Positive Effect: A Randomized Controlled Trial of a Novel Educational Booklet in Primary Care’ (1999) 24 Spine 2484

Cambrosio A and others, ‘Regulatory Objectivity and the Generation and Management of Evidence in Medicine’ (2006) 63 Social Science & Medicine 189

——, ‘Regulatory Objectivity and the Generation and Management of Evidence in Medicine’ (2006) 63 Social Science & Medicine 189

Campbell D, ‘“Control Freak” Jeremy Hunt Accused of Manipulating NHS for Political Ends’ *The Guardian* (27 March 2014) <http://www.theguardian.com/society/2014/mar/27/jeremy-hunt-control-freak-manipulating-nhs> accessed 17 May 2014

Caporaso JA and others, ‘Still a Regulatory State? The European Union and the Financial Crisis’ (2015) 22 Journal of European Public Policy 889

Carpenter DP and Moss DA (eds), *Preventing Regulatory Capture: Special Interest Influence and How to Limit It* (Cambridge University Press 2014)

‘CBO-Richtlijn Voor Diagnostiek En Behandeling van Acute En Chronische Aspecifieke Lage Rugklachten | Nederlands Tijdschrift Voor Geneeskunde’ <http://www.ntvg.nl.proxy-ub.rug.nl/artikelen/cbo-richtlijn-voor-diagnostiek-en-behandeling-van-acute-en-chronische-aspecifieke-lage/ingezonden-mededelingen#1225672> accessed 21 August 2014

Chalmers I, ‘The Cochrane Collaboration: Preparing, Maintaining, and Disseminating Systematic Reviews of the Effects of Health Care’ (1993) 703 Annals of the New York Academy of Sciences 156

Chin-Yee BH, ‘Underdetermination in Evidence-Based Medicine’ (2014) 20 Journal of Evaluation in Clinical Practice 921

Claxton K and others, 'Discounting and Cost-Effectiveness in NICE – Stepping Back to Sort out a Confusion' (2006) 15 Health Economics 1

——, *Methods for the Estimation of the NICE Cost Effectiveness Threshold* (University of York, Centre for Health Economics 2013) <http://www.york.ac.uk/media/che/documents/reports/resubmitted_report.pdf> accessed 15 August 2016

Clinical Standards Advisory Group, *Back Pain.* (HMSO 1994)

Cochrane AL, *Effectiveness and Efficiency: Random Reflections on Health Services* (Nuffield Provincial Hospitals Trust 1972)

Cochrane AL, Cox JG and Jarman TF, 'Pulmonary Tuberculosis in the Rhondda Fach' (1952) 2 British Medical Journal 843

Cohen AM, Stavri PZ and Hersh WR, 'A Categorization and Analysis of the Criticisms of Evidence-Based Medicine' (2004) 73 International Journal of Medical Informatics 35

Committee on Conflict of Interest in Medical Research, Education and Institute of Medicine (U.S.), *Conflict of Interest in Medical Research, Education, and Practice* (Bernard Lo and Marilyn J Field eds, National Academies Press 2009)

'Contributed Poster Presentations' (2006) 9 Value in Health A24

Costa-Font J, Courbage C and McGuire A, *The Economics of New Health Technologies* (Oxford University Press 2009) <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199550685.001.0001/acprof-9780199550685> accessed 1 May 2015

Currie CJ and others, 'Survival as a Function of HbA1c in People with Type 2 Diabetes: A Retrospective Cohort Study' (2010) 375 The Lancet 481

'Databank | GIPdatabank' <https://www.gipdatabank.nl/databank.asp?tabel=85_dia&geg=ddd&item=bijlage&label=00-totaal> accessed 12 October 2016

Davidoff F and others, 'Evidence Based Medicine' (1995) 310 BMJ 1085

De Friesland Zorgverzekeraar, 'Basisverzekering Online Verzorgd Polis 2014' (*Basic Health Policy Online Care Policy 2014*) <https://www.defriesland.nl/Images/Basisverzekering-Online-Verzorgd-Polis-2014-11650.pdf> accessed 26 June 2014

Department of Health, *Equity and Excellence: Liberating the NHS.* (Stationery Office 2010)

'Developing NICE Clinical Guidelines' <http://www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/developing_nice_clinical_guidelines.jsp> accessed 26 May 2014

'Developing NICE Guidelines: The Manual' <https://www.nice.org.uk/process/pmg20/> accessed 15 August 2016

'Development and Validation of an International Appraisal Instrument for Assessing the Quality of Clinical Practice Guidelines: The AGREE Project' (2003) 12 Quality and Safety in Health Care 18

Devlin N and Parkin D, 'Does NICE Have a Cost-Effectiveness Threshold and What Other Factors Influence Its Decisions? A Binary Choice Analysis' (2004) 13 Health Economics 437

'Diabetes-Type-2-Update-Consultation-Table2 Cg-66.Pdf' <https://www.nice.org.uk/guidance/cg66/documents/diabetes-type-2-update-consultation-table2> accessed 14 May 2015

DiMaggio PJ and Powell WW, 'The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields' (1983) 48 American Sociological Review 147

Driskell JE and Mullen B, 'Status, Expectations, and Behavior: A Meta-Analytic Review and Test of the Theory' (1990) 16 Personality and Social Psychology Bulletin 541

Duckworth W and others, 'Glucose Control and Vascular Complications in Veterans with Type 2 Diabetes' (2009) 360 New England Journal of Medicine 129

Duhem PMM, *The Aim and Structure of Physical Theory* (Princeton University Press 1991)

'Dutch Cochrane Centre' <http://dcc.cochrane.org/> accessed 26 May 2014

Eagly AH, 'Gender and Social Influence: A Social Psychological Analysis.' (1983) 38 American Psychologist 971

Egeler C, 'NICE Guidline on Backpain Is Misleading | The BMJ' <http://www.bmj.com.proxy-ub.rug.nl/rapid-response/2011/11/02/nice-guidline-backpain-misleading> accessed 4 July 2016

EMA, ‘European Medicines Agency Recommends Suspension of Avandia, Avandamet and Avaglim. Anti-Diabetes Medication to Be Taken off the Market.’ <http://www.ema.europa.eu/docs/en_GB/document_library/Press_release/2010/09/WC500096996.pdf> accessed 1 September 2016

Everdingen JJE van, *Consensusontwikkeling in de geneeskunde* (Bohn, Scheltema en Holkema 1988)

Everdingen JJE van, *Evidence-based richtlijnontwikkeling: een leidraad voor de praktijk* (Bohn Stafleu Van Loghum 2004)

——, *Handboek evidence-based richtlijnontwikkeling: een leidraad voor de praktijk* (Bohn Stafleu van Loghum, Springer Media 2014)

Evetts J, ‘Sociological Analysis of Professionalism: Past, Present and Future’ (2011) 10 Comparative Sociology 1

Evidence-Based Medicine Working Group, ‘Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine’ (1992) 268 JAMA 2420

Feinstein MD AR and Horwitz MD RI, ‘Problems in the “Evidence” of “Evidence-Based Medicine”’ (1997) 103 The American Journal of Medicine 529

Feyerabend P, *Against Method* (Verso 2010)

Field MJ and others, *Clinical Practice Guidelines Directions for a New Program* (National Academy Press 1990) <http://site.ebrary.com/id/10068353> accessed 25 November 2013

Fournier V, ‘The Appeal to “Professionalism” as a Disciplinary Mechanism’ (1999) 47 Sociological Review 280

Freidson E, *Professionalism, the Third Logic* <http://www.press.uchicago.edu/ucp/books/book/chicago/P/bo3621932.html> accessed 2 June 2016

Fritz JM, Delitto A and Erhard RE, ‘Comparison of Classification-Based Physical Therapy with Therapy Based on Clinical Practice Guidelines for Patients with Acute Low Back Pain: A Randomized Clinical Trial’ (2003) 28 Spine 1363

Gale EAM, ‘Conflicts of Interest in Guideline Panel Members’ (2011) 343 BMJ d5728

General Medical Council, 'Good Medical Practice (2013)' (2013) <http://www.gmc-uk.org/static/documents/content/GMP_2013.pdf_51447599.pdf> accessed 9 January 2014

George JN, Vesely SK and Woolf SH, 'Conflicts of Interest and Clinical Recommendations Comparison of Two Concurrent Clinical Practice Guidelines for Primary Immune Thrombocytopenia Developed by Different Methods' (2014) 29 American Journal of Medical Quality 53

'Goedkeuringsbesluit Kaderbesluit En Drie Specifieke Besluiten van de CHVG Betreffende Huisartsgeneeskunde, Verpleeghuisgeneeskunde En Medische Zorg Voor Verstandelijk Gehandicapten' (*Ratification of the decision by the commission for GP care, social care and care for the mentally disabled on the training requirements*) <http://wetten.overheid.nl/BWBR0017529/geldigheidsdatum_10-02-2014> accessed 26 May 2014

Goldacre B, *Bad Pharma: How Drug Companies Mislead Doctors and Harm Patients* (Fourth Estate 2012)

Goldacre B and Heneghan C, 'How Medicine Is Broken, and How We Can Fix It' (2015) 350 The BMJ h3397

Goldenberg MJ, 'On Evidence and Evidence-Based Medicine: Lessons from the Philosophy of Science' (2006) 62 Social Science & Medicine 2621

——, 'Iconoclast or Creed?: Objectivism, Pragmatism, and the Hierarchy of Evidence' (2009) 52 Perspectives in Biology and Medicine 168

Gorter KJ and others, 'Diabetes: Glycaemic Control in Type 2 (Drug Treatments)' (2012) 2012 BMJ Clinical Evidence <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3462437/> accessed 3 February 2016

'Grading Quality of Evidence and Strength of Recommendations' (2004) 328 BMJ 1490

Grol R and others, 'Implementatie van NHG-Standaarden: Succes of Probleem?' (2010) 53 Huisarts en wetenschap 42

Group TAC, 'Intensive Blood Glucose Control and Vascular Outcomes in Patients with Type 2 Diabetes' (2008) 358 New England Journal of Medicine 2560

Group TAS, 'Long-Term Effects of Intensive Glucose Lowering on Cardiovascular Outcomes' (2011) 364 New England Journal of Medicine 818

‘Guide to the Methods of Technology Appraisal 2013 |’ <https://www.nice.org.uk/article/PMG9/> accessed 4 July 2016

Guyatt GH, ‘Evidence-Based Medicine’ [1991] ACP J Club A

Guyatt GH and others, ‘GRADE: An Emerging Consensus on Rating Quality of Evidence and Strength of Recommendations’ (2008) 336 BMJ 924

Harari E, ‘Whose Evidence? Lessons from the Philosophy of Science and the Epistemology of Medicine’ (2001) 35 Australian and New Zealand Journal of Psychiatry 724

Harrison MI, *Implementing Change in Health Systems : Market Reforms in the United Kingdom, Sweden, and the Netherlands* (SAGE 2004)

Health Consumer Powerhouse, ‘European Health Consumer Index’ <http://www.healthpowerhouse.com/files/EHCI_2015/EHCI_2015_report.pdf> accessed 16 April 2016

Healy D, *Let Them Eat Prozac: The Unhealthy Relationship Between the Pharmaceutical Industry and Depression* (1 edition, NYU Press 2006)

Hemingway H and others, ‘Appropriateness Criteria for Coronary Angiography in Angina: Reliability and Validity’ (2008) 149 Annals of Internal Medicine 221

Herk R van, ‘Artsen onder druk. Het kwaliteitsbeleid van de medische beroepen in Groot-Brittannië, Nederland en België tussen 1970 en 1996, als gevolg van de interne en externe druk op de zelfregulering van artsen.’ (Erasmus Universiteit Rotterdam 1997) <http://repub.eur.nl/pub/18416/971023_HERK,%20Rudi%20van.pdf> accessed 24 November 2014

‘Herziene NHG-Standaard Diabetes Mellitus Type 2 | Farma-Magazine.Nl Onafhankelijk Platform Voor Apothekers’ <http://farma-magazine.nl/nhg-standaard-diabetes-mellitus-type-2/> accessed 22 April 2015

Hester JB, ‘Official Statement from the British Pain Society’ [2015] The BMJ <http://www.bmj.com/rapid-response/2011/11/02/official-statement-british-pain-society> accessed 14 September 2015

‘Hoe de Farmaceutische Marketingmachine Medische Richtlijnen Manipuleert’ (*De Correspondent*) <https://decorrespondent.nl/467/Hoe-de-farmaceutische-marketingmachine-medische-richtlijnen-manipuleert/14363052-1f180d9f> accessed 22 April 2015

Hood C and others, *Regulation inside Government: Waste Watchers, Quality Police, and Sleaze-Busters* (Oxford University Press 1999)

Hopthrow T, Feder G and Michie S, 'The Role of Group Decision Making Processes in the Creation of Clinical Guidelines' (2011) 23 International Review of Psychiatry 358

Huang ES and others, 'Glycemic Control, Complications, and Death in Older Diabetic Patients: The Diabetes and Aging Study' (2011) 34 Diabetes Care 1329

'Huisarts Boycot Dure Diabetespil, Farmaceuten Lopen Geld Mis | Economie | de Volkskrant' <http://www.volkskrant.nl/economie/huisarts-boycot-dure-diabetespil-farmaceuten-lopen-geld-mis~a4351320/> accessed 4 August 2016

Hutchings A and Raine R, 'A Systematic Review of Factors Affecting the Judgments Produced by Formal Consensus Development Methods in Health Care' (2006) 11 Journal of Health Services Research & Policy 172

Ibrahim T, Tleyjeh IM and Gabbar O, 'Surgical versus Non-Surgical Treatment of Chronic Low Back Pain: A Meta-Analysis of Randomised Trials' (2008) 32 International Orthopaedics 107

Insitituut voor Verantwoord Medicijngebruik, 'Kosten Bloedglucoseverlagende Middelen Nederland versus Duitsland' <http://www.medicijnbalans.nl/content/userfiles/files/Kosten%20bloedglucoseverlagende%20middelen%20Nederland%20versus%20Duitsland(1).pdf> accessed 4 August 2016

'International Guideline Library — Guidelines International Network' <http://www.g-i-n.net/library/international-guidelines-library> accessed 27 June 2014

Janis IL, *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes,* vol viii (Houghton Mifflin 1972)

Jones DJ and others, 'Conflicts of Interest Ethics: Silencing Expertise in the Development of International Clinical Practice Guidelines' (2012) 156 Annals of Internal Medicine 809

'Kennisinstituut van Medisch Specialisten - Het Belang van Een Goede Richtlijn' (*Knowledge Institute of Medical Specialists - The importance of a good guideline*) <http://kims.orde.nl/het-kennisinstituut-voor-u/richtlijnen/het-belang-van-een-goede-richtlijn.html> accessed 11 June 2014

‘Klacht NHG Tegen MSD Gegrond Verklaard | NHG’ <https://www.nhg.org/actueel/nieuws/klacht-nhg-tegen-msd-gegrond-verklaard> accessed 14 October 2016

Klazinga N, ‘Compliance with Practice Guidelines: Clinical Autonomy Revisited’ (1994) 28 Health policy (Amsterdam, Netherlands) 51

Klazinga NS, ‘Quality management of medical specialist care in the Netherlands. An explorative study of its nature and development’ (Erasmus Universiteit Rotterdam 1996) <http://www.bmg.eur.nl/fileadmin/ASSETS/bmg/Onderzoek/Promoties/Promoties_1996/Klazinga_1996_web.pdf> accessed 24 September 2014

Kmietowicz Z, ‘President of British Pain Society Is Forced from Office after Refusing to Denounce NICE Guidance on Low Back Pain’ (2009) 339 BMJ b3049

Knaapen L, ‘European Regulation and Harmonization of Clinical Practice Guidelines’, *European Union Public Health Policy: Regional and Global Trends* (Routledge 2013)

——, ‘Being “Evidence-Based” in the Absence of Evidence: The Management of Non-Evidence in Guideline Development’ (2013) 43 Social Studies of Science 681

——, ‘Pragmatic Evidence and Textual Arrangements: A Case Study of French Clinical Cancer Guidelines’ (2010) 71 Social science & medicine (1982) 685

Knaapen LA, ‘In Search of Standards That Avoid Standardization: The Production and Regulation of Evidence Based Guidelines’ (PhD, McGill University 2013) <http://digitool.library.mcgill.ca/webclient/StreamGate?folder_id=0&dvs=1416319090786~744> accessed 18 November 2014

‘KNMG/KNMP Richtlijn Uitvoering Euthanasie En Hulp Bij Zelfdoding’ <http://knmg.artsennet.nl/web/file?uuid=9899cb4d-d598-4f58-974c-10ad38153485&owner=a8a9ce0e-f42b-47a5-960e-be08025b7b04&contentid=119543> accessed 22 October 2013

Koes BW and others, ‘Clinical Guidelines for the Management of Low Back Pain in Primary Care: An International Comparison’ (2001) 26 Spine 2504

Kuhn TS, *The Structure of Scientific Revolutions* ([2d ed, enl, University of Chicago Press 1970)

Kulkarni AV, ‘The Challenges of Evidence-Based Medicine: A Philosophical Perspective’ (2005) 8 Medicine, Health Care and Philosophy 255

Lakatos I, *The Methodology of Scientific Research Programmes* (Cambridge University Press 1999)

Landman GW and others, 'The Relationship between Glycaemic Control and Mortality in Patients with Type 2 Diabetes in General Practice (ZODIAC-11)' (2010) 60 The British Journal of General Practice 172

Latour B, *Reassembling the Social: An Introduction to Actor-Network-Theory* (OUP Oxford 2005)

Lehoux P and Blume S, 'Technology Assessment and the Sociopolitics of Health Technologies' (2000) 25 Journal of health politics, policy and law 1083

Leng G and Baker M, 'NICE Accredits Guidelines Using International AGREE Criteria' (2013) 347 BMJ f4771

Levay C and Waks C, 'Professions and the Pursuit of Transparency in Healthcare: Two Cases of Soft Autonomy' (2009) 30 Organization Studies 509

Levi-Faur D, 'The Rise of Regulatory Capitalism: The Global Diffusion of a New Order' (2005) 598 The ANNALS of the American Academy of Political and Social Science 200

Liebowitz SJ and Margolis SE, 'Path Dependence' [1999] Encyclopedia of law and economics <http://ecsocman.hse.ru/data/018/784/1216/0770book.pdf> accessed 2 November 2016

Lincoff A and others, 'Pioglitazone and Risk of Cardiovascular Events in Patients with Type 2 Diabetes Mellitus: A Meta-Analysis of Randomized Trials' (2007) 298 JAMA 1180

Loveman E and others, 'The Clinical and Cost-Effectiveness of Patient Education Models for Diabetes: A Systematic Review and Economic Evaluation' (2003) 7 Health Technology Assessment <http://www.journalslibrary.nihr.ac.uk/hta/volume-7/issue-22> accessed 29 April 2016

Majone G, 'The Rise of the Regulatory State in Europe' (1994) 17 West European Politics 77

Marshall C, *Designing Qualitative Research* (4th ed, SAGE 2006) <http://www.loc.gov/catdir/toc/ecip0519/2005026958.html> accessed 23 April 2016

May C, 'Mobilising Modern Facts: Health Technology Assessment and the Politics of Evidence' (2006) 28 Sociology of Health & Illness 513

McDonald R and Harrison S, 'The Micropolitics of Clinical Guidelines: An Empirical Study' (2004) 32 Policy & Politics 223

McGuire DK and others, 'Association Between Sitagliptin Use and Heart Failure Hospitalization and Related Outcomes in Type 2 Diabetes Mellitus: Secondary Analysis of a Randomized Clinical Trial' (2016) 1 JAMA Cardiology 126

McKinlay JB and Arches J, 'Towards the Proletarianization of Physicians' (1985) 15 International Journal of Health Services 161

McKinlay JB and Stoeckle JD, 'Corporatization and the Social Transformation of Doctoring' (1988) 18 International Journal of Health Services 191

'Medicijnbalans - NHG-Standaard Diabetes Mellitus Type 2, December 2013' <http://medicijnbalans.nl/video/40243> accessed 3 May 2015

'Medische Richtlijnen per Organisatie' (*Medical guidelines by organisation*) <http://www.artsennet.nl/Richtlijnen/Per-organisatie.htm> accessed 22 May 2014

Melander H and others, 'Evidence b(i)Ased Medicine—selective Reporting from Studies Sponsored by Pharmaceutical Industry: Review of Studies in New Drug Applications' (2003) 326 BMJ 1171

Meyer JW, 'Reflections on Institutional Theories of Organizations', *The SAGE Handbook of Organizational Institutionalism* (SAGE Publications Ltd 2008)

Meyer JW and Jepperson RL, 'The "Actors" of Modern Society: The Cultural Construction of Social Agency' (2000) 18 Sociological Theory 100

Milgram S, 'Some Conditions of Obedience and Disobedience to Authority' (1965) 18 Human Relations 57

Miller D, 'Some Hard Questions for Critical Rationalism' (2014) 15 Discusiones Filosóficas 15

'Mirror, Mirror on the Wall, 2014 Update: How the U.S. Health Care System Compares Internationally - The Commonwealth Fund' <http://www.commonwealthfund.org/publications/fund-reports/2014/jun/mirror-mirror> accessed 17 June 2014

Mirza SK and Deyo RA, 'Systematic Review of Randomized Trials Comparing Lumbar Fusion Surgery to Nonoperative Care for Treatment of Chronic Back Pain': (2007) 32 Spine 816

Moran M, *The British Regulatory State: High Modernism and Hyper-Innovation* (Oxford University Press 2003)

——, *The British Regulatory State: High Modernism and Hyper-Innovation: High Modernism and Hyper-Innovation* (Oxford University Press 2003)

Moreira T, 'Diversity in Clinical Guidelines: The Role of Repertoires of Evaluation' (2005) 60 Social Science & Medicine 1975

Mykhalovskiy E and Weir L, 'The Problem of Evidence-Based Medicine: Directions for Social Science' (2004) 59 Social Science & Medicine 1059

National Collaborating Centre for Chronic Conditions (Great Britain) and Royal College of Physicians of London, *Type 2 Diabetes National Clinical Guideline for Management in Primary and Secondary Care (Update)* (Royal College of Physicians 2008) <http://www.ncbi.nlm.nih.gov/books/NBK53885/> accessed 14 April 2015

NHG, 'Procedureboek' <https://www.nhg.org/sites/default/files/content/nhg_org/uploads/procedureboek_versie_090810.pdf> accessed 17 November 2014

——, 'Zoekstrategieen Lage Rugpijn' <https://www.nhg.org/sites/default/files/content/nhg_org/uploads/zoekstrategieen/lage-rugpijn_m54.pdf> accessed 17 November 2014

'NHG Standaard Diabetes Mellitus Type II' <https://www.henw.org/archief/pdf/id10142-diabetes-mellitus-type-ii-nhg-standaard.html> accessed 6 July 2016

'NHG-Standaard Aspecifieke Lagerugpijn | NHG' <https://www.nhg.org/standaarden/volledig/nhg-standaard-aspecifieke-lagerugpijn> accessed 21 August 2014

'NHG-Standaard Aspecifieke Lagerugpijn, Een Herziening? | Archief | Website van H&W | Huisarts En Wetenschap, Het Maandblad van Het Nederlands Huisartsen Genootschap.' <http://www.henw.org/archief/volledig/id1815-nhg-standaard-aspecifieke-lagerugpijn-een-herziening.html> accessed 20 October 2014

'NHG-Standaarden | NHG' (*Standards of the Dutch Scientifc Association of General Practicioners*) <https://www.nhg.org/nhg-standaarden> accessed 25 May 2014

'NHS England » CCG Authorisation' <http://www.england.nhs.uk/ourwork/commissioning/ccg-auth/> accessed 26 May 2014

NICE, ‘Osteoarthritis (Update): Consultation Comments Table’ (2013) <https://www.nice.org.uk/guidance/CG177/documents/osteoarthritis-update-consultation-comments-table2> accessed 2 November 2016

——, ‘Osteoarthritis (Update): Draft Guideline’ (2013) <https://www.nice.org.uk/guidance/CG177/documents/osteoarthritis-update-draft-guideline2> accessed 2 November 2016

——, ‘Code of Practice for Declaring and Managing Conflicts of Interest’ <https://www.nice.org.uk/Media/Default/About/Who-we-are/Policies-and-procedures/code-of-practice-for-declaring-and-managing-conflicts-of-interest.pdf> accessed 20 June 2016

——, ‘Type 2 Diabetes: Newer Agents’ <http://www.nice.org.uk/guidance/cg87/resources/cg87-type-2-diabetes-newer-agents-a-partial-update-of-cg66-short-guideline2> accessed 21 October 2014

‘NICE Outraged by Ousting of BPS President | The BMJ’ <http://www.bmj.com/rapid-response/2011/11/02/nice-outraged-ousting-bps-president> accessed 17 September 2015

‘NICE Threshold - Centre for Health Economics, The University of York’ <http://www.york.ac.uk/che/research/teehta/thresholds/> accessed 25 February 2015

Nissen SE and Wolski K, ‘Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes’ (2007) 356 New England Journal of Medicine 2457

Norris SL and others, ‘Conflict of Interest in Clinical Practice Guideline Development: A Systematic Review’ (2011) 6 PLoS ONE <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3198464/> accessed 23 June 2015

Ogus A, ‘Rethinking Self-Regulation’ (1995) 15 Oxford Journal of Legal Studies 97

‘Oncoline’ <http://www.oncoline.nl/> accessed 27 June 2014

‘Orde van Medisch Specialisten - Stichting Kwaliteitsgelden Medisch Specialisten (SKMS)’ <http://www.orde.nl/over-oms/wat-doet-de-oms/stichting-kwaliteitsgelden-medisch-specialisten-skms.html> accessed 11 June 2014

‘Osteoarthritis | About-This-Guideline | Guidance and Guidelines | NICE’ <http://www.nice.org.uk/guidance/cg177/chapter/about-this-guideline> accessed 9 October 2014

‘Overzicht Zorgverzekeraars in Nederland’ (*List of Health Insurers in the Netherlands*) <http://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/rubrieken/verzekering/1302-overzicht-zorgverzekeraars-in-nederland/Overzicht+zorgverzekeraars+%28in+Nederland%29.pdf> accessed 25 May 2014

Pagliari C and Grimshaw J, ‘Impact of Group Structure and Process on Multidisciplinary Evidence-Based Guideline Development: An Observational Study’ (2002) 8 Journal of Evaluation in Clinical Practice 145

Pagliari C, Grimshaw J and Eccles M, ‘The Potential Influence of Small Group Processes on Guideline Development’ (2001) 7 Journal of evaluation in clinical practice 165

‘Pallialine’ <http://www.pallialine.nl/> accessed 27 June 2014

Palmer AJ and others, ‘The CORE Diabetes Model: Projecting Long-Term Clinical Outcomes, Costs and Cost-Effectiveness of Interventions in Diabetes Mellitus (Types 1 and 2) to Support Clinical and Reimbursement Decision-Making’ (2004) 20 Suppl 1 Current Medical Research and Opinion S5

Papageorgiou AC and others, ‘Estimating the Prevalence of Low Back Pain in the General Population: Evidence From the South Manchester Back Pain Survey’ (1995) 20 Spine 1889

Pope C, ‘Resisting Evidence: The Study of Evidence-Based Medicine as a Contemporary Social Movement’ (2003) 7 Health: 267

Popper KR, *The Logic of Scientific Discovery* (Psychology Press 2002)

Powell WW and DiMaggio PJ, *The New Institutionalism in Organizational Analysis* (University of Chicago Press 1991)

Power M, *The Audit Society* (Oxford University Press 1997) <https://global.oup.com/academic/product/the-audit-society-9780198289470> accessed 27 May 2016

Price C, ‘NICE Risks Making Itself a “Laughing Stock” over Guidance on Metformin Alternatives, Say Experts’ (*Pulse Today*) <http://www.pulsetoday.co.uk/clinical/diabetes/nice-risks-making-itself-a-laughing-stock-over-guidance-on-metformin-alternatives-say-experts/20009139.article> accessed 5 May 2015

‘Published Clinical Guidelines’ <http://guidance.nice.org.uk/CG/Published> accessed 26 May 2014

Quine WV, ‘Main Trends in Recent Philosophy: Two Dogmas of Empiricism’ (1951) 60 The Philosophical Review 20

Raine R and others, ‘An Experimental Study of Determinants of Group Judgments in Clinical Guideline Development’ 364 The Lancet 429

‘Richtlijnen // Diliguide.Nl’ <http://www.diliguide.nl/> accessed 27 June 2014

‘Richtlijnendatabase’ <http://kims.orde.nl/het-kennisinstituut-voor-u/richtlijnen/richtlijnendatabase.html> accessed 11 June 2014

Rodwin MA, ‘Commentary: The Politics of Evidence-Based Medicine’ (2001) 26 Journal of Health Politics, Policy and Law 439

——, ‘Conflicts of Interest, Institutional Corruption, and Pharma: An Agenda for Reform’ (2012) 40 The Journal of Law, Medicine & Ethics 511

Rutten G and others, ‘NHG-Standaard Diabetes Mellitus Type 2 (Derde Herziening)’ (2013) 56 Huisarts Wet 512

Sackett DL and others, ‘Evidence Based Medicine: What It Is and What It Isn’t’ (1996) 312 BMJ 71

Savigny P and others, ‘Low Back Pain: Early Management of Persistent Non-Specific Low Back Pain’ (London: National Collaborating Centre for Primary Care and Royal College of General Practitioners 2009) <http://www.nice.org.uk/guidance/cg88/evidence/cg88-low-back-pain-full-guideline3> accessed 13 August 2015

Schott G and others, ‘Does the Pharmaceutical Industry Influence Guidelines?: Two Examples from Germany’ (2013) 110 Deutsches Ärzteblatt International 575

Shin JH, Haynes RB and Johnston ME, ‘Effect of Problem-Based, Self-Directed Undergraduate Education on Life-Long Learning.’ (1993) 148 CMAJ: Canadian Medical Association Journal 969

Steinbrook R and Lo B, ‘Medical Journals and Conflicts of Interest’ (2012) 40 The Journal of Law, Medicine & Ethics 488

Stelfox Henry Thomas, Choudhry Niteesh K and Allen S Detsky, 'Relationships between Authors of Clinical Practice Guidelines and the Pharmaceutical Industry' (2002) 287 JAMA 612

Stephenson J, 'Nurses Claim NICE Diabetes Guidance Will Put Patients at Risk' (*Nursing Times*, 15 September 2015) <https://www.nursingtimes.net/clinical-archive/diabetes/nurses-claim-nice-diabetes-guidance-will-put-patients-at-risk/5090329.article> accessed 17 October 2016

Stigler GJ, 'The Theory of Economic Regulation' (1971) 2 The Bell Journal of Economics and Management Science 3

Stone PW, Curran CR and Bakken S, 'Economic Evidence for Evidence-Based Practice' (2002) 34 Journal of Nursing Scholarship 277

Strathern M, 'The Tyranny of Transparency' (2000) 26 British educational research journal 309

Stratton IM and others, 'Association of Glycaemia with Macrovascular and Microvascular Complications of Type 2 Diabetes (UKPDS 35): Prospective Observational Study' (2000) 321 BMJ 405

Sur RL and Dahm P, 'History of Evidence-Based Medicine' (2011) 27 Indian Journal of Urology : IJU : Journal of the Urological Society of India 487

The Action to Control Cardiovascular Risk in Diabetes Study Group, 'Effects of Intensive Glucose Lowering in Type 2 Diabetes' (2008) 358 New England Journal of Medicine 2545

'The Guidelines Manual' <http://publications.nice.org.uk/the-guidelines-manual-pmg6> accessed 26 May 2014

Thomas KJ and others, *Longer Term Clinical and Economic Benefits of Offering Acupuncture Care to Patients with Chronic Low Back Pain* (Gray 2005) <http://www.carolinashealthcare.org/documents/NorthEastInternalIntegrative/research-library/BackAndNeckPain/LongerTermClinicalEconomicBenefit.pdf> accessed 8 October 2014

Thomson S and others, 'International Profiles of Health Care Systems, 2013' (Commonwealth Fund) <http://www.commonwealthfund.org/~/media/Files/Publications/Fund%20Report/2013/Nov/1717_Thomson_intl_profiles_hlt_care_sys_2013_v2.pdf> accessed 14 May 2014

Tsoukas H, ‘The Tyranny of Light: The Temptations and the Paradoxes of the Information Society’ (1997) 29 Futures 827

‘Type 2 Diabetes in Adults: Management | Guidance and Guidelines | NICE’ <https://www.nice.org.uk/guidance/ng28> accessed 17 October 2016

UK BEAM Trial Team, ‘United Kingdom Back Pain Exercise and Manipulation (UK BEAM) Randomised Trial: Effectiveness of Physical Treatments for Back Pain in Primary Care’ (2004) 329 BMJ 1377

van Herk R, Schepers R and Casparie AF, ‘Huisartsen en zelfregulering. De ontwikkeling van intercollegiale toesting en standaarden voor huisartsen tussen 1970 en 1990’ (1994) <https://lirias.kuleuven.be/handle/123456789/161090> accessed 12 June 2014

Vos C de, ‘Kwaliteit, disciplinering en sturing. Een historisch sociologisch onderzoek naar de vormgeving van kwaliteit van zorg in ziekenhuizen = Quality, discipline and governance. A historical sociological study into the design of quality of care in hospitals’ (Erasmus Universiteit Rotterdam 2014)

Walker BF, ‘The Prevalence of Low Back Pain: A Systematic Review of the Literature from 1966 to 1998’ (2000) 13 Journal of Spinal Disorders 205

Watine JC and Bunting PS, ‘Mass Colorectal Cancer Screening: Methodological Quality of Practice Guidelines Is Not Related to Their Content Validity’ (2008) 41 Clinical Biochemistry 459

‘Werkgroep Infectiepreventie (WIP)’ <http://www.rivm.nl/Onderwerpen/W/Werkgroep_Infectiepreventie_WIP> accessed 15 May 2014

Wilson N and others, ‘Governing Healthcare: Finding Meaning in a Clinical Practice Guideline for the Management of Non-Specific Low Back Pain’ (2014) 102 Social Science & Medicine 138

Wincott D, ‘European Political Development, Regulatory Governance, and the European Social Model: The Challenge of Substantive Legitimacy’ (2006) 12 European Law Journal 743

Worrall J, ‘What Evidence in Evidence-Based Medicine?’ (2002) 69 Philosophy of Science S316

Wright JSF, ‘The Regulatory State and the UK Labour Government’s Re-Regulation of Provision in the English National Health Service’ (2009) 3 Regulation & Governance 334

Yin RK, *Case Study Research: Design and Methods* (Sage 2014)

'ZonMw - A Randomised Cost-Effectiveness Study on the Surgical versus Conservative Treatment of Sciatica -' <http://www.zonmw.nl/nl/projecten/project-detail/a-randomised-cost-effectiveness-study-on-the-surgical-versus-conservative-treatment-of-sciatica/hoe-nu-verder/> accessed 19 June 2014

Zorginstituut Nederland, 'Kosteneffectiviteit in de Praktijk' <https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/rapporten-en-standpunten/2015/1506-kosteneffectiviteit-in-de-praktijk/1506-kosteneffectiviteit-in-de-praktijk/Kosteneffectiviteit+in+de+praktijk.pdf> accessed 16 April 2016

——, 'Kwaliteit' (*Quality*) <http://zorginstituutnederland.nl/kwaliteit> accessed 26 May 2014

'Zorgverzekeraars in Nederland' (*Health Insurers in the Netherlands*) <http://www.zorghulpatlas.nl/zorghulpwereld/zorgverzekering/zorgverzekeraars-in-nederland/> accessed 25 May 2014

Zuiderent-Jerak T and others, 'Variatie in richtlijnen. Wat is het probleem?' (*Variation in Guidelines. What is the problem? IBMG,* 2011) <http://www.pgosupport.nl/mailings/FILES/plugin_content/176/Regieraad_rapport_variatie_in_richtlijnen.pdf> accessed 12 November 2013