# Using Quality Indicators in Health Economic Evaluation to Establish the Value for Money of Quality Improvements

*Obstacles and Opportunities*

Vestergaard, Anne Sig

*Publication date:*
2017

*Document Version*
Publisher's PDF, also known as Version of record


# USING QUALITY INDICATORS IN HEALTH ECONOMIC EVALUATION TO ESTABLISH THE VALUE FOR MONEY OF QUALITY IMPROVEMENTS

- OBSTACLES AND OPPORTUNITIES

**BY**
**ANNE SIG VESTERGAARD**

DISSERTATION SUBMITTED 2017

AALBORG UNIVERSITY
DENMARK

# USING QUALITY INDICATORS IN HEALTH ECONOMIC EVALUATION TO ESTABLISH THE VALUE FOR MONEY OF QUALITY IMPROVEMENTS

## - OBSTACLES AND OPPORTUNITIES

by

Anne Sig Vestergaard

AALBORG UNIVERSITY
DENMARK

Dissertation submitted 2017

.

Dissertation submitted: May 12th 2017

PhD supervisor: Prof. Lars Holger Ehlers
Aalborg University

Assistant PhD supervisor: Prof. Torben Bjerregaard Larsen
Aalborg University

PhD committee: Professor Mogens Lytken Larsen
Aalborg Universitetshospital (formand)

Lektor Christian Kronborg
Syddansk Universitet

Professor Janne A. Martikainen
University of Eastern Finland

PhD Series: Faculty of Social Sciences, Aalborg University

ISSN (online): 2246-1256
ISBN (online): 978-87-7112-964-9

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk



Printed in Denmark by Rosendahls, 2017

Standard pages: 78 pages (2,400 characters incl. spaces).

# CV

In June 2013, Anne Sig Vestergaard graduated from Aalborg University with a master degree in Medicine with Industrial Specialization. Her interest in economic evaluation and optimization of the use of the scarce resources in the healthcare system was sparked during the master's degree program, Medical Market Access. Her master thesis revolved around guideline adherence in oral anticoagulant therapy of patients with atrial fibrillation, which inspired her further work in this clinical area. In November 2013, Anne was enrolled as a PhD fellow at the Danish Center for Healthcare Improvements.

Based on her interest in challenging the premise that 'new is always better,' Anne sought to evaluate opportunities for improving existing therapies through the use of economic methods, which became the focal point in this thesis. During her PhD enrollment, Anne has participated in and presented her research at international conferences, including the 10$^{th}$ iHEA World Congress, the 4$^{th}$ Nordic Conference on Research in Patient Safety and Quality in Healthcare, the 37$^{th}$ Nordic Health Economists' Study Group meeting, and the 19$^{th}$ Annual European ISPOR Congress. Anne has been affiliated with and participated in the research environment at the Aalborg Thrombosis Research Unit, Aalborg University Hospital where she has cooperated with leading researchers within atrial fibrillation.

During her enrollment, Anne has been both chairperson and accountant for DELPHI, the network for PhD fellows and scientific assistants at the Faculties of Social Sciences and Humanities.

# ACKNOWLEDGEMENTS

The studies that constitute the background for the present dissertation were designed and conducted during my time as a PhD fellow at the Danish Center for Healthcare Improvements. A lot of people have contributed, directly and indirectly, to the enablement of the present research project, and I owe great thanks to them all for making it possible.

I owe special thanks, first of all, to my supervisor Lars Holger Ehlers for his unceasing enthusiasm, his ability to see new correlations and methodological opportunities, and his ability to direct without leading. I am grateful for his support and our inspiring discussions, which have encouraged me to pursue new scientific areas; it has been a major contributory factor in my academic and professional development.

Furthermore, I would like to thank Torben Bjerregaard Larsen so very much for sharing his expertise within the clinical area of atrial fibrillation. In addition, I would like to thank Flemming Skjøth for his inputs, constructive criticism, and support; they have been invaluable for my understanding of the statistics. I value Torben and Flemming's continuous support and interest in my research and the opportunity they gave me to be a part of the dedicated environment at the Thrombosis Research Unit; thank you so much!

A special thank you goes to my good colleagues and friends at the Danish Center for Healthcare Improvement for our many good hours together, both academic and social. I am grateful for the inspiring environment that we have fostered – it has provided priceless support for my work on this thesis. One should not underestimate the importance of having somebody to call out to across the corridor and to drink coffee with.

Also, I would like to thank my family for their continuous interest in my work and patience; thank you for making me who I am today. Lastly, I would like to thank Kåre. Thank you for your everlasting patience when I start ranting; thank you for your keen mind and your genuine interest in my work; thank you for your continuous, kind support, and your ability to keep both my feet on the ground and focus on what is important. You're the best.

# ENGLISH SUMMARY

Quality improvement is increasingly used in the healthcare sector in an attempt to, amongst other things, contain the increasing resource use. Health economic evaluation informs on the efficient use of resources and hence could provide valuable information on the value for money of quality improvements. However, in quality improvement, quality indicators are often favored for assessment of the impact of interventions. As quality indicators do not necessarily represent an impact on health, per se, this constitutes an obstacle for the economic evaluation of quality improvements. In addition, there are appreciable differences in the aims and epistemologies of economic evaluation and quality improvement. Economic evaluation is intended to inform on the value for money of interventions and evidence on the value of interventions is often derived from evidence-based medicine. In contrast, quality improvements may target other aspects of healthcare than the efficient allocation of resources and evidence is acquired through a multiplicity of methods from different scientific disciplines. These discrepancies constitute substantial barriers to the application of economic evaluation of quality improvements. In consequence, establishment of their value for money may be hampered. The lack of knowledge on the cost-effectiveness of quality improvements may indirectly harm patients by causing opportunity costs – either because cost-effective quality improvements are not implemented or because cost-ineffective interventions are employed.

The present dissertation presents a contribution as to how quality indicators may be employed to estimate the value for money of quality improvements, when evidence on their impact on patient-relevant outcomes is not available. In a framework founded in Bayesian decision theory and value-of-information analysis, quality indicators may be introduced as intermediate links between interventions and patient-relevant outcomes, thereby enabling estimation of the cost-effectiveness of quality improvements. A set of requirements for quality indicators to be applicable in the context of economic evaluation is propounded. These lead to the presentation of a set of methodological considerations, which should be made when studies in quality improvement are designed and economic evaluation is projected.

The empirical case for the present dissertation and the appended papers is within the clinical field of cardiology, specifically on stroke prophylaxis in nonvalvular atrial fibrillation through the use of oral anticoagulant therapy. The focal point of the papers is to evaluate the health economic potential of alternative approaches to improving stroke prophylaxis for this patient population, not focusing on evaluation of one pharmacological treatment versus another.

# DANSK RESUME

Kvalitetsudvikling bliver i stigende grad brugt indenfor sundhedssektoren, blandt andet i et forsøg på at dæmme op for det stigende forbrug af ressourcer. Sundheds-økonomiske evalueringer bruges til at belyse, hvorledes ressourcerne kan anvendes mest efficient, så der opnås mest sundhed for pengene. De kunne derfor være en vigtig kilde til viden om den værdi for pengene, der kan opnås via kvalitetsudviklende tiltag. Kvalitetsindikatorer bruges ofte indenfor kvalitetsudvikling til at påvise effekten af interventioner, men da kvalitetsindikatorer i sig selv dog ikke nødvendigvis afspejler effekt på sundhed, udgør brugen af dem en hindring, når man ønsker at lave informative sundhedsøkonomiske analyser.

Ydermere er der betydelige, grundlæggende forskelle mellem sundhedsøkonomisk evaluering og kvalitetsudvikling; både hvad angår deres formål og epistemologiske baggrund. Formålet med sundhedsøkonomisk evaluering er som sagt at belyse den værdi for pengene, som interventioner leverer, og evidensen, der anvendes, stammer som oftest fra evidensbaseret medicin. I modsætning kan kvalitetsudvikling være rettet mod andre aspekter af sundhedspleje end den efficiente fordeling af ressourcer, og evidens for effekten af kvalitetsudviklende tiltag kan indhentes ved hjælp en mangfoldighed af metoder, der udspringer af flere forskellige videnskabelige discipliner. Disse uoverensstemmelser udgør en betydelig udfordring for sundhedsøkonomisk evaluering af kvalitetsudviklende tiltag og som konsekvens heraf, kan det være svært at vurdere deres værdi for pengene. Manglende viden omkring omkostningseffektiviteten af kvalitetsudviklende tiltag kan være til indirekte skade for patienter, idet det kan påføre alternativomkostninger – enten som følge af at omkostningseffektive interventioner ikke bliver implementeret, eller fordi omkostningsineffektive interventioner bliver anvendt.

Denne afhandling præsenterer et bidrag til, hvordan kvalitetsindikatorer kan anvendes til at estimere omkostningseffektiviteten af kvalitetsudviklende tiltag, når der ikke findes evidens for disses effekt som målt via patient-relevante udfald. Med udgangspunkt i en Bayesiansk beslutningsteoretisk og value-of-information analytisk tilgang, vil kvalitetsindikatorer kunne introduceres som intermediære led i sammenhængen mellem interventioner of patient-relevante udfald. Dette vil efterfølgende kunne gøre det muligt at estimere omkostningseffektiviteten af kvalitetsudviklende interventioner. I afhandlingen opstilles en række betingelser til kvalitetsindikatorer, som skal opfyldes, for at disse vil være anvendelige som inputs i sundhedsøkonomiske evalueringer. Opstillingen af disse betingelser leder ydermere til en række metodologiske overvejelser, som bør tage i betragtning i forbindelse med

designet af studier af kvalitetsudvikling, såfremt det påtænkes at lave efterfølgende sundhedsøkonomiske evalueringer.

Afhandlingen og de tilhørende studier tager empirisk udgangspunkt indenfor det kliniske område kardiologi, nærmere bestemt i den forebyggende behandling af slagtilfælde hos patienter med nonvalvulær atrieflimren ved hjælp af orale antikoagulantia. Omdrejningspunktet for studierne er at evaluere det sundhedsøkonomiske potentiale ved alternative tilgange til forbedring af den forebyggende behandling af slagtilfælde for denne patientgruppe, hvor der ikke fokuseres på sammenligningen af et medikament overfor et andet.

# ABBREVIATIONS

| | |
|---|---|
| **AF** | Atrial fibrillation, nonvalvular |
| **CBA** | Cost-benefit analysis |
| **CEA** | Cost-effectiveness analysis |
| **$CHA_2DS_2$-VASc** | Congestive heart failure, hypertension, age≥75 [doubled], diabetes, stroke [doubled], vascular disease, age 65–74, and sex category (female) |
| **CUA** | Cost-utility analysis |
| **DAM** | Decision-analytic modeling |
| **DKK** | Danish kroner |
| **EBM** | Evidence-based medicine |
| **ESC** | European Society of Cardiology |
| **ICER** | Incremental cost-effectiveness ratio |
| **ICH** | Intracranial hemorrhage |
| **INR** | International normalized ratio |
| **IOM** | Institute of Medicine |
| **HAS-BLED** | Hypertension, abnormal hepatic and/or renal function, stroke, bleeding, labile INR; TTR<60%, elderly; age>65, drugs; influencing coagulation and/or alcohol abuse |
| **HRQoL** | Health-related quality of life |
| **MB** | Major bleeding |
| **NB** | Net benefit |
| **NOAC** | Non-vitamin K oral antagonist |
| **OAC** | Oral anticoagulant therapy |
| **VOI** | Value of information |
| **PSA** | Probabilistic sensitivity analysis |
| **QALY** | Quality-adjusted life year |
| **QI** | Quality indicator |
| **RADS** | Rådet for Anvendelse af Dyr Sygehusmedicin |
| **RCT** | Randomized controlled trial |
| **SSE** | Stroke/systemic embolism |
| **TTR** | Time spent in therapeutic range, percent |
| **WHO** | World Health Organization |

# LIST OF INCLUDED PUBLICATIONS

**Paper I:** **What is the value of quality indicators for informed decision-making in healthcare? A Bayesian decision theoretical and value of information analysis perspective**

AS Vestergaard, LH Ehlers
*Submitted*[1]

**Paper II:** **Effect of Anticoagulation on Hospitalization Costs After Intracranial Hemorrhage in Atrial Fibrillation – A Registry Study**

AS Vestergaard, F Skjøth, GYH Lip, TB Larsen
*Stroke (2016) 47(4):979-85*[2]

**Paper III:** **A Health Economic Evaluation of Stroke Prevention in Atrial Fibrillation: Guideline Adherence Versus the Observed Treatment Strategy Prior to 2012 in Denmark**

AS Vestergaard, LH Ehlers
*PharmacoEconomics (2015) 33:967–79*[3]

**Paper IV:** **The importance of mean time in therapeutic range for complication rates in warfarin therapy of atrial fibrillation: a meta-regression analysis**

AS Vestergaard, F Skjøth, TB Larsen, LH Ehlers
*Submitted*[4]

## RELATED ABSTRACTS BY THE AUTHOR

- **Health economic consequences of resuming anticoagulation after intracranial hemorrhage in patients with atrial fibrillation (80)**

  AS Vestergaard, F Skjøth, GYH Lip, TB Larsen
  *4th Nordic Conference on Research in Patient Safety and Quality in Healthcare 2016. University of Eastern Finland, 2016. 80.*
  *ISBN 978-952-61-2129-1.*

- **Application of quality indicators to enable economic evaluation of quality improvements – a Bayesian decision theoretical and value of information analytical framework (PRM 219)**

  AS Vestergaard, LH Ehlers
  *Value in Health (2016) 19 (7):A397.*

# TABLE OF CONTENTS

# TABLE OF FIGURES

Unless stated otherwise, figures are produced by the author.

# CHAPTER 1.
# INTRODUCTION

## 1.1. THE INSATIABLE HEALTHCARE MONSTER

In the continuous pursuit of improving the health of patients, innovative healthcare solutions are essential. Innovation in healthcare and healthcare technology incessantly increases the range of effective treatment options available for alleviation of the ailments of patients. Amongst other due to the progress in medical technology, massive changes and advancements have occurred with respect to health and healthcare within the last 150 years. Since the 1900s, infectious and parasitic diseases have no longer been the leading causes of morbidity and mortality worldwide and, particularly since the 1950s, longevity has increased in developed and developing countries alike. Instead, the prevalence and incidence of chronic diseases, such as coronary heart disease, type 2 diabetes, cancer, and stroke, have increased and now constitute the leading causes of morbidity and mortality. The increase in chronic diseases has occurred in concert with a change in the world demography, as the population ages.[5] Though enhanced longevity is inherently desirable, the concomitant epidemiological transition places healthcare systems and economies under substantial pressure. In Denmark, it has been estimated that approximately two thirds of the population suffer from one or more chronic diseases[6], thereby emphasizing the great public health challenge chronic conditions pose today. Improved knowledge on disease progression and how it may be abated has increased the call for, and availability of, treatments for chronic conditions, both prophylactically and to diminish the risk of late complications of diseases[5].

Within the last decade, healthcare costs have increased globally as well as nationally[7]. In 2014, the total expenditure on health in Denmark accounted for 10.6 percent of the gross domestic product. The 2014 expenditure of publicly funded healthcare alone added up to more than 170 billion Danish kroner (DKK) (2017 value) in Denmark, accounting for approximately 30 percent of the total public expenditure.[7,8] If the current developments continue, projections suggest that the healthcare expenditures will have increased by 47 percent by 2060[9]. As with other costs in the healthcare system, the costs of medications inside and outside hospitals have increased over time; in 2014, the cost of all medications amounted to more than 13 billion DKK[10]. In Denmark, a Beveridge model for organization of the healthcare system has been adopted, in which the majority of the integrated healthcare

is completely financed by taxes, with the exception of some costs related to dental treatment and medications. Thus, the healthcare authorities are the single payer and are responsible for the supply of healthcare and the containment of the costs related to it.[11]

The progress in new medical technology permits delivery of improved healthcare of both acute and chronic conditions, but usually it comes at an additional cost. The increased longevity of population puts healthcare systems and economies under pressure and the use of relatively costly new treatment for chronic conditions only exacerbates the scarcity of resources.[9,12–14] The availability of an arsenal of effective though costly treatments makes demands on decision-makers when budgetary constraints preclude the utilization of all potential treatments. The need for healthcare never ceases. Unless further resources are added, which would only transfer the pressure to other public sectors, the scarcity of resources within the healthcare system represents an impediment for the sustained utilization of clinically optimal therapies, if they come at too high a cost. For this reason, prioritization is expedient[11,13]. Prior to policy decisions on which interventions should be approved for utilization in a publicly funded healthcare system, such as the Danish one, it is highly relevant to acknowledge and evaluate the expected value for money of interventions to enable informed decision-making on the consequences of the allocation of the scarce resources. Economic evaluation, identifying the value for money of medical technology, is a requirement for reimbursement decisions to be made in many countries with publicly funded healthcare systems[13,14]. In Denmark, attention to the cost-effectiveness of interventions has increased of late and different steps are taken to contain the costs and reduce the escalation of the expenditure[9,15].

The pursuit of constraining healthcare costs, while still providing high-quality healthcare, has stimulated alternative approaches to healthcare optimization, aside from the utilization of more effective and safe medical technology. These approaches include quality improvement within existing healthcare[16–18].

## 1.2. QUALITY IMPROVEMENT – A KEY TO COST CONTAINMENT?

Since the turn of the millennium, the focus in healthcare research has increasingly been turned towards the possibility of achieving increased patient safety, improved therapy, and cost savings through quality improvements within existing healthcare. The increased interest in quality of care has been spurred by changes in public healthcare systems and their reimbursement systems that might affect the quality of

care.[18–20] Consequently, quality in healthcare and quality improvement have been a part of the official, political agenda both nationally and internationally since approximately 2000. In Denmark, national strategies for quality improvement in the healthcare system are therefore continuously presented that include explicit quality aims dictating the trends in quality improvement work and research. Prior to the introduction of centralized management of quality improvement in Denmark, quality improvement was managed in a decentralized way by dedicated healthcare personnel.[19,21,22]

One of the main purposes of assessing quality and quality improvement has been to support administrative and social policy decisions, by indicating which initiatives might produce beneficial results[23,24], and the methods for evaluating quality improvements have been developed and utilized with this in mind. Assessment of the quality of healthcare is increasingly important to providers, purchasers, and regulators of healthcare, as budgetary constraints enforce considerations of the expected benefit of interventions and their subsequent value for money[20]. Consequently, quality improvement may in the future be subjected to the same requirements for evidence of cost-effectiveness as those for medical technology. Explicit evaluation of the cost-effectiveness of interventions has, however, not traditionally been a part of the evaluation of quality improvements and the methods for economic evaluation of quality improvement remain mostly unexplored[16,18,25]. Discrepancies in the currently used methods for economic evaluation and the very methodology of quality improvement in healthcare challenges the design of economic evaluation within this field of research. Incongruence in the scientific foundation may cause problems as health economic methods have been developed to address other challenges than those that present themselves in research on quality improvement[24,26]. For instance, a prevalent and increasing challenge is the use of quality indicators (QIs) for assessment of effect of quality improvements[20,27]. These effect measures often do not reflect impact on health, per se, which represents an obstacle to economic evaluation of the value for money of quality improvements. This may complicate informed decision-making within healthcare if the expected value for money of quality improvements cannot be established. It may ultimately cause suboptimal policy decisions leading to opportunity costs if inferior interventions are opted for due to a lack of knowledge of their comparative cost-effectiveness.

## 1.3. OPTIMIZING ORAL ANTICOAGULANT THERAPY IN ATRIAL FIBRILLATION

Within the last decade, several advancements have occurred with respect to the treatment of patients with the chronic, cardiac arrhythmia, nonvalvular atrial fibrillation (i.e. excluding valvular atrial fibrillation, hereafter only referred to as AF). Greater understanding of the disease itself and the availability of new therapies have led to the rapid development of clinical practice guidelines on clinically appropriate treatment of this condition. The guidelines recommend that the majority of patients with AF should be treated prophylactically with continuous oral anticoagulant therapy (OAC) to avoid stroke and other thromboembolic complications, which the condition predisposes to.[28,29] Sustained thromboprophylaxis through the use of OAC for eligible patients with AF constitutes the clinical focus for the present dissertation.

In Denmark, thromboprophylaxis through the use of OAC is mainly achieved by the use of the vitamin K antagonist warfarin and the newer OAC agents, non-vitamin K oral anticoagulants (NOACs)[29–31]. The emergence of the NOACs for stroke prophylaxis in AF is assumed to be the triggering factor for a raised awareness of the clinically appropriate treatment of this patient population and for an increased spread of appropriate treatment. Both in Denmark and internationally, more patients with AF are now treated in accordance with clinical practice guidelines on stroke prophylaxis in AF than was observed just four years ago. This has also entailed an increased use of NOACs for stroke prophylaxis in AF and a stagnation in the utilization of warfarin therapy, which has otherwise been used for stroke prophylaxis for decades.[29,32,33]

A main difference between the agents is the level of monitoring that is necessary to ensure that the effectiveness and safety of the therapies are maintained. Whereas warfarin therapy necessitates relatively frequent monitoring, which may be a nuisance to patients and represent an additional cost of therapy delivery, this is not necessary for the NOACs[34]. Though increased use of OAC, and especially the NOACs, for stroke prophylaxis in AF represents a clinical improvement of the therapy of patients with AF, it also increases the current pressure on healthcare budgets, as the purchase cost of NOACs is substantially higher than that of warfarin[35].

Within the last decade, the comparative cost-effectiveness of the NOACs versus conventionally managed warfarin therapy has been investigated thoroughly[36–38]. These analyses have typically investigated the cost-effectiveness of substituting one therapy with another, based on evidence from large randomized controlled trials (RCTs)[39–42]. However, other initiatives in OAC exist that likewise may improve the health of the AF patient population as a whole. For instance, adherence to current

clinical practice guidelines may be improved, ensuring a larger proportion of patients receive clinically appropriate therapy. Quality improvements, such as improving guideline adherence, may represent important approaches when the aim is to improve the health of patients.

# CHAPTER 2.
# HYPOTHESIS AND THESIS AIMS

The basic hypothesis in the present dissertation is that economic evaluation may represent a useful tool for establishing the value for money of quality improvements in healthcare and that, though economic evaluation cannot encompass all aspects relevant to informed policymaking, it constitutes a rational premise for the decision-making to be sound. This dissertation presents a contribution to exploring that hypothesis within the empirical field of research on OAC stroke prophylaxis in AF. The present thesis thus intends to:

1) INVESTIGATE OPPORTUNITIES FOR IMPROVING ORAL ANTICOAGULANT THERAPY OF PATIENTS WITH ATRIAL FIBRILLATION

and

2) INVESTIGATE THE OPPORTUNITIES AND OBSTACLES ASSOCIATED WITH ECONOMIC EVALUATION OF QUALITY IMPROVEMENTS WHEN PATIENT-RELEVANT OUTCOMES ARE NOT AVAILABLE AND PRESENT A CONTRIBUTION AS TO HOW THE OBSTACLES MAY BE OVERCOME

Quality in healthcare may be increased by a vast number of methods and knowledge of how their health economic impact should be estimated is currently incomplete. The clinical aim of the present dissemination is to seek cost-effective approaches to optimize therapy for this patient population, not necessarily focusing the use of new medical technology.

In the present dissertation, the focus is on static, comparative evaluation of quality improvements. As such, the approach to economic evaluation is based on the available evidence on the impact of quality improvements without projections of potential changes in the impact over time. The impact of quality improvements is expected to be constant over time. Unsubstantiated assumptions on potential increase or decrease, as would likely be observed in real life, are not included. This is consistent with the methods for economic evaluation of medical technology, ensuring comparability of the expected cost-effectiveness of interventions irrespective of the type of intervention under investigation. Increasing knowledge in this area may have social and scientific bearing as it may aid informed policymaking in healthcare. The appended papers (Appendix A–D) are intended to support the exploration of the hypothesis and aims

of the present dissertation. They are presented as illustrative examples of how evaluation of potential quality improvements may be enabled and should not be considered exhaustive.

Chapter 3 of this dissertation presents an account of the purpose of the healthcare system, which both the methods of health economics and quality improvement in healthcare are intended to support. The chapter therefore includes a brief introduction to the foundations and epistemological background of the two scientific fields of quality improvement and health economics, with a specific focus on the methods for economic evaluation, to illustrate the discrepancies and similarities between them. Acknowledgement of these is pivotal to the understanding of the opportunities and obstacles that exist for the application of economic evaluation to quality improvements. The chapter aims to provide an overview of the dominant theories and premises that form the foundation for the methods for economic evaluation and modern quality improvement. It is not intended to cover all aspects and methodologies of either health economics or quality evaluation and improvement.

Chapter 4 includes an introduction to the clinical field of AF and research on stroke prophylaxis in AF through the use of OAC. The chapter aims to highlight the necessity and challenges of OAC in AF, the current developments in OAC stroke prophylaxis, the challenges in procuring sufficient evidence on the effectiveness and safety of agents, and potential focus areas for quality improvement. The prevalence of AF is relatively low and, luckily, complications occur only relatively infrequently. Consequently, the effectiveness and safety of interventions to stroke prophylaxis is sometimes evaluated through the use of QIs. OAC in AF thus represents a relevant empirical case for the present thesis.

Chapter 5 includes summaries and discussions of the individual papers included in the thesis. It also includes a contextualization of them within the frame story of the dissertation.

Lastly, Chapter 6 discusses the potentials of applying economic evaluation to the area of quality improvement in healthcare, as supported by the papers and the dissertation itself. It includes suggestions of a set of considerations that should be included in future work to potentially enable the application of economic evaluation to quality improvement, when patient-relevant outcomes are not available.

# CHAPTER 3.
# METHODOLOGY

## 3.1. PURPOSE OF THE HEALTHCARE SYSTEM

The main objective of the Danish healthcare system is explicated in the first section of the Danish Health Care Act, which states that "*The healthcare system aims to increase the health of the population and to prevent and treat illness, suffering, and infirmity for the individual*" [author's translation][43]. Thus, the main purpose of the healthcare system and therefore implicitly also interventions provided by the healthcare system, is to bring relief of pain and illness in order to improve health to the benefit of individuals and the society as a whole[26,43]. As such, the aim of healthcare is not to increase the products of the healthcare system, such as surgery, bed days, or medication, etc. These entities do not bring patients utility, per se, but constitute the means to an end, i.e. the obtainment of health. [11] The World Health Organization (WHO) has defined health as "*a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity*"[44]. According to this definition, health comprises more aspects than just the physical health of individuals, interpreted as their bodily functioning. This interpretation of health is pivotal when evaluating the effectiveness of interventions that are aimed at maximizing the health of patients.

Policymakers within the healthcare system direct the focus of healthcare and decide its purpose, as explicated in the Danish Health Care Act[43]. There is a widespread assumption that decision-makers intend to maximize the health of the population, but they must do so within existing budgetary constraints[11,14]. Accepting the premise that the purpose of the healthcare system is to maximize health outcomes and that this must be done with only a certain amount of available resources, the average expected health benefits and additional costs related to actions within the healthcare system should be decision-makers' primary concern[13].

## 3.2. FOUNDATIONS OF HEALTH ECONOMICS

Health economics concerns the application of economic methods within the area of health and healthcare. As the field of the discipline, i.e. health and healthcare, is not embedded within the scientific field of normative economics, health economics comprises elements from various scientific disciplines, including social, health, and natural sciences. Consequently, application of health economic methods often requires a multidisciplinary approach that appreciates and incorporates the complexity of the functioning of the healthcare system to generate analyses that may substantiate sound policy decisions.[11]

Identified as a subdiscipline within economics, specialized in health and healthcare delivery, health economics is founded in microeconomics and welfare economics. As such, the methods used in health economics adopt the expectation that resources, both monetary and nonmonetary, are scarce and must be allocated to competing purposes. This is on a par with the observed reality in healthcare. Furthermore, it is expected that all decision-making concerning health and healthcare is based on rationality and an inherent desire for maximization of 'benefit'. This entails all players within the healthcare system, including patients, clinicians, and policymakers, choosing the courses of actions that provide them with the largest maximand and being able to rank, for instance, interventions based on their desirability under budgetary restrictions. It follows that policymakers who make decisions on the functioning of the healthcare system and utilization of its scarce resources are expected to be rational and have an inherent desire to maximize the 'benefit' for the society that they represent.[11,45]

Welfare economics constitutes the normative basis for health economics and provides the value-based, theoretical framework in which economic evaluation is performed. Within normative health economics, the objective of actions, e.g. policy decisions, is predefined via the philosophical foundation of welfare economics. The actions are valuated against that objective. The philosophical foundation and ethical theory of welfare economics are often recognized to be, basically, utilitarianism. In utilitarianism, the ethical optimum is the situation that generates the total maximum amount of utility of all individuals in a society, i.e. the greatest social welfare[46]. Roughly described, "*Utility is a measure of* [the individual's, ed. by author] *level of satisfaction with various combinations of consumer goods. It includes a market basket filled with a combination of housing, food, transportation, and so on, with perhaps many types of each*"[46]. Each individual has an index of preferences towards a multitude of 'goods' or states of the world, which may change depending on the conditions that he or she experiences. An individual's preferences towards the goods or states of the world and ability to weigh their importance against each other generate

a utility function, where some elements are given a higher preference than others. When an individual is able to attain a state of the world that he/she has a higher preference for than his/her prior state, this indicates that the individual has achieved a higher level of utility. As preferences for states of the world are individually decided, the utility that may be expected from different states is not transferable between individuals. Accordingly, in the welfarist evaluation only individual utility can be used as an outcome for evaluation. The aggregate of all individuals' experienced utility within a society constitutes the level of social welfare. Therefore, if an individual achieves a state of the world that he or she prefers compared to the prior state this will, consequently, also increase the total level of social welfare.[46,47]

Adopting a utilitarian ethical foundation, the dominant framework in welfare economics for the assessment of, for example, interventions has adopted four explicit, normative key tenets presented in Box 1. These comprise the cornerstones of the framework for welfare economist evaluation.[11,46,47]

I. *The utility principle.* The expectation that individuals are able to rationally order options based on their expected benefit and will maximize their own welfare accordingly by choosing the preferred option. Individuals' utility is unaffected by what happens to other individuals.

II. *Individual sovereignty*. Individuals are the best judges as to what brings them the greatest utility. Individuals' preferences towards different 'goods' differ. Utility is contingent on the individuals' preferences towards the 'goods'. Consequently, it is not possible to compare or aggregate utility across individuals.

III. *Consequentialism*. Welfare is derived from the final result of actions and processes and not the processes themselves.

IV. *Welfarism*. The benefit of a situation should be evaluated based on the level of social welfare it provides.

*Box 1. Normative key tenets for the welfare economist framework for evaluation*[47].

Welfare economics is used in the pursuit of maximizing total social welfare. It then follows that the immediate purpose of welfare economics is to maximize benefit within budgetary constraints, i.e. to identify the optimal resource allocation that ensures the lowest cost per unit of the maximand. [11,47] The optimum, the Pareto-efficient situation is achieved when the maximand is the highest possible and when no reallocation of resources can be made where the benefit of one individual is optimized without making that of another individual worse. Traditionally, it is assumed that Pareto efficiency, also known as optimality, may be achieved in a perfectly competitive market. The importance of this assumption in welfare

economics is emphasized by the reference to it as 'the first fundamental theorem of welfare economics'[46]. However, the requirements that substantiate the competitive model, and to some degree the very tenets that substantiate the framework for welfare economic evaluation, are not observed in the healthcare system.[11,46] This has led to criticism of the welfare economist approach to the evaluation of healthcare.

A particularly influential critique of the welfare economist approach to the evaluation of healthcare has been put forward by Sen, who criticized the very foundation of welfare economist evaluation by calling into question the use of individual utility as the sole outcome in the evaluation of healthcare. By definition, utility reflects an individual's emotional reaction, e.g. satisfaction (Box 1), to the possession of goods and capabilities[46]. However Sen argues that the usefulness, i.e. the enablement and potential capabilities, that 'goods' or states of the world would impart to individuals is of higher relevance as an outcome for measurement[47]. In traditional welfare economics, health is considered a commodity equal to any other goods that may be valuated, purchased, and consumed. It could, however, be argued that sound, or at least adequate, health is a prerequisite for the consumption of other goods. Thus, health holds an 'extra' value, which distinguishes it from other commodities. If individuals do not possess a certain level of health, they may not be able to value or utilize other goods, for which reason health and consequently healthcare provision as a means for obtaining health is a necessity rather than a normal good. Furthermore, the health of the population is also meritorious for the maintenance of a functioning workforce and the productivity in society. Therefore, health represents an entity of such high importance for the very functioning of society that it could be argued that health and other commodities should not be given the same weight.[46,47] This is also used as argument for the view that the maintenance of the health of the population should be subsidized by the state, e.g. by adopting the Beveridge model for the organization of the healthcare system, as done in Denmark[11].

The issues with the welfare economist approach to evaluation in healthcare has spurred the development of extra-welfarism[11]. The extra-welfarist approach is not exclusively used in healthcare[47], but in the present dissertation it is discussed only in this context. In recognition of the fact that market forces fail in the healthcare system, these are disregarded in the extra-welfarist approach to evaluation[11]. Nonetheless, extra-welfarism adopts the same main objective for actions in healthcare as welfare economics, i.e. the pursuit of increasing a maximand within budgetary constraints. One of the principal differences between welfare economics and extra-welfarism is the notion of what should be maximized.

In the extra-welfarist approach to evaluation, observance of the normative key tenets comprising the cornerstones for the framework for welfare economist evaluation is relaxed (Box 1). To accommodate, amongst others, the criticism expressed by Sen, under extra-welfarism, outcomes other than individual utility may be used in evaluations[45,47]. In the most pragmatic approach to extra-welfarist evaluation, health constitutes the maximand and distribuendum of the healthcare system, which is in immediate agreement with the primary objective of the healthcare system, i.e. health maximization[13,43,45]. This implies that under extra-welfarism, health is considered to have intrinsic value, i.e. the value is not determined by the potential utility that it may entail for the individual. For this reason, health and elements that affect health status are often used as maximand in extra-welfarist evaluation.[13,43,47] In agreement with the welfare economist approach (cf. tenet III, Box 1), under extra-welfarism, the value of interventions is only determined by their expected final impact on health and not by the process, i.e. intervention such as surgery or medical treatment through which health maximization is achieved. This is consistent with the expectation that policymakers and patients do not seek medical care per se, but the health benefit that it conveys[11].

Accepting health and elements that affect health as maximand for evaluation under extra-welfarism challenges the expectation of individual sovereignty in the welfare economist framework for evaluation (Box 1). Due to incomplete knowledge of the expected health value of different healthcare interventions, the individual may not be the best judge of how to achieve the highest level of health and may consequently not be able to give rational preferences to different healthcare interventions. Furthermore, under extra-welfarism, preferences for, for example, different health statuses may be derived from sources other than the individual who experiences the health state, such as samples representing the general population. The use of societal preferences towards health statuses can be justified by considering that decision-makers in publicly funded healthcare systems should safeguard their ‘stakeholders’’ interests; i.e. the society that subsidizes the healthcare system. Hence, it follows that by valuating outcomes by societal preferences, consistency between payer and payer preference is achieved. Application of pre-specified, transferable preferences for health states facilitates interpersonal comparability of outcomes. This enables comparison of the expected benefit of changes in health states irrespective of the individual achieving it, which would not be permitted under the welfare economist approach.[45,47]

Despite the discrepancies between the welfare economist and extra-welfarist approaches to evaluation, extra-welfarism is still considered a subspecies of normative welfare economics[47]. Although the primary aim of the healthcare system is to improve health when possible, there are additional social and societal concerns that

are not explicitly incorporated in the pursuit of achieving Pareto efficiency, i.e. either under the welfare economist or the extra-welfarist approach[43,47]. This includes concerns of, for example, equal access to healthcare and the provision of high-quality care for all individuals. Sen directed a sharp critique of considering the Pareto optimum as a representation of the optimal allocation of resources, as Pareto optimality may be achieved in situations of high inequity and inequality, where resource allocation is highly skewed with some individuals receiving all resources and others none. By all other ethical standards, this would represent a suboptimal resource distribution.[47] Equity and equality concerns are explicitly incorporated into the stated aim of the healthcare system, i.e. in the Danish Health Care Act[43], which causes the rejection of a strict welfare economist approach to resource allocation in real-life policymaking.[47]

Nonetheless, evaluation methods based on the extra-welfarist approach to evaluation, such as economic evaluation, are suitable for informing decisions in healthcare due to the common foundation and purpose of the healthcare system and the aim of extra-welfarism. With its foundation in normative welfare economics, economic evaluation can be used to inform on – if not ensure – efficient use of resources.[13]

## 3.3. ECONOMIC EVALUATION AS A DECISION AID FOR HEALTHCARE PRIORITIZATION

The purpose of economic evaluation is to inform on whether the amount of extra health benefit that healthcare interventions provide justifies the additional costs related to the interventions.[13] The existence of budgetary constraints and the need to contain costs emphasize the prudence of applying these methods to inform policy decisions in healthcare.

In full economic evaluation, both the costs and consequences of interventions are considered to enable evaluation of the expected, i.e. average, incremental cost related to acquiring an extra unit of the effect measure, i.e. health benefit related to the interventions. Usually, the results of economic evaluations are presented via the incremental cost-effectiveness ratio (ICER)[13]:

$$\frac{Cost_A - Cost_B}{Effect_A - Effect_B} = \frac{\Delta C}{\Delta E} = ICER =$$

$$incremental\ cost\ per\ incremental\ unit\ of\ effect, i.e.\ health\ benefit$$

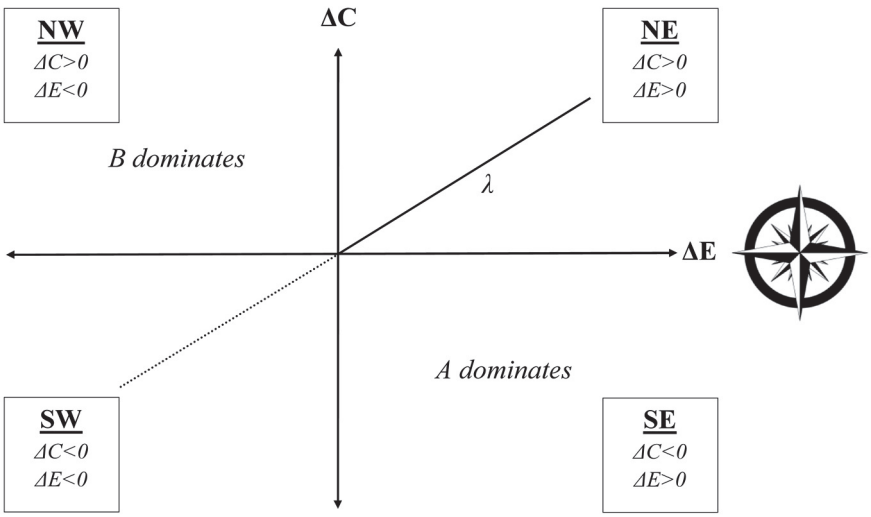



*Figure 3-1. The incremental cost-effectiveness plane given intervention A vs. intervention B. λ represents the threshold for opportunity costs. ΔC: incremental cost, ΔE: incremental effect.*

There is an inherent expectation in the evaluation tradition that new medical technology (designated A in the equation and in the incremental cost-effectiveness plane in Figure 3-1) provides a higher level of health benefit than the existing technology (designated B in the equation and in the incremental cost-effectiveness plane). If a course of actions, intervention A, provides positive incremental effects and negative incremental costs compared to intervention B, it dominates its comparator and is situated in the southeast quadrant of the incremental cost-effectiveness plane in Figure 3-1, and vice versa for the northwest quadrant. If both the incremental costs and effects of intervention A versus B are eitherpositive or negative, the ICER is situated in either the northeast or southwest quadrant, respectively. In that case, the ICER may be compared to a cost-effectiveness threshold, i.e. a threshold for opportunity costs ($\lambda$) (Figure 3-1). In evaluations where dominance is not observed a threshold for opportunity costs is necessary to enable the interpretation and establishment of cost-effectiveness of the investigated interventions. Otherwise, the economic evaluation only provides information on the expected costs and effects of interventions, and cannot inform on the comparative value for money of the interventions.[13]

The threshold for opportunity costs reflects the monetary value that society attaches to one unit of the chosen effect measure and is specific for that effect measure. The threshold represents the shadow price of the budget constraint, that is, the benefits and

costs alike that are displaced if the intervention under investigation is adopted. For instance, for the effect measure quality-adjusted life year (QALY), the National Institute for Health and Care Excellence under the National Health Service in the United Kingdom has stated a threshold of £20,000–£30,000 per QALY[48]. Hence, for each £20,000 that are spent on healthcare resources in one clinical area, one QALY is displaced elsewhere in the healthcare system if no further resources are brought into the healthcare budget. This implies that if the ICER of intervention A vs. intervention B is above the threshold (Figure 3-1), the additional benefit provided by intervention A is smaller than the benefit that is displaced elsewhere in the healthcare system. This causes inefficient allocation of resources. Likewise, if the ICER related to the intervention under investigation is below the threshold, more than one unit of effect, e.g. QALY, can be achieved at the threshold value.[13]

Opportunity costs will always occur under budgetary constraints, as the allocation of resources to one area obstructs the use of the same resources in another area. What is of interest is whether the benefit achieved by an allocation supersedes the benefit that could have been achieved elsewhere. The extra-welfarist decision rule of cost-effectiveness provides that if the ICER is lower than the threshold, the course of action under investigation is considered cost-effective and should be accepted, whereas if it supersedes $\lambda$, it is considered cost-ineffective and should be rejected. Thus, the threshold reflects the maximum that the healthcare system, i.e. society, should be willing to pay for further benefit in order to avoid inefficient resource allocation.[13,14]

Despite its widespread use, the ICER framework suffers from deficiencies. It is not suitable for when incremental effectiveness approximates zero nor is it possible to generate apprehensible ICER intervals to evaluate possible dispersion of cost-effectiveness estimates[13,49]. However, when a threshold for opportunity costs exists, the ICER and the decision rule for cost-effectiveness may be rearranged as:

$$\frac{Cost_A - Cost_B}{Effect_A - Effect_B} < \lambda \Leftrightarrow \frac{\Delta C}{\Delta E} < \lambda \Leftrightarrow \Delta C < \lambda \Delta E \Leftrightarrow$$
$$\lambda \Delta E - \Delta C > 0.$$

This is known as the net benefit (NB) framework, in which $\lambda \Delta E - \Delta C$ is the incremental NB. If interventions are cost-effective, based on an established $\lambda$, the incremental NB is positive. The NB framework amongst other things, allows for evaluation of the uncertainty surrounding the cost-effectiveness of interventions, as it may be described via a distribution.[13,49,50]

If applying a strict extra-welfarist approach, the decision rule means that if a course of actions provides a negative incremental NB, it should not be adopted[11,46]. The threshold introduces a social commitment where all resources cannot be allocated to a particular intervention if the opportunity costs are considered too substantial. Economic evaluation may be used to illustrate the potential opportunity costs related to the introduction of interventions and help making prioritization explicit regarding the acceptance or rejection of interventions based on their expected costs, health benefits, and potential opportunity costs. The full economic evaluations include the cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), and cost-utility analysis (CUA), which is sometimes recognized as a subspecies of CEA. Common among the evaluation types are the methods for accounting costs, which in all analyses are represented in monetary units. What differs is the chosen measure of effect.[13,14]

### 3.3.1. OUTCOMES IN ECONOMIC EVALUATION

The effect measures used in extra-welfarist evaluation usually include indicators of well-being and measures of health as these reflect the main priority of policymakers within healthcare, whom the analyses are intended to inform.[13,43,47]

CBA is closer associated with traditional welfare economics than CEA and CUA as the main purpose of CBA is to identify interventions that increase welfare, and not specifically health. Thus, in CBA health is not employed as the sole maximand. In CBA, health benefits are considered a commodity, whose value may be estimated based on the equivalent amount of consumption of other goods that individuals might otherwise have. Hence, the health benefits of interventions are established in monetary units, based on individuals' preferences towards them, i.e. the monetary value that each individual attaches to the benefit. Commoditizing and monetizing health and health gain enables comparison of the benefits of interventions across different sectors. Consequently, the budgetary constraints that would normally apply to consumption of resources within the healthcare system are relaxed and CBA can be used to identify the allocation of resources that provides the greatest amount of total social welfare irrespective of sector origin.[11,13] There are, however, challenges involved in asking individuals to state their preferences towards, commoditizing, and monetizing health and interventions to increase health[11,13,46]. These challenges compose substantial obstacles for the application of CBA to health and healthcare. CBA remains the least used type of economic evaluation in healthcare[11,51].

In contrast to CBA, the methods of CEA and CUA have a stronger foundation in extra-welfarism, recognizing of the failure of the market forces in the healthcare system. In CEA and CUA the budgetary constraints of the healthcare system are sustained. These form the restrictions under which CEA and CUA may be used to identify efficient allocation of resources.[13,45,46] Due to the limited use of the CBA in evaluations of healthcare [11,51], in this remainder of the present dissertation the term economic evaluation will refer to CEA and CUA only.

The effect outcomes employed in CEA are measured in countable, natural units, which are expected to eventually affect health (see Box 2 for exemplification of CEA[52]). Effect measures could include life years gained by increased survival, a decrease in the risk of ischemic stroke in AF, or the number of days with an adequate level of quality in anticoagulant therapy[13,52]. These effect measures immediately meet the criteria for extra-welfarist evaluation outcomes as they are expected to ultimately affect health. However, their application gives rise to some challenges with respect to the interpretation of analyses. Application of disease- or treatment-specific effect measures obstructs the comparison of interventions that do not have identical purposes and, therefore, dissimilar effect measures, or even if a similar effect measure is defined slightly differently in studies, resulting in incomparable evidence. This complicates the identification of efficient resource allocation across clinical areas and thus oppose the fundamental purpose of economic evaluation. Furthermore, as the purpose of the economic evaluation is to inform on the optimal decision in order to improve health within budgetary constraints, the application of a one-dimensional effect measure may potentially be misleading if the chosen effect measure does not contain all impact on health. For many diseases and treatments, the application of a single effect measure, such as improved survival achieved through therapy, would lead to incomplete evaluation of health benefit, as health is likely to be affected by multiple aspects of treatment simultaneously. Lastly, thresholds for opportunity costs are not established for the vast majority of clinical effect measures, which renders the results of CEAs with nondominance uninterpretable.[11,13,52]

Effect measures for CEA also include clinical intermediate endpoints, such as changes in blood pressure, the level of serum cholesterol, etc. These may be applied as surrogates of final, patient-relevant outcomes. Their application may be justified if they are known to be predictors of patient-relevant outcomes that are known to affect health. Their application may be accepted if unreasonably long follow-ups or unrealistically large patient samples were required to observe direct impact on health. Changes in clinical intermediate endpoints do, however, not hold intrinsic health-related value and are thus not optimal outcome measures in economic evaluation. Their direct application as outcome measures in CEA may lead to uninterpretable or misleading analyses and results. Instead, there must be a link between the clinical,

intermediate endpoints and final, patient-relevant outcomes, which enables modeling of changes in final impact on health (cf. section 3.4 for elaboration).[13,53]

The WHO definition of health[44] provides the foundation for research in health-related quality of life (HRQoL), which conceptualizes the effect measure in CUA.[11] Hence, the results of CUA may best inform decision-makers within the healthcare system, as the outcome most succinctly reflect the maximand of the healthcare system. According to the WHO definition, health is multidimensional and comprises more than the physical dimension of health, namely mental and social well-being, which should be taken into account to fully appreciate the impact on HRQoL provided by interventions.[11,44] The functional capacity, i.e. the ability to perform usual every-day activities and the behavior that the functional capacity enables, is often equated with quality of life, where HRQoL refers to the functional capability that is contingent on health[54]. As the majority of the population suffers from one or more chronic diseases[6], the functional capacity that individuals experience, despite the presence of disease, becomes increasingly more important than, for example, survival.[11]

Claes et al.[52] performed a CEA to evaluate the costs related to increasing quality in warfarin therapy by four different quality-improving initiatives in general practice, which targeted the organizational model for management of OAC. Cost-effectiveness results were expressed as the incremental cost per extra day in which therapy was considered within the therapeutic range (cf. section 4.1.2 for elaboration). Claes et al. found that multifaceted education and the use of decision-assisting equipment dominated the previous organizational management model for OAC in general practice by improving the quality of therapy at a lower cost.[52]

However, if incremental costs and effect had been positive (or negative), it would have been impossible to evaluate the cost-effectiveness of the intervention as no threshold for the applied effect measure is available and the health-related value of more days with therapy within the therapeutic range is vague. This illustrates the difficulties of interpreting the results of CEA when dominance is lacking.

*Box 2. Exemplification of a cost-effectiveness analysis by Claes et al.*[52].

A generally accepted – though not globally implemented – generic measure of HRQoL is the QALY. The use of QALY as an effect measure operationalizes HRQoL and enables summary valuation of both the expected quality and the quantity of the life lived, i.e. morbidity and mortality, respectively, into a single index measure. This enables evaluation of interventions that may affect the quality and quantity of life simultaneously and potential trade-offs between gains in quality and quantity of life. The application of QALY as an effect measure of health benefit allows for comparison

of diseases and interventions across different clinical areas by 'translating' their impact on, for example, mortality and the occurrence of final, patient-related outcomes, such as stroke, myocardial infarction, etc., into impact on the single-unit measure. Hence, CUA meets the aim of economic evaluation of providing information on efficient allocation of resources within the entire healthcare system and not just a specific clinical area. By employing a single-unit measure to embrace all impact on health, the application of one threshold for opportunity costs is enabled. [13,14,49] Currently, in Denmark, there is no recognized threshold for opportunity costs for QALYs. For this reason, a threshold equal to £20,000–£30,000 per QALY[48] is often assumed.

When operationalizing HRQoL through the use of the QALY, different health states may be ordered according to their desirability, i.e. the QALY weight that is attached to them. The weights are generally ordered in relation to two health states, 'perfect health' and death, which comprise the reference points for setting the QALY weights of further health states. Normally, the maximum QALY weight that may be attached to a health state is 1, which equals perfect health. Hence, a year in the 'perfect health' state would yield one QALY. The corresponding QALY weight of death is set to 0 and, consequently, irrespective of the time spent in that health state, the accumulating QALYs would be zero, which is intuitively correct. The QALY weights that are attached to health states should be based on preferences, under the expectation that better health states are given higher preferences, which is consistent with the maximization approach in normative health economics. Any conditions or events befalling an individual that worsen the 'perfect health' state, such as disease or treatment complications, would lower the preferences for that health state, which would consequently be given a lower QALY weight. This would subsequently affect the accumulation of health benefit, i.e. QALYs, when individuals experience the health state over time. In disagreement with the welfare economist approach (cf. Box 1), the preferences for health states may be elicited from representatives of the general public and not only individuals experiencing the specific health states. Valuation and comparison of health states across individuals is thus enabled by the application of preference weights for the health states, elicited from the population rather than the individuals.[13,45,47]

Although use of the QALY ensures comparison of the effectiveness of interventions across different therapeutic fields and the setting of a single threshold for opportunity costs, it is still a disputed effect measure. The estimation and utilization of QALYs is enabled by a number of assumptions, whose bearing in real life is debatable[13,45]. One of the main assumptions concerns the rationality of decision-makers. Individuals are expected to demonstrate rational, risk-neutral, and consistent preferences towards health states[11,45]. For instance, a QALY should hold the same value, irrespective

of the therapeutic field in which it is obtained. Thus, increasing the HRQoL from a weighting of 0.2 to 0.4 should be equal to increasing the weighting from 0.8 to 1.0, based on an interval scaling of the QALY weights. This may not represent real-life preferences, as society may prefer to help those in poor health (QALY weight 0.2) to achieve a better HRQoL than those in relatively good health (QALY weight 0.8) to obtain perfect health. In addition, the assumption of rational decision-making requires a certain 'time independence' of the QALY weighting. Hence, a QALY should hold the same value irrespective of the age at which it is obtained, which may not reflect the opinion of society. Also, the QALY weight of health states should be independent of the duration of time, in which individuals are expected to experience the health state. The required preference consistency may thus be violated in real life as 'a QALY is not necessarily a QALY' in the eyes of the population whose preferences decision-makers should safeguard when prioritization is made.[13,45,46]

Simple aggregation of QALYs may therefore not be consistent with the normative purpose of maximizing the benefit to society, as the aggregated QALYs may not necessarily hold social value. Furthermore, the QALY has a rather narrow focus by only embracing impact on *health*-related quality of life[54]. The QALY has been criticized for not including other aspects that might equally affect individuals' quality of life, such as their perceptions and satisfaction with the convenience of care, etc. (insofar as these elements are not reflected in the QALY weights resulting from disease and treatment).[13] Consequently, it has been argued that use of the QALY does not necessarily reflect individuals' true, resulting capacity when such aspects are not explicitly included in evaluation. As a counterargument, it may be stated that the QALY is a pragmatic, reasonable, and operable approximation of the explicated maximand in the healthcare system, i.e. health, which is the primary concern of the policymakers, whom the analyses are intended to inform. As individuals' capability is contingent on the existence of a reasonable level of health, it has been argued that the QALY therefore provides a reasonable approximation of the expected capability of individuals.[13,46,47] Despite the perhaps insufficient reflection of true social value of the QALY, its development is still founded in extra-welfarism. Consequently, its use reflects the underlying utilitarian ethics and philosophy of normative health economics and reflects an efficiency-maximizing approach to decision-making. Utilitarianism and the ethical considerations that it represents remain disputed as a foundation for resource allocation, as "*an economy can be Pareto optimal, yet still perfectly disgusting by any standards*"[47]. CUA and QALY are accordingly subject to the same critique[11,46,47].

Despite the criticism of the QALY, CUA nonetheless is currently the preferred type of economic evaluation in healthcare and is increasingly applied.[45,51,55]

### 3.3.2. THE SCIENCE AND ART OF EVIDENCE-BASED MEDICINE

As health economics revolves around health and healthcare, the methods for economic evaluation have, to a large degree, evolved in accordance with the views and approaches seen in research in medicine, specifically evidence-based medicine (EBM). EBM is characterized by a rigorous approach to the design and interpretation of studies, with a strong focus on determining the potential causality between interventions and outcomes and statistical inference. Accordingly, the methods for economic evaluation have been developed to embrace the challenges and epistemological views that apply to EBM. As a result, evidence on particularly outcomes in CEA and CUA often derives from research in medicine.[11,24,56]

When the concept of the RCT was established in the 1980s, formal evaluation of new medical technology prior to its introduction became a reality. Since then, evidence of the effect and safety of medical technology is increasingly requested prior to its application in real-life practice.[14,26] In EBM, the RCT is regarded as the optimal design for evaluating the effect of interventions when the purpose is to investigate possible causal relationships between interventions and expected outcomes. The RCT design allows evaluation of the effect of interventions in a counterfactual, controlled environment, thereby enabling detection of the 'true' effect of interventions, in theory unaffected by other influencing factors. Hence, it is used to evaluate the relative safety and effect of one intervention, *ceteris paribus*, versus a comparator[39–42,57].

The RCT design is optimized for the evaluation of potential causality between a single or few determining factors and a limited number of outcomes.[26,56] The estimate of effect from clinical trials is referred to as the efficacy of interventions, which comprises the 'technical frontier', i.e. the maximum benefit from interventions that may be achieved under ideal circumstances[23,56]. By design, efficacy data may not hold the influence of other important non-medical factors that influence the effect of interventions outside the controlled trial environment[23,56]. Hence, when efficacy data are used to inform on the cost-effectiveness of interventions, the results may not reflect what will be observed in real life. This may especially apply, if the quality in within-trial therapy and in real-life practice differs or if trial populations are not representative of the patients treated in practice. Consequently, data on effectiveness rather than efficacy would, theoretically, constitute the better data grounds for economic evaluation when the purpose is to provide information on the expected NB of interventions in real life.[13]

EBM has been developed with the purpose of supporting clinical decision-making. In contrast, economic evaluation is to a high degree intended to support healthcare

prioritization and policymaking. In decision-making in clinical practice, decision-makers handle tangible decision problems concerning specific patients, whereas in policymaking, the individuals whom decisions concern are 'faceless.' Social decisions require empathy, but also a certain degree of detachment to provide decisions that are optimal for the common good and should therefore not be founded on the grounds of single, clinical cases. Consequently, policy decisions should be made based on what *on average* comprises the optimal solution. This may not be consistent with what constitutes the optimal solution in a specific clinical situation, for which reason the two decision contexts cannot be equated.[11,13]

### 3.3.3. WHAT YOU ASK IS WHAT YOU GET – THE IMPORTANCE OF STUDY DESIGN

Under extra-welfarism, the design of the economic evaluation should be contingent on and reflect the decision context, in which it is used. As such, the result of an economic evaluation differs depending on the question that it is to answer[47].

Often economic analysis comprises only a few, often only two, comparators. Of these, the currently used intervention or gold standard should be represented to ensure that the NB of new technology is compared to the NB of a realistic, otherwise used course of actions. Ideally, however, all relevant comparators for a specific decision problem should be compared to ensure full disclosure of the comparative cost-effectiveness of relevant interventions[13]. The applied perspective in an economic evaluation determines the costs that are included in the analysis and should be relevant to the decision-maker whom it is intended to inform. The broadest – and most demanding to the analyst – perspective is the societal perspective, in which all costs should be included, irrespective of who defrays them. When a narrower perspective than the societal one, e.g. a healthcare sector or hospital perspective, is applied, the cost analysis does not necessarily reflect the true cost to society if other payers outside the scope of the analysis defray the cost. In Denmark, the societal perspective is recommended as first choice to avoid fallible conclusions regarding the true cost-effectiveness of interventions.[13,14,58] Likewise, economic evaluation should include all incremental costs and effects, irrespective of their temporal occurrence, to avoid under- or overestimation of the cost-effectiveness of interventions. For interventions that have a long-term impact on health and cost accumulation, e.g. by increasing survival or decreasing a continuous risk of debilitating complications for treated patients, the appropriate time horizon for the evaluation of costs and effects

would potentially be patients' lifetime. This is generally the case for treatment of chronic diseases.[13]

These requirements immediately highlight the problems with using, for example, an RCT as the sole source of evidence on which to build a sound economic evaluation. Thus, clinical trials often fall short of including all relevant comparators, they often cannot capture all relevant impact on health and costs due to a different focus and a restricted follow-up and may combine clinical events into aggregate outcomes that become uninformative regarding the appertaining costs and impact on HRQoL[14,56]. This should not be considered a failing of EBM but illustrates that the discipline was not developed with the aim of informing policy decisions. To facilitate sound economic evaluation, it may therefore be necessary to compile evidence from multiple sources and to extrapolate data beyond the available evidence[13,14]. Decision-analytic modeling (DAM) provides the means of doing so[13,49].

## 3.4. DECISION-ANALYTIC MODELING – A TOOL TO AID DECISION-MAKING UNDER UNCERTAINTY

DAM is a tool developed upon the principles of statistical decision theory and Bayesian decision analysis and provides a systematic approach to decision-making under conditions of uncertainty[49,59,60]. DAM has its roots in welfare theory and was developed with the explicit purpose of supporting decision-making and prioritization when resources are scarce. DAM is used to identify the optimal alternative included in the economic evaluation, e.g. by identifying whether an intervention is expected to be cost-effective compared to the gold standard. As a vehicle for economic evaluation, DAM enables many of the requisites of a sound economic evaluation. It enables comparison of all relevant comparators, inclusion of all relevant impact on health and costs by allowing relevant evidence to be compiled from multiple sources, and extrapolation of clinical data beyond their observation to reflect an appropriate time horizon. In DAM, a model is built in which all relevant elements are included systematically to reflect the decision problem at hand. A model is composed of a set of mathematical relationships between the entities that constitute the alternatives under investigation. The use of mathematical relationships between the entities enables, amongst other things, the linking of clinical, intermediate endpoints to patient-relevant outcomes and the linking of patient-relevant outcomes to expected impact on HRQoL.[13,49,60]

The results from DAM are based on expected values of the different alternatives under investigation. In the models, the consequences of the alternatives, i.e. events occurring and both their potential impact on health and costs, are contingent on the probability of the events occurring. Thus, the expected values of the different decision options in a model are contingent on the consequences of the events included in the model multiplied by the probability of the events occurring. It follows that the expected values reflect the mean values of effects and costs of the different decision options. Decision-analytic models present simplifications of the alternatives and their impact that would be expected in the real world. Construction of the models requires explicit decisions regarding the inclusion of expectedly essential elements in the model – and exclusion of elements that are expected to not affect the results and hence are irrelevant in the decision-making context.[13,49,60]

The different parameters included in a decision-analytic model, including cost, effect, and probability parameters, may be included as distributions reflecting the uncertainty pertaining to them. The impact of these uncertainties may be included and investigated in sensitivity analyses. Hence, DAM enables the identification of what, on average and based on the included evidence, constitutes the most efficient alternative. Furthermore, DAM entails an explicit acceptance that decisions must be made under uncertainty and actively enables the incorporation and evaluation of the uncertainty that pertains to the utilized evidence and how it may affect decision uncertainty. In addition, via the construction of the decision-analytic model and retrieval of evidence to populate it, any gaps in current evidence and highly uncertain parameters may be highlighted.[13,49]

### 3.4.1. BAYESIAN DECISION THEORETIC APPROACH TO DECISION-MAKING

Irrespective of the origin of the evidence used in economic evaluation, some degree of uncertainty always pertains to it. As such, decisions must be made under conditions of uncertainty, which introduces the risk of making suboptimal decisions. Making the ‘wrong’ decision would impose opportunity costs. Either because cost-effective therapy is detained and thus unavailable to patients who would otherwise have benefitted from it, or because cost-ineffective therapy is reimbursed, thereby decreasing the total health benefit that the total patient population might have achieved if the resources had been used elsewhere. Therefore, the risk of making the wrong decision carries a cost in itself.[14,49,50]

In a policy context, decisions should be made on the mean expected NB of interventions, irrespective of the uncertainty that pertains to the included evidence[13,50]. If applying the rules of classical statistical inference, any intervention demonstrating a positive, but not statistically significant incremental NB versus a relevant comparator should be rejected. However, as the mean incremental NB was positive, this rejection would impose opportunity costs. Thus, it has been argued that in a decision-making context, no particular attention should be paid to the level of statistical significance[26,49,50]. Although uncertainty, therefore, should not immediately affect the decision as to whether to adopt an intervention, it may still affect the decision-making process. If the evidence base for an evaluation is insufficient or of poor quality, the validity of any evaluation relying on it may be disputed. Applying a Bayesian decision theoretic approach to decision-making, two conceptually different decision questions present themselves in the decision-making process that should be distinguished from each other. The first question of whether to adopt an intervention, given the existing evidence should be distinguished from the second question of whether further evidence should be requested to reduce the uncertainty that pertains to the included evidence.[49,50,59] The latter might be the case if the evidence substantiating the evaluation is considered to be too uncertain.

DAM may be utilized to evaluate the uncertainty that pertains to the different parameters included in a model and how it may affect the certainty of cost-effectiveness of the interventions under investigation and, consequently, the validity of the evaluation results. The individual parameters may be included in the model by using appropriate distributions that reflect the uncertainty applying to them. Probabilistic sensitivity analysis (PSA) may subsequently be used to evaluate all parameter uncertainty simultaneously and how it affects the decision uncertainty. In PSA, repeated sampling from the included parameter uncertainty distributions generates a NB distribution of interventions, typically consisting of 10,000 individual NB estimates. PSA may be used to indicate the probability that the optimal decision, as established via the expected values in the model, is, in fact, the optimal decision when taking into account the combined parameter uncertainty in the model. If the optimality of the decision is highly uncertain based on the current level of evidence, it might be considered whether further evidence should be requested to decrease the risk of making the 'wrong' decision. That is, recommending interventions that, in fact, have a negative incremental NB compared to their comparator. As the combined parameter uncertainty affects the decision uncertainty, the presence of poor-quality evidence increases the risk of making suboptimal decisions.[13,49,59] The potential cost of uncertainty, i.e. the potential cost of making the wrong decision due to existing uncertainty, constitutes a 'payer uncertainty burden;' the burden that the payer and decision-maker must carry if their decision is wrong.

From a decision-oriented view, the uncertainty that revolves around particular parameters is only of interest if it is able to change the decision that would otherwise be made, based on the expected values from the analysis. Consequently, sufficient evidence is not necessarily perfect evidence. Parameters may be highly uncertain, but not necessarily affect the decision uncertainty, for which reason their uncertainty is inconsequential to the decision and therefore – from a decision-oriented viewpoint – unimportant.[49,50,61]

The evidence base may be improved by acquiring further information via, for instance, further research. It is not possible to eradicate all uncertainty, though; some will always remain[14,49]. Conducting studies to generate more evidence draws on limited resources; if more precise estimates of a parameter are not expected to affect the decision uncertainty substantially and thereby the decision, the execution of such studies would represent inefficient use of resources.[14,59,61] This highlights the potentials of identifying parameters that affect the decision uncertainty, and on which uncertain parameters it might be beneficial to procure information. Value-of-information (VOI) analysis may be used for this purpose[59]. VOI analysis can be used to calculate the total cost of uncertainty related to all parameter uncertainty in a model, but also the cost of uncertainty pertaining to specific parameters in the model, via the expected value of perfect information and the expected value of perfect parameter information, respectively. These estimates may be used to evaluate the potential cost of uncertainty of specific parameters and, consequently, determine the amount of resources that maximally might be allocated to eradicate the uncertainty.[49,60] The use of VOI analysis is further elaborated on in the paper presented in Appendix A.

## 3.5. QUALITY IMPROVEMENT IN THE HEALTHCARE SYSTEM

To enable the evaluation of quality improvement, it is necessary to generate a consensus about what may be understood by the concept of 'quality' and quality improvement and subsequently how it may be evaluated. Consequently, the purpose of the presentation in this section is not to delve into the intricacies and challenges in quality evaluation and improvement in practice, but only outline the concepts that comprise the foundation for the majority of research in quality in healthcare.

### 3.5.1. THE ELEMENTS OF QUALITY

The foundation of modern quality evaluation is often ascribed to the physician Avedis Donabedian. In the 1950s, Donabedian developed a conceptual model that provides a framework for the assessment of quality and quality improvement in healthcare. In the Donabedian framework, no definition is supplied as to what 'quality' in healthcare actually comprises, but it may ambiguously be said to be a reflection of the values and purposes of the medical care system and, on a larger scale, the society as a whole.[23] In this interpretation, the focal point of quality in healthcare is the health of patients and the population as a whole[23,43], but it also implies that quality in healthcare entails other aspects of healthcare delivery. In the Donabedian framework certain attributes of quality are attached to healthcare interventions that can be used to evaluate their quality (Box 3)[24,62]. The attributes pertain to different aspects of healthcare delivery and simultaneous consideration of these is required to achieve a thorough evaluation of the quality of healthcare interventions. As they represent elements of quality, it is desirable to improve the attributes, which makes them objects for quality improvement. Evidently, some of the attributes are synergistic. However, some are also contradictory and their individual importance should therefore be considered and weighted when interventions are evaluated.[62]

What is perceived as quality in different situations may differ, and hence a uniform definition and evaluation of the quality of healthcare across interventions is precluded under the Donabedian model. The lack of a clear, specific definition of quality in the Donabedian framework for assessment of quality and quality improvement is both a strength and a weakness. It does not confine the use of the model to evaluations where a certain interpretation of the concept of quality is applied, but it also carries the risk that interpretation and subsequent measurement and comparison of quality in different decision contexts are hampered.[24,62]

*Efficacy* The effect of an intervention under optimal circumstances. The efficacy of an intervention constitutes the technological frontier, i.e. the maximal improvement that may be expected from the intervention given specific circumstances and patients' characteristics.

*Effectiveness* The observed effect of the an intervention under real-life circumstances as experienced in daily practice, i.e. the degree to which the demonstrated efficacy may be achieved under conditions of, potentially, alternate quality levels.

*Efficiency* A measure of the cost that is required to gain improvement through the intervention. If more interventions with the same gain are compared, the intervention with the lowest cost is the more efficient.

*Optimality* When the value of the improvement of interventions is compared to the costs related to gaining it. At the optimum, the cost related to gaining improvement is at the lowest. With optimality, the concept is introduced that the marginal value of increasing quality may by counterbalanced by the costs required to gain the value, potentially rendering the costs related to a quality improvement unreasonably high.

*Acceptability* Patients' subjective valuation of effectiveness, efficacy, and optimality, but also of their experience and preferences regarding the physician-patient relationship, accessibility of care, etc.

*Legitimacy* The acceptance of the society of interventions. Legitimacy may differ from individual patients' acceptability, as it concerns the welfare of the collectivity, which may be different from the desires of the individual.

*Equity* The just distribution of healthcare to the members of the population. That is what is considered acceptable to the individuals of society, but also legitimate to the society as a whole. Equitable distribution of healthcare may consequently not necessarily be the most cost-effective.

*Box 3. Attributes of quality of healthcare as described for the Donabedian framework*[62].

In the Donabedian model, three dimensions of quality are furthermore presented to ease classification of information on the quality of interventions and, hence, evaluation. These dimensions – and associated classifications – are *structure*, *process*, and *outcome*.[23,24] In this conceptualization, the dimension structure refers to the system characteristics that form the structural and organizational framework of a setting. Structure encompasses the attributes of the setting, including the available conditions and resources that restrict and determine the ability of the system to meet

healthcare needs. Thus, structure includes the organizational preconditions of healthcare provision, e.g. the organization of healthcare personnel and financial management, and physical surroundings – both material resources, such as facilities and equipment, and human resources, such as qualified personnel. Process refers to the activities – or services – that are carried out by healthcare personnel in the setting under evaluation. The process dimension, therefore, entails all clinical processes in relation to the continuity of care of patients, including diagnosis, therapy initiation, continuous care, rehabilitation, etc. Lastly, outcome refers to the effect of the intervention under evaluation on the health status of patients and the population. The health status is affected by changes in patients' HRQoL, mortality, morbidity, complication rates, functional level, etc. With the application of a broader definition, changes in patient behavior, understanding, and satisfaction may also be included in the definition of outcome.[20,23,24]

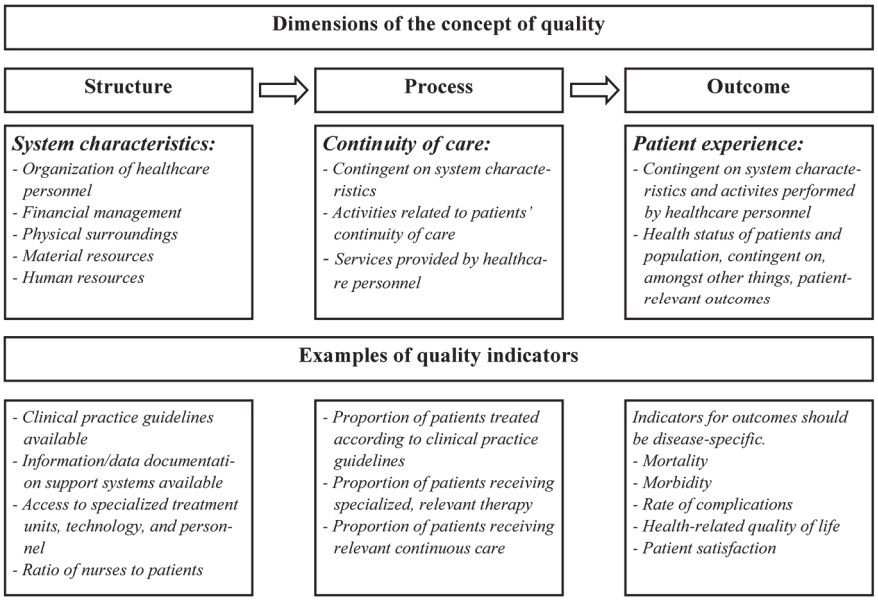



*Figure 3-2. Dimensions of the concept of quality and examples of quality indicators that may inform on the quality of healthcare. Based on the Donabedian model.*[20,24]

The three dimensions of quality in healthcare are often presented as a unidirectional causal chain (Figure 3-2), in which the quality of prior dimensions conditions the quality of later dimensions. Hence, the existence of suboptimal quality in, for instance, structural quality, e.g. available technical equipment, will be propagated through the model and obstruct the attainment of high quality in patient-relevant outcomes

affecting health. Thus, the ultimate effect on health status is contingent on the preceding dimensions to be of sufficient quality. If deficiencies in a prior dimension are present, these will manifest themselves throughout the rest of the causal chain.[23] The three-part approach to classification of information on quality of healthcare may improve intelligibility when establishing the possible impact of changes in the different dimensions of quality on patient-relevant outcomes and health status. The Donabedian model remains a cornerstone in modern approaches to quality evaluation and improvement in practice. Today, it is the dominant paradigm for quality improvement assessment and research.[19,20]

The influential American Institute of Medicine (IOM; now known as the National Academy of Medicine) has proposed a framework for quality improvement that is perhaps more workable than the conceptual model by Donabedian. In substance, the IOM framework is based on the Donabedian model. It does, however, include a more explicit definition of the concept of quality in healthcare.[63] Hence, the IOM has defined quality as "*the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge*"[63]. In this definition, the focal point of quality in healthcare is the achievement of desirable health outcomes. This is in general agreement with the Donabedian definition[23], although the societal considerations related to quality of healthcare represented by the attributes '*legitimacy'* and '*equity*' in the Donabedian framework are deprioritized (Box 3)[63]. In 2001, the IOM adopted six constituent elements of quality that may be subject to improvement: safety, effectiveness, patient-centeredness, timeliness, efficiency, and equity (Figure 3-3). With the IOM definition of quality in healthcare and the explication of the six aims for quality improvement, the IOM instigated a focus on what real-life quality improvement may entail, thereby providing a framework for quality improvement in practice.[63,64]

In essence, most approaches to quality evaluation and improvement reflect similar interpretations of the concept of quality and entail more or less the same constituent elements as advanced in the Donabedian framework and the IOM conceptualization of quality in healthcare[20,62,63]. All approaches adopt health status – and satisfaction – as the focal component of quality in healthcare. Thus, it could be presumed that the ultimate purpose of quality improvements must be to improve the health of patients and the population, which is in line with the purpose of economic evaluation[23]. When appraising the applicability of economic evaluation to the area of quality improvement, the IOM aims for quality improvement can be used to aid contextualization[63]. Consequently, given an appropriate research design, economic evaluation accommodates the assessment of the elements of effectiveness, safety, efficiency, and timeliness, the outcomes of which may affect patients' health status and/or satisfaction. In contrast, the most prevalent approaches to economic evaluation

do not explicitly include the elements equity and patient-centeredness in their methodological framework. Thus, it may be inferred that quality in healthcare is composed of more components and reflects a higher complexity than is traditionally embraced by the maximand used in economic evaluations. This discrepancy between the potential targets for quality improvements and what is used as maximand in economic evaluation should be deliberated when economic evaluation is considered for quality improvements.

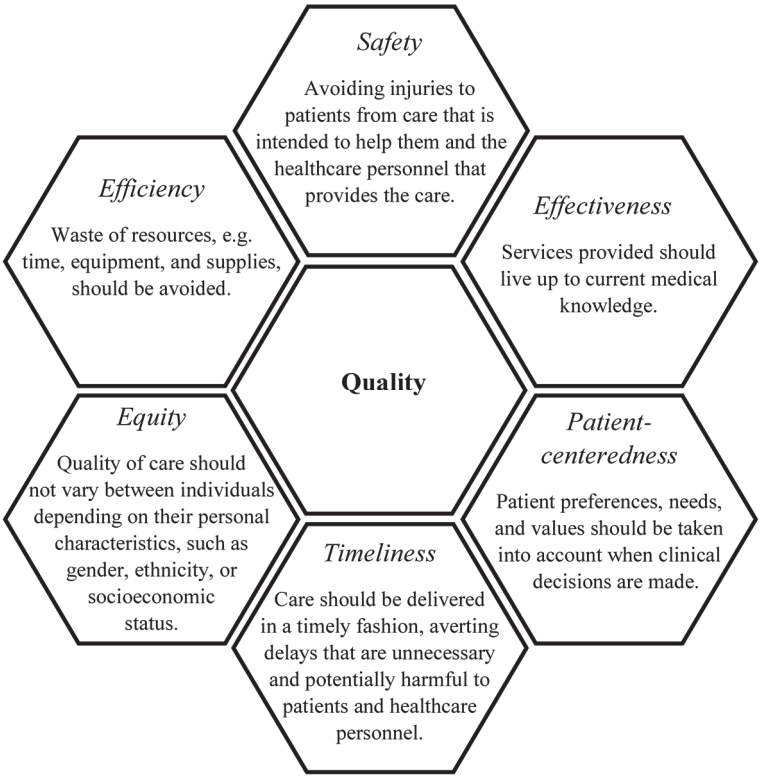



*Figure 3-3. Constituent elements of quality and aims for quality improvement proposed by the Institute of Medicine*[63,64].

### 3.5.2. APPLICATION OF QUALITY INDICATORS FOR QUALITY EVALUATION

Quality and quality improvement in healthcare may be assessed by qualitative and quantitative methods, depending on the purpose of the assessment[19,23,24]. In the present dissertation, the focus is on the quantitative assessment of quality improvement, which requires instruments to facilitate measurement and evaluation of possible effects. QIs provide useful, quantitative measures for the evaluation of quality and any change achieved through improvements in care. QIs may be used for multiple purposes including documentation of care, benchmarking, priority setting, and evaluation of the effect of quality improvement.[20,23] The present dissertation focuses on the use of QIs as measures of quality improvement.

Most definitions of QIs relate to the dimensions of quality presented in the Donabedian model and the information they provide is classified accordingly[20,24]. Structural QIs reflect preconditions that are conducive to healthcare provision; for instance, they may indicate the availability of clinical practice guidelines that are conducive for evidence-based therapy. Process QIs are used to assess healthcare delivery under the circumstances provided by the system characteristics, e.g. the proportion of patients treated according to clinical practice guidelines. Finally, outcome QIs may be used to evaluate the impact of healthcare on the health status of patients, for instance through the occurrence of patient-relevant outcomes, such as changes in the mortality or disability of patients.[20,24] Outcome QIs, furthermore, include intermediate indicators of biological functioning, i.e. biomarkers that are known to have longer-term effect on the health status of patients[20]. Evidently, there is consistency between the classification of outcome QIs and the preferred outcomes in EBM. Examples of QIs related to the different dimensions of quality are given in Figure 3-2.

QIs may be employed with different objectives. Sentinel QIs indicate occurrences that are inherently undesirable, such as severe adverse events, and are used to flag incidents where the course of events leading up to the occurrence requires further scrutiny. In the present dissertation, however, the focus is on QIs used to evaluate the quality of everyday practice, i.e. related to events that occur with a certain frequency within the expectations of standard care. Ideal QIs are reproducible, measurable, generally accepted by practitioners and researchers, evidence-based, and outcome-validated, i.e. the relationship between the QI and patient-relevant outcomes should be established[20,23]. Despite this, all QIs are to some degree sensitive to measurement error, which affects their validity. Greater uncertainty is introduced when subjective judgment is required to indicate quality. In particular, measurement

of process QIs may be susceptible to some degree of valuation and, therefore, be less stable than outcome QIs. For instance, for a process QI, e.g. "proportion of patients treated according to clinical practice guidelines," a certain degree of interpretation may be required to assess whether sufficient adherence to the guidelines is demonstrated. In contrast, patient-relevant outcomes, such as mortality and rates of clearly defined complications, constitute unquestionable QIs of effectiveness and safety, for which impact on health cannot be disputed.[23] The request for objective measures of quality favors the use of estimable endpoints, which are less prone to valuation[20].

When quality and quality improvement are evaluated by the use of QIs, the measures may be compared to quality standards, which may indicate aims that should be achieved for the quality to be considered acceptable. The standards should be evidence-based and derive from reliable sources. Their origin may be empirical or normative, where empirical standards derive from evidence on the level of quality in healthcare that may be attained in real-life practice. In contrast, normative standards are often agreed upon by experts in the therapeutic field, for which the standard is developed and may reflect levels of both acceptable and optimal care. Standards may furthermore be explicit and implicit. Explicit standards set a pre-specified quantitative aim that should be achieved for quality to be deemed acceptable, whereas implicit standards may be set for individual patients and be adapted to specific cases to adjust for factors that may affect the level of quality that may be realized.[20,23,24]

To enable sensible and informative evaluation of interventions, the QI(s) that are applied should be sensitive to changes imposed by the intervention and should reflect the main purpose of the intervention[23,24]. For instance, if the purpose of a quality improvement were to enhance rehabilitation after ischemic stroke, mortality would be a poor choice of indicator of effectiveness. Accepting the basic assumption that the main purpose of the healthcare system and quality improvement is to improve and maintain health, the best QIs of effect would offhand be outcome QIs that impact patients' health status directly[23]. However, for some quality improvements the application of patient-relevant outcomes as QIs would not be feasible due to them occurring only sporadically or after long follow-ups, or if only small patient populations were available for study, resulting in underpowered analyses.[20,24,27] For some interventions, structural or process QIs may be more sensitive to improvement caused by interventions and may, therefore, be favored to enable detection of statistically relevant changes. For instance, process QIs are more sensitive to improvements in the act of caring that may go undetected by an outcome QI because, within reasonable study implementations, the incidence of patient-relevant outcomes would be impacted by relatively little. In such cases, meaningful assessment of possible quality improvements through the use of outcome QIs would require

unreasonably large samples or follow-ups.[27] With the application of structural, process, and intermediate outcome QIs evidence on effect may be procured without delay[20,53]. The National Quality Measures Clearinghouse, which is an initiative of the Agency for Healthcare Research and Quality under the U.S. Department of Health and Human Services, holds more than 2000 evidence-based QIs available to the public[65]. More than 60 percent of the registered QIs are process QIs[66], which emphasizes the currency of the use of these QIs in research in quality improvement.[23]

Structural and process QIs do not possess intrinsic value. They only signify quality if change in them ultimately reflects change in patient-relevant outcomes and health status.[23,67] Therefore, the relationship between QIs and the impact on health status should be validated prior to the application of the QIs for assessment of the quality in healthcare[20,24,67]. However, the evidence supporting the relationship between structure, process, and outcome is often sparse. Thus, there are QIs and standards used in practice that have not been outcome-validated (exemplified in Box 4)[67]. At best, the use of QIs and standards that has not been validated may lead to vague and questionable conclusions on the performance of interventions. At worst, their use may lead to erroneous conclusions on the benefit of interventions.

The interval between tests of international normalized ratio (INR) during warfarin therapy of vulnerable elders has been propounded as a process QI, with a maximum time interval of six weeks between tests representing the standard. However, the correlation between testing intervals and outcomes is only suspected, and not validated, and the six-week standard is not based on evidence suggesting that safety and effectiveness should be improved by this testing interval.[23,67]

If accepting the premise that the main purpose of the healthcare system is to improve health, this QI and standard do not necessarily reflect quality in healthcare and may be erroneously applied.

*Box 4. Exemplification of the use of a non-validated quality indicator and quality standard within oral anticoagulant therapy*[67].

If there are many antecedents between applied QIs and patient-relevant outcomes, it may affect the ability to establish causal inference between changes in the QI and changes in health status. Antecedents include explanatory factors for the elements under investigation and may include, for instance, the organization of continuity of care, the competencies of healthcare personnel, the doctor-patient relationship, patient and disease characteristics, etc.[20] As more antecedents may take effect with greater distance in the causal chain between the QI and patient-relevant outcomes, the distance may impede the establishment of causality. The same applies to the

relationship between the intervention and expected change in the applied QI: If the QI may be affected by more antecedents, the attribution of effect to the intervention may be questioned (Figure 3-4). The existence of antecedents thus attenuates the causal inference between interventions and ultimate change in health status due to the antecedents. There is an inverse relationship between the distance between the intervention and patient-relevant outcomes and the ability to make causal inferences about the relationship between the intervention and ultimate health gain.[23] As a result, quality 'improvements' targeting the structural quality of healthcare are likely not to result in noticeable changes in patient-relevant outcomes within reasonable study setups[24]. The deployment of such interventions may therefore be disputed, if their impact on health cannot be established. This poses an objection to the sole use of patient-relevant outcomes for the evaluation of quality improvements as these will be influenced by all antecedents preceding their occurrence in the causal chain[27].

Antecedents may be equally or more influential than the quality improvement itself[27]. Unfortunately, the presence of these explanatory factors is often not included in analyses[23]. Therefore, erroneous conclusions may be drawn about the possible impact of the intervention on ultimate health if the potential effect of antecedents is not sufficiently taken into account when the relationship between QIs and patient-relevant outcomes is estimated. If the relationship between applied QIs and impact on health status is not adequately validated, the application of particularly structural and process QIs may hamper reliable evaluation of ultimate effect on patients' health of potential quality improvements.

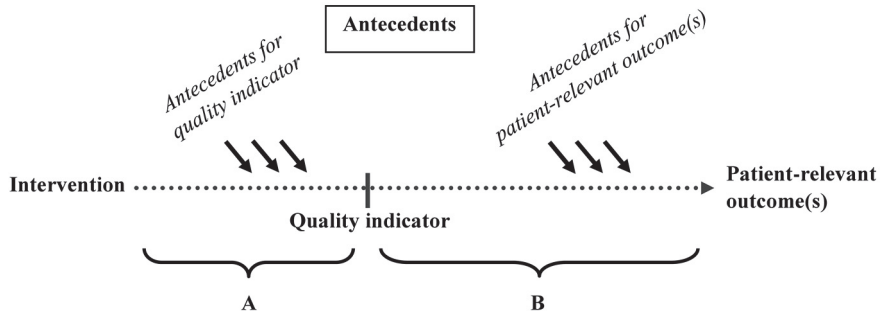



*Figure 3-4. Impact of antecedents on the usefulness of quality indicators for evaluation of quality improvements.*
*A: antecedents impeding the establishment of causality between an intervention and a quality indicator. B: antecedents impeding the establishment of causality between a quality indicator and patient-relevant outcome(s).*

What provides optimal QI(s) for a specific intervention depends on the objective of the evaluation. When assessing the reliability of a potential QI for evaluation of a quality improvement, the antecedents between the intervention and the QI and between the QI and patient-relevant outcomes and their potential to disturb the causal inference between the intervention and health status ought to be weighted. As quality in healthcare is determined by multiple attributes (Box 3), the impact of quality improvements on all attributes is unlikely to be sufficiently embraced by a single QI[20,24,62]. Comprehensive evaluation of the impact on all dimensions of the quality of healthcare might, therefore, require the application of multiple QIs. Likewise, it is possible that a single QI would not adequately capture all impact of a quality improvement on a certain aspect of quality, e.g. health status. Again, this would necessitate the application of more QIs to reflect all impact.

To obtain more tangible estimates of the relationship between changes in QIs and impact on health, the expected correlation between the QI and outcomes should be calculated numerically. With the application of appropriate statistical methods, it might be possible to adjust for the presence of antecedents and to quantify the uncertainty pertaining to the correlation. This might be usable in analyses of the expected cost-effectiveness of interventions.

### 3.5.3. CHALLENGES FOR ECONOMIC EVALUATION OF QUALITY IMPROVEMENTS

Acknowledgment of the importance of the quality of healthcare for economic evaluation may best be stirred when considering the differences between the observed efficacy in, for example, RCTs and the subsequent, observed effectiveness achieved under real-life conditions, where the quality of care may not be as high as observed within trial[30,34,62]. It is furthermore recognized that inefficiency in healthcare provision and low quality in healthcare coexist, as low quality is either directly harmful to patients or because it causes inefficient use of healthcare resources, thereby entailing opportunity costs. It could be hypothesized that if the quality of healthcare is increased by investment in a quality improvement, the health benefit related to the quality improvement will, likewise, increase. [18,24,62] As quality in healthcare thus is a determinant of the expected effectiveness and potentially safety of interventions, this encourages evaluation of interventions that narrow the gap between real-life effectiveness and the technological frontier. Such quality-improving interventions may potentially represent cost-effective interventions to increase health gain.

The correlation between cost and health benefit related to the level of quality is likely to be nonlinear (Figure 3-5)[24]. This creates the foundation for the concept of optimality (Box 3), in which the health benefit achieved by investment in a marginal quality improvement may not counteract the cost related to the marginal quality improvement. Thus, the optimally effective level of quality may not be the maximally effective level (Figure 3-5)[62]. At the optimally effective quality level, an equilibrium is reached above which further investment in quality will not be counteracted by the gain achieved by the investment and where the cost related to 'a unit of quality improvement' is the lowest. By the introduction of the concept of optimality, it is implied that there may be limits to the acceptable costs of quality improvement.[13,47] This supports the hypothesis of the present dissertation that economic evaluation may be applicable to the area of quality improvement. It also highlights the prudence of establishing the value for money of quality improvements to substantiate informed decision-making.

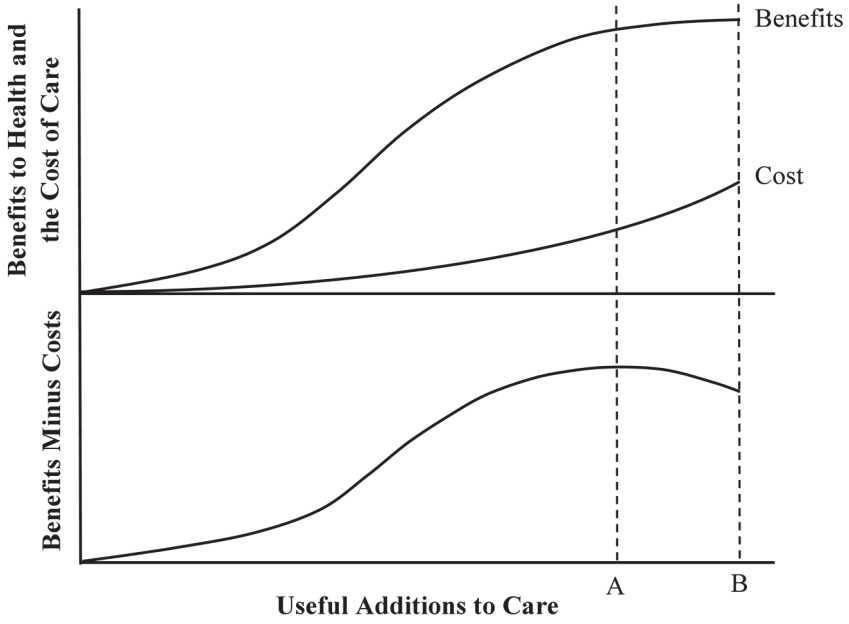



*Figure 3-5. Outline of hypothetical relationship between health benefit achieved by increasing quality and the related costs of attaining that quality. A: optimally effective quality, B: maximally effective quality. Adapted from Donabedian*[62] *and reproduced with permission from the American Medical Association.*

Nonetheless, economic evaluation of quality improvements is rarely undertaken[16,25]. This may be due to a resistance towards valuation of quality improvements because of (misinterpreted) ethical concerns or due to a lack of evidence on the impact on health and costs. Often quality improvement is intended to increase patient safety by decreasing the risk of treatment-related harm that befalls patients or healthcare personnel. To some, it may be of invaluable priority to avoid medical errors or medical inadequacy that lead to harm. It might be argued that avoiding afflicting further harm on patients cannot be valued. If adopting this attitude towards the necessity of avoiding iatrogenic complications, economic evaluation would have no bearing on policy decisions on whether to adopt quality improvements or not. However, in adopting a normative approach to evaluation, the opportunity costs that would be introduced by accepting quality improvements without consideration of their expected health benefit and associated costs, might be equally harmful to other patients.[18,25,63]

An additional recurring obstacle to economic evaluation of quality improvement remains an apparent lack of acceptable data on resource consumption and the effect of interventions[18].

Considering the diversity of the constituent elements of quality in healthcare, it is hardly surprising that approaches to research and evaluation within the 'informal art of quality improvement' originate from different scientific disciplines. Accordingly, the methodologies applied for evaluation of quality improvement derive from, amongst others, healthcare science, engineering, and behavioral sciences. The different scientific disciplines all have the potential to provide information on the value of interventions. They do, however, take different epistemological approaches to research and evidence, which may cause disagreements on the validity of evidence and the interpretation of results of studies.[24,26,68]

Whereas the purpose of research in medicine is to improve the clinical evidence on the safety and efficacy of interventions, i.e. focused on the final impact on health, research on quality improvement in healthcare is to a greater extent focused on improvement via the processes of care and the processes themselves[26]. Quality improvement is often a dynamic, data-driven process, in which changes are introduced continuously to improve outcome and multiple factors are often introduced concomitantly. Thus, quality improvements may comprise large, multifactorial interventions that cause system changes through complex mechanisms and affect more dimensions of quality simultaneously.[19,26,69] It follows that multiple antecedents may take effect and hamper the establishment of causal relationships between interventions and patient-relevant outcomes. Continuous improvement though actions and changes in behavior inherently go against the fundamentals of

EBM, in which the causality between a single or a few factors and outcomes is determined while other elements are adjusted for. Although EBM – and the RCT in particular – may be superior for evaluating medical technology, for which the intervention and outcome are clearly defined, it may be less apt for the evaluation of multifactorial, complex interventions. Given that the foundation for economic evaluation is suffused profoundly by the epistemological basis for EBM, the epistemological discrepancy between EBM and research in quality improvement in healthcare may be transmitted to the application of economic evaluation to the area of quality improvement.[26,69]

Valid data on costs and effects related to changes in processes of care may thus be difficult to obtain. Quality improvement may optimize processes and, in some cases, reduce variable costs, but the impact on fixed costs and costs related to a reduced risk of patient-relevant outcomes may be difficult to establish with accuracy within the study time frames that are often imposed on research in quality improvements.[18,25] Furthermore, as previously explained, the effectiveness of quality improvement is often measured via QIs. This complicates the evaluation of cost-effectiveness via conventional methods, as the QIs may not possess intrinsic value, i.e. not reflect change in health, her se. Estimations of cost-effectiveness, in which QIs have been applied for effect measure may thus generate unintelligible results.[18,20,65]

The lack of credible data on health-related effects and costs of quality improvements and the methods that are applied to obtain the data constitute a hurdle for economic evaluation. However, the increasing use of QIs emphasizes the need for methods to establish the cost-effectiveness of such interventions.

# CHAPTER 4.
# ATRIAL FIBRILLATION

The increasing pressure on healthcare budgets worldwide is partly due to the increased prevalence of chronic diseases, such as cardiovascular diseases – including AF[5]. AF is the most common cardiac arrhythmia worldwide with an estimated prevalence of approximately 3 percent in the general population. The prevalence increases with age and more than 6 percent of individuals over 70 years of age suffer from AF. Hence, the majority of AF patients are over-65s and the mean age of patients at first diagnosis of AF is higher than 70.[70,71] In Denmark, the estimated prevalence of AF is 120,000 individuals and the current annual incidence is between 15,000 and 20,000 individuals [29]. The prevalence is expected to rise dramatically in the foreseeable future due to the increase in longevity and an increased prevalence of predisposing conditions, such as diabetes mellitus, obesity, and prevalent heart failure. Consequently, it has been estimated that a quarter of all 40-year-olds in Europe and the US will develop AF in their lifetime. [5,70–73] The exact prevalence and incidence of AF are unknown as some patients may be undiagnosed due to 'silent AF,' i.e. asymptomatic AF. In other patients, the disease only manifests itself intermittently and the patients may not seek medical attention. Consequently, some patients are not diagnosed with AF until they suffer from severe complications related to the condition, such as stroke, heart failure, or peripheral arterial embolism, which require urgent treatment and may have long-lasting consequences.[28,74]

In addition to the great personal cost that AF may impose on patients and their relatives, the economic burden of AF is substantial[75–78]. Very conservative estimates indicate that the direct cost of AF may amount to at least 1 percent of the total healthcare expenditure.[75] The complications related to AF, such as stroke, constitute the main cost driver[75–77]. In a Danish cost-of-illness analysis from 2016 Jakobsen et al.[79] found that the three-year societal costs related to first-incident ischemic stroke in patients with AF amounted to approximately DKK210,000 (2016 value) per patient with hospitalizations being the main cost driver. As the prevalence of AF is only expected to increase in the future, the costs related to complications are expected to also rise[76].

## 4.1. ANTICOAGULATION – THE CHOICE BETWEEN SCYLLA AND CHARYBDIS

In AF, the muscle of the heart atria do not contract efficiently or at a normal pace but tremble, which prevents an efficient blood flow through the atria. This increases the risk of blood stasis within the atria, which increases the risk of thromboembolism. The anomalous contraction of the atria may, furthermore, affect the contraction of the ventricles and cause symptoms. The symptoms of AF can be diverse and unpleasant, and intrinsically lower patients' HRQoL, but may be mitigated by appropriate rate- and rhythm-controlling medication that stabilizes the erroneous contraction of the heart. Despite these therapies, AF still substantially increases all-cause mortality and morbidity.[28,80] In a Danish register-based study from 2011, Olesen et al.[81] investigated the event rate of thromboembolism (i.e. peripheral artery embolism, ischemic stroke, and pulmonary embolism) leading to hospitalization or death of AF patients who were not receiving OAC. For uncoagulated AF patients with an intermediate to high risk of stroke and systemic embolism (SSE) (cf. section 4.2 for elaboration), the rate of thromboembolism ranged from 1.45 to 5.72 per 100 patient-years [81]. Patients with AF suffer a five-fold increased risk of ischemic stroke, compared to patients with sinus rhythm and it has been estimated that 20 to 25 percent of all strokes are attributable to AF.[29,80]. As the majority of strokes in AF are cardioembolic, they furthermore tend to be more severe and entail greater disability and mortality than strokes that are not related to AF[82,83]. The event rate of ischemic stroke has been estimated to range from 0.10 to 2.00 per 100 patient-years for uncoagulated AF patients with an intermediate to high risk of SSE [84]. Consequently, stroke prophylaxis is a critical objective of AF therapy and lifelong OAC is recommended for the majority of patients[28,29]. Other nonpharmacological approaches to stroke prophylaxis exist, but the use of OAC remains critical in all strategies[28].

OAC reduces the coagulability of the blood and thus prevents stagnant blood from coagulating within the fibrillating atria, effectively reducing the risk of thromboembolism that may lead to ischemic strokes and other complications. Reducing the coagulability of the blood naturally induces an increased risk of bleeding. The most severe adverse event of OAC is intracranial hemorrhage (ICH), which may entail the same persistent and dire consequences for patients as ischemic stroke[85–87]. The rate of ICH has been estimated to range from 0.10 to 0.30 for uncoagulated AF patients with an intermediate to high risk of SSE; OAC may increase this rate by up to 50 percent.[84] In spite of this, a positive net clinical benefit of OAC has been found for the majority of AF patients with an intermediate to high risk of SSE, irrespective of their risk of bleeding (cf. section 4.2 for elaboration)[84,88,89].

### 4.1.1. WARFARIN THERAPY

In RCTs, the long-used vitamin K antagonist warfarin (Marevan®, Waran®) has been shown to reduce the risk of ischemic stroke by approximately two thirds and all-cause mortality by one-third for patients with AF compared to placebo and no treatment. It thus carries great potential for alleviating the complications that AF may entail[90,91]. However, warfarin therapy may be cumbersome for treatment providers and patients alike as continual monitoring and dose adjustment of the medication are necessary to maintain adequate effectiveness and safety in using the therapy. Warfarin therapy has a narrow therapeutic window, meaning that the difference between the effective and toxic dose is relatively small, and it therefore entails a nonnegligible risk of hemorrhagic adverse events. In addition, warfarin has a substantial number of drug-drug and drug-food interactions and its effect may be affected further by the genetics and concomitant disease of patients. Consequently, there are large intra- and interindividual differences in the anticoagulant effect of warfarin therapy, which necessitates the continuous routine monitoring and potential dose adjustments to maintain the therapeutic effect of the treatment and to avoid serious adverse events.[28,29,92] It follows that the effectiveness and safety of warfarin are also highly contingent on the achieved quality of therapy.

### 4.1.2. QUALITY INDICATORS IN WARFARIN THERAPY

The quality in warfarin therapy is often described by the use of the QI, time in therapeutic range (TTR)[93]. TTR refers to the percentage of time that patients' international normalized ratio (INR) of the blood is within the therapeutic range, that is, the target range where the optimal balance between safety and effectiveness is achieved[92]. The INR denotes the coagulability of the blood, where an INR of 0.8–1.2 indicates that the blood coagulates at a normal rate; with a higher INR, the coagulability is lower. In warfarin therapy, the optimal intensity of anticoagulation for the majority of AF patients obtains a therapeutic range of INR 2 to 3[28,92]. To maintain a therapeutic INR level for an individual AF patient, the interval between INR measurements is customized to allow adjustment of medication dosages when needed. Adjustments are based on the co-medication and -morbidity status and coagulation status at previous measurements of the patient. There are no evidence-based recommendations concerning the optimal interval between measurements. The average interval between INR measurements for the majority of patients is

approximately three weeks[67,76,94]. Guidelines recommend that the TTR should be calculated by linear interpolation between measurements of patients' INR levels, i.e. the Rosendaal principle[29,95]. Thus, the TTR is contingent on both the actual coagulation status of patients, i.e. the INR level, and the INR measurement control regimen. The TTR is often referred to as an intermediate outcome QI[93], although, with reference to the Donabedian framework[23], it could be argued that it is neither a true process nor outcome QI, as it depends on both biological functioning (INR) and the process of care (the level of coagulation control).

Nonetheless, the TTR is a well-established QI and demonstrates many of the characteristics of the ideal QI; it has the potential to be improved via interventions[52,67] and is generally accepted, not based on subjective valuation, measurable and replicable, evidence-based, and outcome-validated[20,67,93]. A negative relationship between the TTR and the rates of stroke, bleeding, and mortality has been established in multiple studies[96–99]. Hence, the longer a patient's INR is within the therapeutic level, i.e. the TTR is higher, the more the risk of thromboembolic and hemorrhagic complications is reduced. The importance of quality of therapy for the expected safety and effectiveness of warfarin therapy is emphasized by evidence suggesting that the TTR of patients should reach 58 percent to ensure a net clinical benefit. If a TTR of 58 percent cannot be maintained, the potential adverse events related to the therapy will counteract and supersede the benefits.[96] In alarming comparison, Mearns et al.[100] found that from 1990 to 2013 the mean TTR of AF patients in community settings and anticoagulation clinics was as low as 55 and 63 percent, respectively. Evidence, furthermore, suggests that the quality of therapy has not improved substantially over the last few decades[100,101]. Today the quality of warfarin therapy in Denmark is, in general, expected to be fairly good and better than in many other countries, with no substantial differences in the quality of therapy delivered in general practice and hospital-based clinics[102,103]. Nonetheless, in a Danish registry study on data from 1997 to 2011, Nissen Bonde et al.[104] found the median TTR for a large Danish AF population to be only 64 percent. Although the data are not completely up to date, arguably, the effectiveness and safety of warfarin therapy in real-life practice may still be enhanced today if the quality of therapy were improved.

Due to the multiple factors influencing the effect of warfarin, the ability to achieve a high TTR depends on a number of patient-related characteristics, including their sex, age, and medical history. Generally, patients with more comorbidities and poorer health are less likely to achieve a high quality in warfarin therapy as evaluated by the TTR. Consequently, not all patients are likely to be able to obtain a high quality of warfarin therapy.[57,104,105] In addition, the quality of warfarin therapy is contingent on the organizational setting in which the therapy is monitored and

managed. In general, monitoring at anticoagulation clinics is expected to entail a higher quality of therapy than community practice, and the quality of therapy delivered in relation to RCTs is generally expected to be yet higher[100,101]. Likewise, self-monitoring of warfarin therapy has been shown to increase patients' TTR compared to conventionally managed warfarin therapy, though only in highly heterogeneous groups[100,106]. This highlights the potential for quality improvement in warfarin therapy by, for example, organizational changes. However, the impact of clinical and organizational factors on the expected effectiveness and safety of warfarin also hampers the transferability of results between settings with dissimilar quality of therapy and patient populations.

The dependency of the TTR on the time between INR measurements introduces some uncertainty regarding the TTR as a reliable QI of warfarin therapy. TTR, via linear interpolation, does not necessarily reflect the true intensity of anticoagulation between the points of measurement. Particularly if the interval between INR measurements is large, the INR may fluctuate outside the therapeutic range between coagulation controls, placing patients at risk of complications. This would not necessarily be detected in the measurements, thereby complicating evaluation of actual quality in the therapy via the TTR. The deficiencies of TTR as a reflection of true quality in warfarin therapy has spurred the search for other potential quality measures in warfarin therapy, including the proportion of INR measurements within the therapeutic range, i.e. without interpolation, the variability in INR measurements, and patterns of anticoagulation control[107–109]. These measures may improve the prediction of adverse events in warfarin therapy compared to the TTR. However, currently, no other method for estimating the quality in warfarin therapy has gained the same currency as the TTR. TTR is the most often applied QI for warfarin therapy and is used for evaluation and setting standards of quality in warfarin therapy in clinical practice guidelines[29]. It also represents a reported measure of quality of therapy in research settings[39–42].

### 4.1.3. NON-VITAMIN K ORAL ANTICOAGULANTS

The disadvantages of warfarin therapy have spurred the development of NOACs, including dabigatran etexilate (Pradaxa®)[39], apixaban (Eliquis®)[40], rivaroxaban (Xarelto®)[41], and edoxaban (Lixiana®)[42], which within the last six years have become available alternatives to warfarin therapy in Denmark[29]. NOACs have the same function as warfarin, but take different mechanisms of actions. Hence, they do not exhibit the same number of drug-food and drug-drug interactions as warfarin,

making their anticoagulant effect more predictable. This enables the administration of fixed-dose medication that consequently reduces the need for frequent routine monitoring of coagulation related to warfarin therapy.[34,39–42] In a meta-analysis of the phase III trials on the four NOACs versus treatment with dose-adjusted warfarin, Ruff et al.[34] found the NOACs to have a preferable balance between efficacy and safety compared with warfarin. However, newer observational studies from Denmark indicate that, though the NOACs may have a slightly better safety profile compared to warfarin, in a routine care setting the effectiveness of the NOACs is in general only similar to that of warfarin[30,31].

Indirect comparisons have been made[37,110,111] that agree that the NOACs are more or less comparable, though some find apixaban to be slightly more efficient in attaining health[37]. In the RCTs on the NOACs, SSE was used for primary efficacy outcome and major bleeding (MB) was used for safety outcome[39–42]. As SSE and MB – fortunately – only occur relatively infrequently in AF, the sample size required to provide sufficient statistical power to establish noninferiority, let alone superiority, is considerable. The patient populations in the RCTs on the NOACs versus warfarin were sizeable, ranging from 14,264 to 21,105 included patients [41,42] with median follow-ups ranging from 1.8 to 2.8 years[40,42]. The quality in warfarin therapy as evaluated by the median TTR of individual patients, furthermore, ranged from 58 percent to 68 percent[41,42] in the trials. As the relative efficacy and safety of NOACs are contingent on the quality of warfarin therapy, transfer of the results from the RCTs to settings with dissimilar therapy quality should be performed with caution. It could be hypothesized that the relative benefits of the NOACs would be challenged if they were compared to truly high-quality warfarin therapy[112].

Within the last five years, the use of NOACs has increased in Denmark and internationally, whereas the use of warfarin has declined. NOACs may be preferred to warfarin therapy, due to the slightly improved efficacy and safety profile, but also likely due to the greater convenience that these medication options provides via the fixed-dose administration with the reduced need for monitoring. An added benefit of this may be increased patient adherence to OAC.[29,32,113] One disadvantage of the NOACs remains the medication costs, which for all the NOACs are more than five times higher per defined daily dose than the costs of warfarin[35].

## 4.2. GUIDELINES ON STROKE PROPHYLAXIS THROUGH THE USE OF ORAL ANTICOAGULANT THERAPY

Throughout the last decade, clinical practice guidelines for the management of AF have developed rapidly[28,29,114,115]. In 2012[114] and 2016[28], the European Society of Cardiology (ESC) published updated clinical practice guidelines on pharmacological stroke prophylaxis in AF with a more inclusive treatment strategy rendering more patients eligible for OAC than previously recommended. The Danish clinical practice guidelines on OAC in AF, published by the Danish Council for the Use of Expensive Hospital Medicines [Rådet for Anvendelse af Dyr Sygehusmedicin; RADS; now the Medicines Council [Medicinrådet]], are in general agreement with these guidelines[29,115].

According to the guidelines, the decision on whether to initiate OAC should depend on patients' risk of SSE, which can be stratified by the use of the $CHA_2DS_2$-VASc (**c**ongestive heart failure, **h**ypertension, **a**ge≥75 [doubled], **d**iabetes, **s**troke [doubled], **v**ascular disease, **a**ge 65–74, and **s**ex (female)) scoring system. In the $CHA_2DS_2$-VASc scoring system, each present risk factor yields one point, unless doubled[28,81]. The components of the risk stratification scheme represent clinical risk factors for SSE, and with higher scores, the risk of SSE increases[81]. Patients with a $CHA_2DS_2$-VASc score of zero, i.e. zero points, would not benefit from OAC[84] and should, according to the guidelines, not receive the treatment, whereas patients with a score of 1+ for men and 2+ for women are likely to benefit from the therapy and should be treated[28,29]. Considering the relatively high prevalence of the risk factors for SSE in an elderly population[116,117], it is hardly surprising that the majority of AF patients are considered eligible for OAC[81]. Naturally, clear contraindications against OAC should be considered when evaluating the need for therapy. A high risk of bleeding is not a contraindication against OAC per se, but should make apparent the need to identify the optimal OAC with a low risk of bleeding[28–31]. Some risk factors for ischemic stroke and bleeding overlap, including a higher age, hypertension, and prior stroke. Relevant risk factors for bleeding are summarized in the HAS-BLED (**h**ypertension, **a**bnormal hepatic and/or renal function [if both; doubled], **s**troke, **b**leeding, **l**abile INR; TTR<60%, **e**lderly; age>65, **d**rugs; influencing coagulation and/or alcohol abuse [if both; doubled]) scoring system[29,118].

A change in the Danish clinical practice guidelines from RADS has been introduced from the 2013 edition to the 2016 edition. In 2013, all NOACs were assumed to have similar effectiveness and safety profiles and were ranked alongside each other whereas, in 2016, RADS no longer equated dabigatran etexilate with the other

NOACs due to an apparent increased incidence of adverse events and more drug-drug interactions compared to the other NOACs[29,115]. This distinction between the NOACs is not replicated in the guidelines from the ESC nor in guidelines from the Danish Society of Cardiology[28,119]. Thus, in 2016, RADS equated apixaban, rivaroxaban, and edoxaban with warfarin, contingent on an expected compliance with an *a priori* explicit, empirical quality standard for warfarin therapy of a TTR of a minimum of 70 percent for treated patients. Based on an expectation that the TTR for the majority of patients will not reach the quality standard for warfarin therapy, it is, rather arbitrarily, recommended that 75 percent of new patients should begin treatment with NOACs and 25 percent should begin warfarin therapy. Furthermore, patients who do not currently achieve a TTR of 70 percent should be switched to a NOAC instead.[29]

In the clinical practice guidelines from RADS, it is furthermore commented that there is a lack of randomized head-to-head studies in which the NOACs are compared with high-quality warfarin therapy, i.e. where the mean TTR is over 70 percent[29]. However, it does not seem very likely that such studies will be performed in real life, considering the resources – and patient population – that would be required to enable the execution of such studies.

## 4.3. ROOM FOR IMPROVEMENT

Clinical practice guidelines summarize the medical knowledge about what is currently regarded as the optimal treatment of patients, thereby providing an aid for clinicians to make evidence-based decisions regarding optimal healthcare delivery to patients. Adherence to the current clinical practice guidelines on stroke prophylaxis in AF, as established by a consistency between patients' risk factors for SSE and subsequent observed treatment, has been shown to improve patient outcomes with respect to mortality and the incidence of thromboembolism and MB[120,121]. Guidelines adherence thus represents an optimization of treatment, when the purpose is to achieve the greatest health gain possible.

With the introduction of NOACs, greater awareness has been raised about the appropriate treatment of patients with AF and, consequently, more AF patients in Denmark as well as internationally are now treated prophylactically against stroke[32,33,122]. In a Danish registry study, Mikkelsen et al.[117] found that in the period from 2002 to 2011 only approximately 50 percent of patients newly diagnosed with AF in inpatient and outpatient settings were started on OAC within 180 days

after the initial diagnosis[117]. By 2011, in a Danish cross-sectional survey Brandes et al.[116] found that 66 percent of AF patients who started on antithrombotic therapy in general practice were treated with OAC, corresponding to 75 percent of the patients being treated in accordance with the 2012 ESC guidelines. Based on this, Brandes et al. concluded that the quality of antithrombotic therapy offered to patients with AF in general practice in Denmark is reasonably high.[116]

Despite this, nonadherence to the clinical practice guidelines on stroke prophylaxis in AF is still a common difficulty. This leads to patients being un-, under-, or overtreated, thereby placing them at unnecessary risk of thromboembolism and bleeding, respectively. For instance, acetylsalicylic acid is now considered an obsolete thromboprophylactic agent in AF as its benefits do not match those of OAC, but it entails the same risk of adverse events[28,91]. Nonetheless, it is still used in clinical practice and represents a common undertreatment of AF patients[33,122]. Contemporary, international evidence indicate that a substantial proportion – up to 60 percent – of AF patients receive treatment that is not in agreement with the current clinical practice guidelines and that elderly patients in particular remain undertreated.[33,116,120–122] Thus, there is still room for improvement.

The lack of guideline adherence may have multiple causes. In addition to considerations of the efficacy and safety of therapies, the clinical practice guidelines on stroke prophylaxis in AF also encourage incorporation of patient preferences to increase patient satisfaction and persistence with treatment[28,29]. However, the level of guideline adherence is often evaluated by a comparison of the clinical indications of patients to the subsequent, observed choice therapy for them irrespective of patient preferences and satisfaction[33,120,122]. Evidence suggests that up to 30 percent of AF patients disprefer OAC despite a clinical indication, which may lead to nonadherence to guidelines[123]. If treatment of an individual patient strays from the recommendations, it might be due to regard for patient preferences that oppose the clinically indicated therapy.[74,123,124] Patient aversion to OAC, furthermore, manifests itself in an often-poor persistence with therapy, which is a prevailing problem in OAC. Thus, a substantial proportion of AF patients newly started on OAC discontinue therapy within the first year of treatment.[32,125–127] Furthermore, patients who are not treated with OAC when clinically indicated tend to be older and suffer from multiple comorbidities. Nonadherence to clinical guidelines may occur as a result of an aversion amongst clinicians to prescribing OAC that might induce iatrogenic, hemorrhagic complications in already fragile individuals[33,120,122]. This may be particularly true for patients perceived to be at high risk of bleeding, such as frail patients, patients at risk of falls, or whose who are clinically disposed for bleeding (cf. the HAS-BLED scoring system)[28,74,123].

The routine monitoring and difficulties in keeping therapy safe and effective may be explanatory factors for potential poor adherence and persistence with warfarin therapy[124,127]. These nuisances may be alleviated by the use of NOACs, and studies indicate that therapy persistence is improved by using these agents[32,126]. Patient-related lack of adherence to, and persistence with, OAC when clinically indicated may, however, also be due to ignorance of the risks related to nontreatment, which might be ameliorated by improved communication between clinicians and patients[74,113].

Stroke prophylaxis in AF, thus, still holds substantial potentials for improvement. Care of this patient population as a whole may be improved by multiple approaches. Thus includes increasing the proportion of patients who are initially prescribed clinically appropriate therapy and by increasing patient adherence to, and persistence with, therapy. Targeting these elements would likely diminish the gap between the efficacy frontier of OAC in AF and the observed effectiveness in real life and might potentially present cost-effective interventions to the benefit of patients[128,129].

# CHAPTER 5.
# FIELD OF RESEARCH

This chapter includes summaries of the appended papers and a contextualization of the studies into the frame story of the present dissertation. The implications of the results of studies II–IV[2–4] are furthermore discussed in a clinical and decision-oriented context.

The following section, 5.1, summarizes study I[1], which presents a contribution as to how the methods for economic evaluation might be applied to estimate the cost-effectiveness of quality improvements despite a lack of patient-relevant outcomes. Sections 5.2 and 5.3 include contextualization of studies II[2] and III[3] within the framework presented in section 5.1. They are used to illustrate examples of hypothetical QIs that might be applied to research in interventions intended to improve stroke prophylaxis in AF. In section 5.4, study IV[4] exemplifies the establishment of a correlation between an outcome-validated QI, the TTR, and patient-relevant outcomes, SSE and MB. It serves as an illustration of the importance of finding an appropriate model for the correlation between QIs and patient-relevant outcomes and the importance of including relevant antecedents for analysis.

The overriding purpose of this chapter is to highlight the opportunities for applying economic evaluation to quality improvement along with the obstacles to doing so. in In Chapter 6, this will lead to a discussion of the overall potentials for doing this and a proposal of considerations that should be undertaken prior to the application of economic evaluation to quality improvements to optimize the research process.

# 5.1. THE VALUE OF QUALITY INDICATORS FOR INFORMED DECISION-MAKING

The present section pertains to study I[1] presented in Appendix A.

## 5.1.1. SUMMARY OF STUDY I

Study I advances a framework built on Bayesian decision theory and VOI analysis. It concerns the application of DAM to enable evaluation of the NB of quality improvements when patient-relevant outcomes are not available and is intended to support economic evaluation when QIs are used for indication of effect. The aims of the study were to 1) investigate the requirements for QIs to be used in economic evaluation and 2) investigate under which circumstances they may be used to enhance informed decision-making.[1]

DAM enables estimation of the expected NB of interventions by utilizing the mathematical relationship between the elements in the model. It thus may facilitate the inclusion of QIs in the relationship chain between interventions and NB. The use of a QI introduces an extra, intermediate link in the relationship between interventions and their expected NB, partitioning the correlation between interventions and impact on patient-relevant outcomes into two separate parameters: $\varphi_1$ and $\varphi_2$ (Figure 5-1). The following requirements were identified for QIs to be useful in economic evaluation:

- The intervention should be expected to ultimately affect patient-relevant outcomes (i.e. a correlation, $\varphi_{tot}$, exists between the intervention under investigation and patient-relevant outcomes)
- A quantifiable correlation, $\varphi_1$, exists between the intervention and the QI
- A quantifiable correlation, $\varphi_2$, exists between the QI and patient-relevant outcome(s)

The correlations $\varphi_1$ and $\varphi_2$ may be included in DAM as distributions, reflecting the uncertainty pertaining to them. If a global correlation $\varphi_2$ between a QI and patient-relevant outcome(s) is established, this may subsequently be used to estimate the NB of more interventions and studies, in which the QI is used as effect measure and a distribution of $\varphi_1$ is established. To contain the uncertainty related to the correlation multiple estimations of the 'global' correlation parameter $\varphi_2$ may be necessary for patient subgroups with expectedly different risk profiles for the patient-relevant outcomes. As the uncertainty pertaining to the parameters $\varphi_1$ and $\varphi_2$ is propagated

through the model (Figure 5-1), stronger correlations, i.e. less parameter uncertainty of $\varphi_1$ and $\varphi_2$, will diminish the impact on final decision uncertainty. When possible, relevant antecedents should be adjusted for when $\varphi_2$ is specified[67]. This may aid in containing the uncertainty of the parameter and ensure better reflection of the patient population under investigation. If the correlation $\varphi_2$ is misspecified, it may cause incorrect estimations of the expected NB of interventions, leading to erroneous conclusions regarding the cost-effectiveness of interventions.[1]

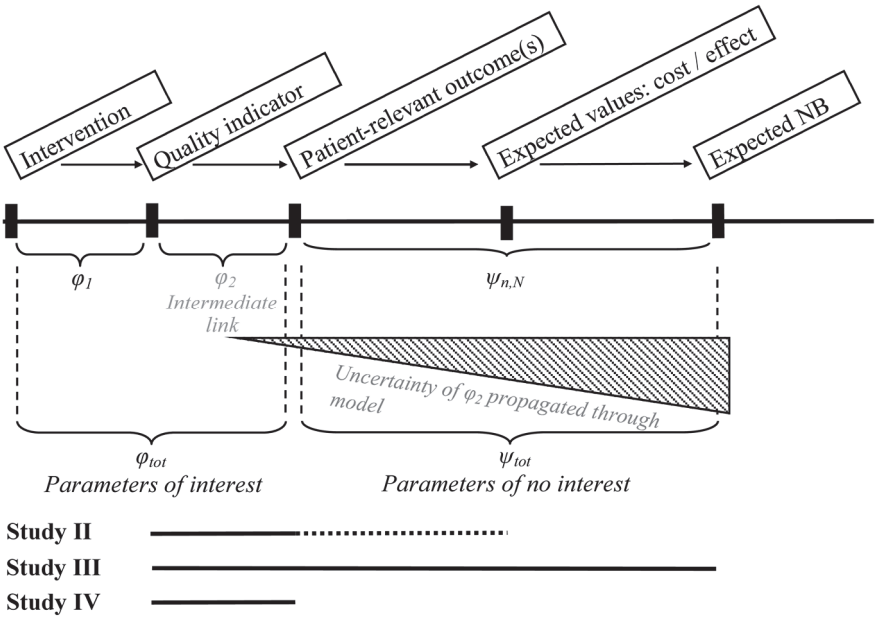



*Figure 5-1. Inclusion of quality indicator(s) as an extra parameter in the relationship between an intervention and its final expected NB.*
*In study II, the correlation $\varphi_2$ between a hypothetical quality indicator and the occurrence of patient-relevant outcomes is estimated, including the impact on the cost of the patient-relevant outcomes. Study III exemplifies how the cost-effectiveness of an intervention may be estimated based on hypothetical a quality indicator and subsequent decision-analytic modeling. Study IV exemplifies the establishment of the correlation parameter, $\varphi_2$, between an acknowledged quality indicator and patient-relevant outcomes. NB: Net benefit. Adapted from Vestergaard and Ehlers*[1].

The use of DAM enables explicit evaluation of parameter uncertainty in the model by PSA and ensuing VOI analysis including calculation of the potential cost of uncertainty pertaining to the use of the QI, i.e. the expected value of perfect parameter information of $\varphi_2$. Taking a Bayesian decision theoretic approach, it should be considered whether the decision uncertainty related to the NB estimated from the use

of the QI indicates that further research might be needed to decrease the risk of making the 'wrong' decision. The cost of uncertainty of the different parameters in the model should be compared to the expected cost of resolving the uncertainty. This entails considerations of what comprises a sufficient data foundation for decision-making – whether research to inform on $\varphi_{tot}$ is warranted to resolve the uncertainty pertaining to the basis for a decision, otherwise provided by the NB estimated through the use of a QI ($\varphi_2$).[1]

### 5.1.2. ELABORATION OF STUDY I

The use of QIs as presented in the framework in study I enables the estimation of the NB of quality improvements when evidence on the impact on patient-relevant outcomes is not available, a prevalent issue for quality improvements. Studies in quality improvement are rarely powered to detect differences in patient-relevant outcomes [27,67], yet the need to be able to establish the value for money of quality improvements stands unabated and is likely to increase in the future. The focus in study I and this dissertation is on the requirements to QI to be applicable in DAM and the establishment of the correlation $\varphi_2$. Less attention is on the establishment of the correlation parameter $\varphi_1$, which depends, amongst other things, on the study design used in research in quality improvement. This comprises a topic for further research outside the scope of the present thesis.

The framework builds on the NB framework, which necessitates the existence of a threshold for opportunity costs (cf. section 3.3). This, de facto, entails the utilization of QALYs as the measure of effectiveness, as a generally accepted threshold exists only for this outcome[48]. Consequently, utilization of the framework presented in study I requires acceptance of the assumptions that underlie the use of QALYs.

The chosen QI for a particular model should be expected to embrace all impact on health imposed by the intervention. Alternatively, more QIs may be applied, but this necessitates their impact on health being mutually exclusive in order to avoid overestimation of effect. Once a relevant QI is identified, correct specification of the $\varphi_2$ is pivotal to avoid misleading estimations of NB. VOI analysis only illustrates the uncertainty of the parameters and not their correctness. Consequently, VOI analysis cannot be used to uncover an erroneous $\varphi_2$ specification, emphasizing the need for careful consideration when it is specified. If the 'distance' between the applied QI and patient-relevant outcome is substantial, more antecedents may influence the correlation $\varphi_2$ and inflate the uncertainty pertaining to it, causing mistrust in the results

of analyses. Antecedents may induce effect modification and confounding, which stresses the need for adjustment. If not all relevant influences are recognized and included, the model suffers the risk of unobserved confounding. For this reason, it is important to identify potential antecedents and, if possible, adjust for their influence. To achieve this, a good overview of potential causal relationships and correlations between elements in the research setting under study is required. When possible, the accuracy of the correlation between the QI and patient-related outcomes, $\varphi_2$, should be validated against real-life data prior to its application.

The framework presented in study I reflects the same approach to decision-making under uncertainty as observed for economic evaluation incorporating clinical, intermediate endpoints. A viable clinical, intermediate endpoint should constitute a substitute for final, patient-relevant outcomes and be predictive of these For the use of clinical, intermediate endpoints in economic evaluation it is, likewise, suggested that a quantifiable and stable relationship between the endpoints and patient-relevant outcomes should exist and all uncertainty related to assumptions and parameter inputs should be highlighted. Likewise, VOI analyses are recommended.[13,53] One principal difference between the framework presented here and the existing recommendations on use of clinical, intermediate endpoints for DAM consists in the interpretation of what may constitute the intermediate link in the relationship between interventions and NB. Research in clinical, intermediate endpoints takes its origin in EBM and clinical, intermediate endpoints are often biomarkers or indicators of physiological changes known to eventually affect health.[53] The recommendations thus embraces only outcome QIs. In contrast, the present framework accepts the incorporation of any links not necessarily related to physiological functioning, i.e. also structural process QIs, in the relationship between interventions and NB under the expectation that the uncertainty it entails is sufficiently illustrated. It is likely that some of the experiences of the incorporation of clinical, intermediate outcomes in DAM may be transferred to economic evaluation of quality improvements where QIs are used for outcome.

## 5.2. LINKING QUALITY INDICATORS TO PATIENT-RELEVANT OUTCOMES

The present section pertains to study II[2] presented in Appendix B.

### 5.2.1. SUMMARY OF STUDY II

Clinical practice guidelines recommend that AF patients resume OAC following an ICH unless contraindicated[28,114]. The recommendation is based on the expectation that this treatment course, on average, provides a greater net clinical benefit than discontinuation, as long as relevant risk factors for bleeding and the cause of the initial ICH have been sufficiently treated[28]. The primary aim of study II was to estimate the three-year hospitalization cost (2015 value) related to thromboembolism and MB for two patient groups with AF who survived the first 90 days following an ICH: 1) patients who resumed warfarin therapy within the 90 days following the ICH and 2) patients who did not resume therapy within this period.[2]

The study was performed as a registry study based on the nationwide Danish registries, providing information on patients' baseline characteristics, diagnoses, hospitalizations, and medications. The effect of warfarin therapy resumption on three-year hospitalization costs and survival was estimated by the use of regression analysis with adjustment for between-group differences in baseline characteristics. The marginal effect of therapy resumption on mean hospitalization costs was estimated through the use of a two-part model, comprising a logistic regression model and a generalized linear model. The impact on survival was evaluated through the use of the Cox proportional hazard ratio. All analyses were adjusted for the individual components of the $CHA_2DS_2$-VASc and a modified HAS-BLED score at day 90 post-ICH. The study design is illustrated in Figure 5-2.[2]

When adjusting for differences in baseline characteristics, the odds ratio for hospitalization of patients resuming warfarin therapy was 0.92 [95%CI: 0.65;1.31] compared to patients who did not resume therapy. The marginal effect of therapy resumption on hospitalization costs for hospitalized patients was DKK-12,101 [95%CI: -22,289;-1913]. When combining the models, the marginal effect of therapy resumption on hospitalization costs was DKK-3101 [95%CI: -6211;15]. The baseline-adjusted hazard ratio for mortality was 0.79 [95%CI: 0.61;1.01] for patients resuming therapy versus patients not resuming therapy.[2]

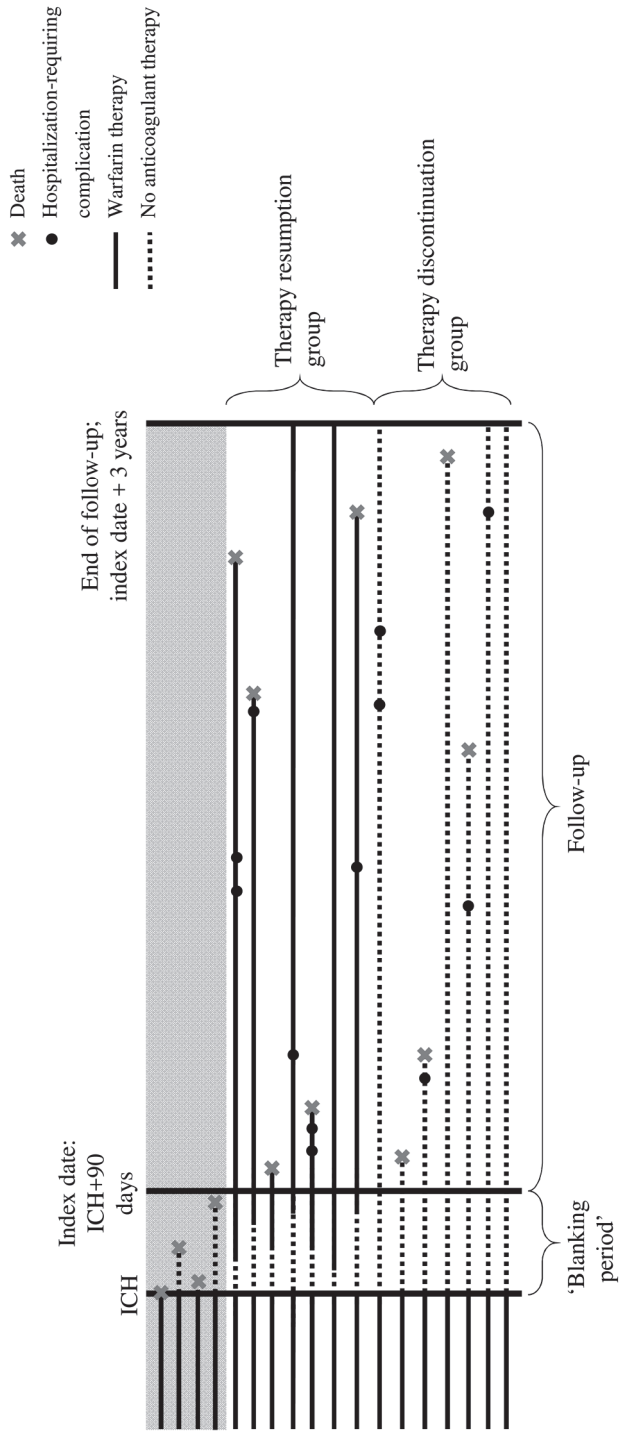



*Figure 5-2. Illustration of the study design of study II.*

### 5.2.2. CONTEXTUALIZATION OF STUDY II

Trial-based evidence on the effectiveness and safety of OAC following ICH in AF is scarce as prevalent ICH is a common reason for excluding patients from trials, including the major trials on the NOACs[39–42]. However, the results of study II are in agreement with other observational studies indicating that the resumption of warfarin therapy following an ICH in AF may represent the optimal treatment course for the majority of AF patients[130,131]. Thus, evidence suggests that patients resuming therapy experience a higher survival and fewer hospitalization-requiring complications. Study II provides both a clinical and financial incentive to motivate guideline adherence, i.e. OAC resumption after ICH. The study does, however, most likely suffers from unobserved confounding, including confounding by indication. The Danish registries do not hold information on the subtype of the ICH, which may decide for whether OAC resumption is expected to be beneficial or not[132–134]. This information may have been a deciding factor in whether to resume therapy or not. Likewise, patients' health status and preferences towards treatment after the ICH is unknown, which may have influenced the decision.[2]

As resumption of OAC after bleeding is generally recommended in guidelines[28], OAC resumption could, hypothetically, constitute a QI that might be used to evaluate the quality of care of patients with AF. In that case, it would be recognized as a process QI referring to the initial decision on whether to resume therapy or not. It would be a binary indicator of 'resumption of therapy following ICH, yes/no' and accordingly, not indicate the level of quality of subsequent continuous therapy of patients. Study II confirm that there is a (nonsignificant) correlation between resumption of OAC and patient-relevant outcomes ($\varphi_2$) and the costs related to them, thereby validating it as a potential, though uncertain, QI. It furthermore illustrates how the correlation parameter, $\varphi_2$, between the QI and the outcomes may be estimated by using regression analysis with adjustment for *a priori* chosen covariates – or antecedents – known to affect the occurrence of patient-relevant outcomes and potentially also the initial decision to resume therapy. The occurrence of patient-relevant outcomes may, however, also be affected by a multiplicity of factors after the initial resumption decision, e.g. patient nonadherence to therapy. These cannot and should not be adjusted for in analysis. These factors may be contributory to the inability to establish a statistically significant relationship between the process QI and the occurrence of thromboembolism and MB.

If therapy resumption after ICH were to be used as a QI, it might not make sense to use it as an indicator of quality of therapy for the individual patient, as a number of patient-specific elements may influence the decision as to whether to resume therapy.

Instead, it might be used for patient cohorts as the 'proportion of AF patients resuming therapy following ICH.' It may, however, prove difficult to set quality standards, e.g. a minimum proportion of patients that should resume therapy, as the decision to resume therapy or not should reflect the clinical situation at hand[131].

Study II exemplifies the difficulties in establishing statistical significance that may arise when studying relatively rare conditions and the occurrence of relatively rare complications in these. Even with one of the largest database sources available it was not possible to generate statistically significant results when applying unaggregated outcome measures[2]. Small samples and very infrequent occurrence of patient-relevant outcomes are, likewise, a prevalent issue in research on quality improvement. Nonetheless, decisions must be made, which supports a Bayesian approach to decision-making with less focus on statistical inference and a greater focus on highlighting the potential cost of uncertainty related to decision-making under uncertainty.

Although both the hospitalization costs and mortality were lower for the group resuming OAC, the health economic potential of the intervention, resumption of therapy, remains obscure. As stated in the *discussion* in study II[2], the savings in hospitalization costs imposed by therapy resumption would not counteract the costs related to the delivery of warfarin therapy and, arguably, neither the applied perspective nor the time horizon enables the inclusion of all relevant costs. In addition, the use of the one-dimensional effect measure, survival, likely also underestimate the impact on health under the two treatment courses, given that thromboembolism and MB negatively influence HRQoL. Although a tendency towards dominance of therapy resumption was observed in study II, i.e. lower costs and improved survival, DAM would be warranted to enable the acquisition of interpretable cost-effectiveness results in a broader perspective. It is expected that the benefits of therapy resumption would be augmented in such an analysis.

## 5.3. MODELING VALUE FOR MONEY BASED ON A QUALITY INDICATOR

The present section pertains to study III[3] presented in Appendix C. During her master thesis, the author of the present dissertation developed a preliminary model for the decision-analytic model applied in study III. As a part of the Ph.D., the model has ensuingly been refined for publication in study III[3].

### 5.3.1. SUMMARY OF STUDY III

In the clinical practice guidelines on pharmacological stroke prophylaxis in AF published by the ESC in 2012[114] a larger proportion of AF patients were considered eligible for OAC than previously estimated. Study III evaluated the cost-effectiveness of optimizing therapy of the Danish AF patient population by complete adherence to the 2012 ESC clinical practice guidelines compared to the treatment strategy observed prior to 2012.[3]

A CUA was performed to compare the guideline-adherent treatment strategy to the strategy observed prior to 2012. Guideline-adherent treatment included the use of a NOAC (dabigatran etexilate), warfarin, and no treatment; the comparator included warfarin, acetylsalicylic acid, and no treatment. A Markov model was designed to reflect the expected courses of treatment under the two strategies, including potential complications arising from the different treatments. A patient cohort was modeled with a risk profile reflecting that of the Danish AF patient population as evaluated by patients' $CHA_2DS_2$-VASc scores, established from register-documented conditions. The $CHA_2DS_2$-VASc scores were used to allocate patients into different therapies under the guideline-adherent treatment strategy. The model was populated with information on input parameters from the literature. The data were chosen based on the criterion that they should expectedly best reflect the Danish context and local cost data (2014 value). The applied perspective was that of the Danish healthcare sector, including patient-paid cost of medication. The model ran over a 10-year time horizon with a three-month cycle length.[3]

Based on the decision-analytic model, the ICER for the guideline-adherent treatment strategy versus the strategy observed prior to 2012 was approximately DKK 26,500 per QALY, which would be well below a potential threshold for opportunity costs of DKK 185,000 per QALY (≈£20,000/QALY). Both deterministic sensitivity analyses and PSA indicated that the result was largely robust to reasonable changes in assumptions and input parameters.[3]

### 5.3.2. CONTEXTUALIZATION OF STUDY III

Study III presents an alternative approach to improving the health of a patient population by comparing treatment strategies including the utilization of multiple medication types instead of comparing two or more individual treatments as is more common in economic evaluation. The cost-effectiveness of the individual medications has been established previously[36–38], for which reason it was not the main concern of the study. The purpose of clinical practice guidelines is to indicate what is clinically optimal. Hence, they do not explicitly include considerations of cost-effectiveness of neither the individual medications nor the treatment course they recommend. Therefore, the focus was on identifying the cost-effectiveness of adhering to the guidelines, as this would represent an improvement of the quality of therapy. It should be noted that though the pre-2012 treatment strategy represents suboptimal treatment by current standards, it might reflect acceptable treatment at the time when it was provided, although this is unlikely[23]. However, if the pre-2012 strategy also reflects current treatment practice, this would represent suboptimality, which could be ameliorated by enhanced guideline adherence.

Whereas economic evaluation is often applied to identify the incremental cost-effectiveness of utilizing one treatment instead of another, study III also incorporates the impact of treating a larger proportion of the patient population in accordance with guidelines. In the pre-2012 strategy, only 37 percent were treated with warfarin, 19 percent were treated with acetylsalicylic acid, and 44 percent were untreated, which, by current standards, reflects a substantial undertreatment. In the guideline-adherent treatment strategy, only 9 percent would be appropriately untreated, based on their $CHA_2DS_2$-VASc score. This illustrates that optimizing prophylaxis for this patient population by treating a larger proportion of the AF patient population with OAC is highly cost-effective (DKK26,500/QALY); irrespective of the fact that medication costs would increase. Sensitivity analyses, furthermore, indicated that the cost-effectiveness was not contingent on what subtype of OAC, warfarin or NOAC, was utilized. Study III thus indicates that it may be clinically and health economically beneficial to improve the quality of therapy for the patient population as a whole by focusing not only on which intervention to use, i.e. warfarin or NOACs, but also for whom. That is, identifying all eligible patients and treating them according to current medical knowledge.

To enable the construction of the decision-analytic model, a number of simplifications of reality were necessary, which are open to discussion. The Markov model reflecting the guideline-adherent treatment strategy is built on the assumption that patients remain treated throughout the investigated time horizon, which in real life may not be

realistic. As the patient population ages, conditions may furthermore arise, which are not included in the model, that preclude the continuation of OAC. This is not included in the model. In addition, the distribution into different therapies under the guideline-adherent treatment strategy is solely based on the risk profile, the $CHA_2DS_2$-VASc distribution, of the Danish patient population and does not take into account potential contraindications for OAC, nor patient preferences.

Hypothetically, treatment of patients in accordance with clinical practice guidelines could be used as a process QI for treatment in AF, reflecting continuous adherence to guidelines over time. The correlation $\varphi_2$ between the QI and patient-relevant outcomes is established implicitly via the use of medications, for which the relationship with the outcomes has been established previously. Thus, study III illustrates how the cost-effectiveness may be established based on a QI, here the binary QI, 'treated according to guidelines, yes/no.' Given the model structure presented in study III[3], this potential QI cannot be used for the evaluation of quality of therapy of the individual patient, but may reflect quality of therapy for cohorts. As such, it might be employed as the 'proportion of patients treated according to guidelines over time.' In this, the correlation, $\varphi_2$, between the QI, via utilization of relevant medications, and patient-relevant outcomes is used to estimate the potential cost-effectiveness of perfect guideline adherence by conditional linking of the consequences for costs and effects that perfect guideline adherence incurs (Figure 5-1). The decision-analytic model in study III does not explicitly incorporate the potential influence of antecedents. However, the evidence used to populate the model was intended to reflect the Danish patient population, for which reason the correlation between the two strategies, through the use of the various medications, and patient-relevant outcomes is expected to represent what would be observed in real life.

In practice, 100 percent guideline adherence for cohorts is unobtainable, but the model could be expanded to evaluate the cost-effectiveness of interventions seeking to improve guideline adherence. To include an intervention that influences the QI, i.e. 'proportion of patients treated according to guidelines over time', in the model, the additional uncertainty parameter $\varphi_1$ should be included to the left in the model structure presented in study III (Figure 5-3). This would enable indication of expected proportions of patients treated according to the 2012 ESC clinical practice guidelines and the pre-2012 treatment strategy, respectively, with and without an intervention to improve guideline adherence. This model structure also enables inclusion of the expected costs related to the intervention under investigation (Figure 5-3). The same structure could be set up to include an intervention and $\varphi_1$ for the hypothetical QI presented in study II.

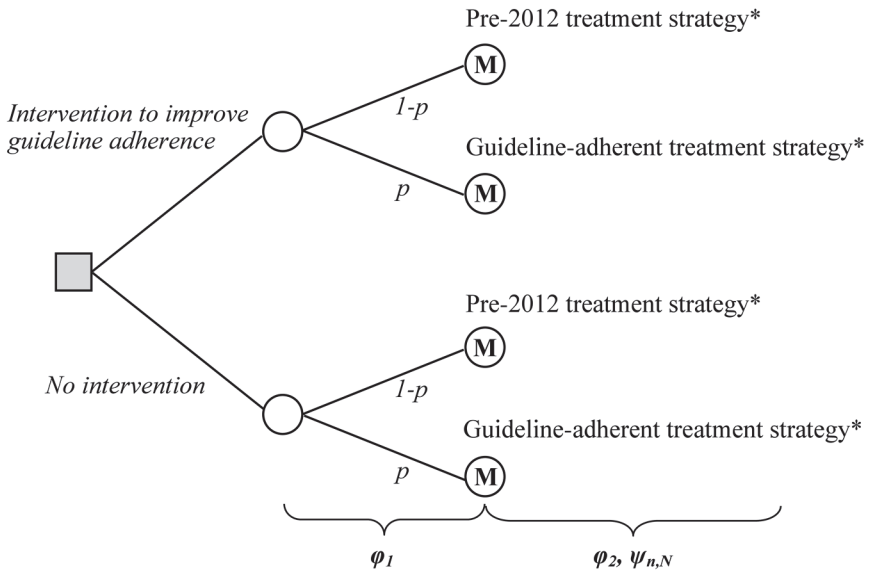



*Figure 5-3. Illustration of the introduction of a hypothetical intervention to improve guideline adherence. In all probability, 100 percent guideline adherence cannot be achieved, but it may be increased by an intervention (p in the upper branch of the decision tree). The costs related to the intervention should be included in the model. *Confer study III*[3] *for explication of the Markov models.*

## 5.4. EMPIRICAL INVESTIGATION OF A QUALITY INDICATOR

The present section pertains to study IV[4] presented in Appendix D.

### 5.4.1. SUMMARY OF STUDY IV

As previously explicated, TTR is a generally accepted QI in warfarin therapy and a negative linear relationship between TTR and complication rates has been demonstrated[67,96–99]. Increasing TTR therefore represents a quality improvement by decreasing the risk of serious adverse events. The occurrence of thromboembolism and bleeding is, however, also contingent on patients' risk profile, often summarized by the $CHA_2DS_2$-VASc and HAS-BLED scores. Previously, patients' risk profile has not been taken into account when the correlation between TTR and complications has been estimated. Evidence, furthermore, suggests that the relationship between TTR

and complications may not be completely linear, as increasing the TTR at lower levels of TTR reduces the risk of complications, whereas this effect is not observed for higher TTR levels, i.e. at TTRs above 70 percent[135,136]. Combined, this calls into question the accuracy of the previously established relationship. The aim of study IV was to evaluate the correlation between the mean TTR of AF patient cohorts and the occurrence of patient-relevant outcomes with adjustment for patient cohort characteristics known to affect patients' risk of events.[4]

The data used for study IV were retrieved through a focused, structured literature review. Thirty-five papers with information on 31 patient cohorts, including information on more than 100,000 AF patients, were extracted based on pre-specified criteria. A linear random-effects meta-regression with and without adjustment for *a priori* chosen covariates was performed to illustrate the association between mean TTR in the study cohorts and the occurrence of SSE and MB. Supplementary analyses were performed for hemorrhagic and ischemic stroke, respectively. Analyses of the hemorrhagic outcomes were adjusted for the mean age and proportion of the population with previous stroke or transient ischemic attack. For analyses of the ischemic outcomes, the proportion of the population that was female was, furthermore, included. A modified double arcsine transformation was applied to the outcome data for normalization, to avoid overdispersion, and to stabilize variances.[4]

As expected and previously shown, in the univariable meta-regression with mean TTR as single predictor, higher mean TTR was statistically significantly and negatively related to the double arcsine transformed rates of both MB and SSE. In adjusted analyses, mean TTR was still statistically significantly related to the rate of MB, but not with the rate of SSE. Study IV indicates that increasing the quality in warfarin therapy by increasing the mean TTR of a patient cohort is clinically beneficial. However, when adjusting for clinically relevant characteristics of the cohorts, increasing quality in therapy mainly improves the safety of therapy by decreasing the MB rate, whereas the rate of SSE, though still inversely associated with mean TTR, is only negligibly decreased.[4]

### 5.4.2. CONTEXTUALIZATION OF STUDY IV

Other potential QIs for warfarin therapy exist that could have formed the basis for evaluation in study IV, some of which may represent better QIs than TTR (cf. section 4.1.2). However, as TTR is generally accepted as a QI in warfarin therapy, greater knowledge about it and its association with complications may potentially be used to

aid decision-making. When taking a pragmatic, application-oriented approach, the widespread and acknowledged use of TTR merits its application in the present study. Furthermore, TTR is often reported in studies, which provided a larger evidence base for the study than would have been possible for other potential QIs. To the best of the author's knowledge, no real-life studies have been performed with the specific purpose of establishing the relationship between mean TTR and patient-relevant outcomes. Consequently, the health economic impact of increasing quality in warfarin therapy as measured via mean TTR is currently unknown. For this reason, in study IV evidence was compiled from the literature in an attempt to establish the correlation. The ulterior motive for the conducting of the study was to make the estimated correlation(s) accessible for further use in the framework presented in study I to subsequently enable economic evaluation of interventions intended to increase TTR.

A pragmatic approach was taken to the study design, balancing the desire for the right evidence with the need for sufficient evidence to support the analyses. The primary outcomes investigated in the study, MB and SSE, are largely suboptimal in a health economic context as the related costs and effects of the aggregated outcomes may be highly dissimilar. This may be especially true for the outcome MB, which includes ICH that on average entails more severe consequences for patients' expected disability level, costs, and very survival compared to extracranial hemorrhage. Preferably, from a health economist perspective, the study should have included less aggregate outcomes, but the evidence base on separate outcomes was too small to substantiate valid analysis. Furthermore, the current adjustments in the multivariable meta-regressions were performed based on the availability of the data. The majority of the included studies did not report on all parameters of interest and in order to keep a sufficient evidence base for analysis, adjustment was only performed for basic characteristics of the patient populations. [4] Ideally, the analyses might have included adjustment for more factors known to influence the occurrence of outcomes. As it is, study IV likely suffers from some effect modification and unobserved confounding. Multiple assumptions were made for the execution of study IV, which should induce caution when making inferences based on the study results. The studies used in the meta-regression were heterogeneous and the reporting applied in the studies was, likewise, dissimilar. Nonetheless, data from the studies were combined to enable analysis of the reported mean value. As such, ecological bias may exist.[4] For future research, the correlation between TTR and complications might be performed on the individual level with adjustment for a more complete set of relevant variables[67]. This might validate the rates predicted from study IV, which is currently lacking.

The hypothetical QIs investigated in studies II and III were binary indicators reflecting 'resumption of therapy following ICH, yes/no' and 'treated according to guidelines, yes/no,' resulting in quality indication by the 'proportion of patients resuming

therapy' and 'proportion of patients treated according to guidelines,' respectively. In contrast, the QI, TTR, represents the level of quality in continuous therapy and is used to evaluate the quality of ongoing therapy. The meta-regressions indicating the rates of SSE and MB as functions of TTR constitute two correlation parameters, $\varphi_2$, which may be applied simultaneously in the same decision-analytic model to more sufficiently include the impact on health of increasing mean TTR. As the data in study IV were aggregated as the mean values of cohorts and the covariates were included as proportions in the meta-regressions, the results of study IV are likewise only applicable on a cohort level, i.e. reflecting the correlation between mean TTR and outcomes on a population level and not for the individual.

In study IV, the double arcsine transformation was applied to the rates of MB and SSE, producing a slightly nonlinear relationship between the mean TTR and the outcomes in the original scale (Figure 5-4). This estimation of the correlation to some degree facilitates the real-life observations that TTR is not correlated with complication rates at high values[135,136]. Consequently, in the present model, the clinical benefit of increasing TTR at high values would not be as high as increasing it at low values. It could be hypothesized that the costs of increasing mean TTR when quality of therapy is already high are similar to, or higher than, the costs related to quality improvements at lower levels of quality. This would be in agreement with the expectation presented by Donabedian that the marginal benefit (here, decreasing rates of MB and SSE) of improving quality may potentially not counteract the costs related to attainment of the quality(Figure 3-5)[24,62]. Evidence that the rate of complications does not correlate with the level of TTR above 70 percent[135] suggests that the optimally effective quality in warfarin therapy may exist at a mean TTR of 70 percent. Above this level of quality, the marginal health gains achieved by quality improvement may be bought too dearly if the cost of quality improvement increases at higher levels of quality. Thus, it is possible that above a mean TTR of 70 percent further investment in quality improvement in warfarin therapy may not be cost-effective. This remains to be elucidated.

The correlation between mean TTR and patient-relevant outcomes found in study IV might be used to model the expected cost-effectiveness of warfarin therapy at different levels of quality, i.e. different mean TTRs, through the use of the framework presented in study I. Increasing quality in warfarin therapy may be achieved through different initiatives, including the use of self-monitoring in eligible patient groups[100,106].

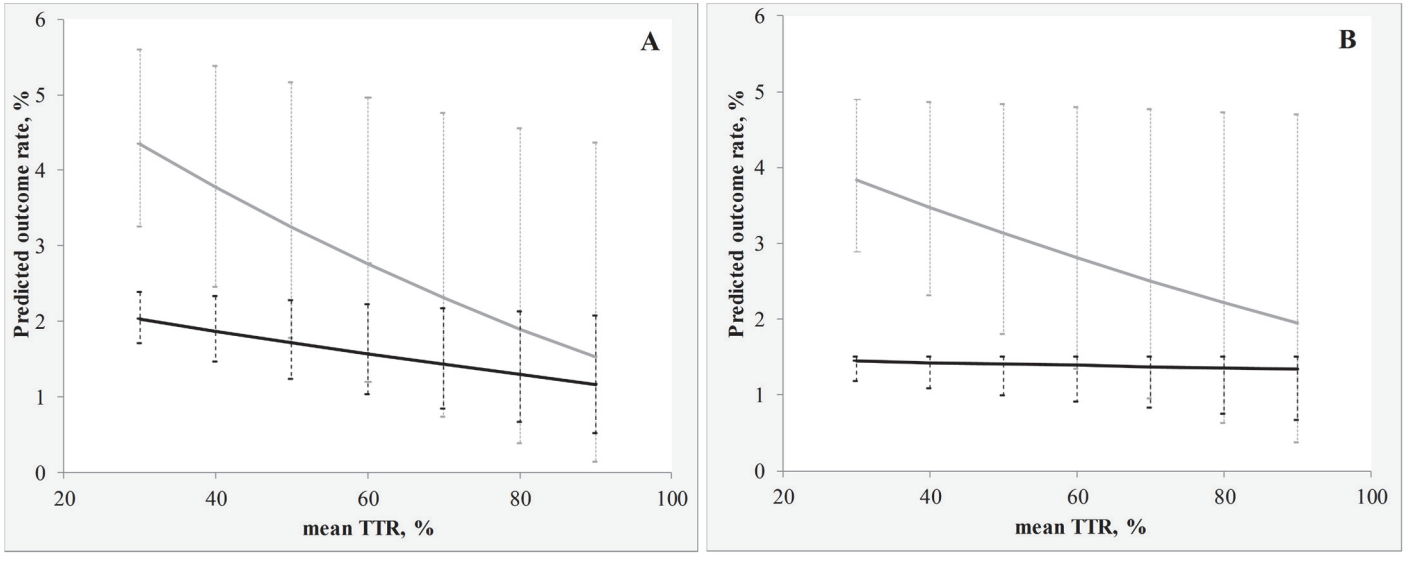



*Figure 5-4. Outcome rates, reported in original scale, as a function of mean TTR of cohorts. Gray lines indicate the outcome major bleeding. Black lines indicate stroke/systemic embolism. Figure A: predicted rates from univariable meta-regression analyses. Figure B: predicted rates from multivariable meta-regression analyses. For the multivariable meta-regression analyses, unweighted mean values for age (74 years), proportion with prior stroke/transient ischemic attack (21 percent), and proportion of females (41 percent) reported in study IV* [4] *were used. Vertical lines indicate the estimated 95% confidence intervals. TTR: Time in therapeutic range, percent. Adapted from Vestergaard et al.*[4].

Currently the cost-effectiveness of self-monitored warfarin therapy in AF is obscure, as no studies provide sufficient evidence on the occurrence of patient-relevant outcomes in self-managed therapy for AF patients alone. Evidence on TTR does exist, however, which may be used to model the expected NB, if the correlation presented in study IV is accepted. The correlation $\varphi_2$ may be used in DAM to compare, for instance, self-managed warfarin therapy, conventionally managed warfarin therapy, and the use of NOACs. This may indicate whether quality improvement in warfarin therapy provides a viable alternative to the use of new medication for eligible patients[112,135]. An outline of a Markov model-based decision analysis with the utilization of the correlation found in study IV is exemplified in Figure 5-5. In the model, the impact of differing quality of therapy as evaluated by the mean TTR would be reflected in the transition probabilities for SSE and MB in self-managed and conventional warfarin therapy. Modeling of the expected cost-effectiveness of warfarin therapy at different levels of quality has been done before[128,137]. However, previously only univariable, linear regression analyses have been utilized to estimate the impact of increasing TTR on complication rates, thereby disregarding the potential impact of other factors. Furthermore, the potential cost of uncertainty related to the use of the expected correlation has not been investigated sufficiently, which would be highly relevant for investigating the validity of the primary results and impact on the decision uncertainty.

Study IV illustrates how the mathematical correlation between an outcome-validated QI for the level of quality and patient-relevant outcomes may be estimated with adjustment for relevant antecedents. The study also highlight some of the obstacles to doing so. Due to the rarity of events, a common issue in research in quality improvement[27], multiple studies may need to be aggregated to enable estimation of the correlation $\varphi_2$. Preferably, however, data should be retrieved on the individual level allowing for more accurate data and informative analyses. Study IV illustrates the importance of adjusting for factors that affect the occurrence of the patient-relevant outcomes; not adjusting for relevant variables may bias the association and the correlation may be misspecified[67,105].

Utilization of misspecified correlations could lead to erroneous conclusions on the cost-effectiveness of interventions that increase the quality of therapy. If, for instance, a linear relationship between mean TTR and complication rates were applied or if appropriate adjustments were not performed, the benefits of increasing mean TTR might be erroneously inflated, potentially leading to unreasonably optimistic results on the impact of improving quality in warfarin therapy.

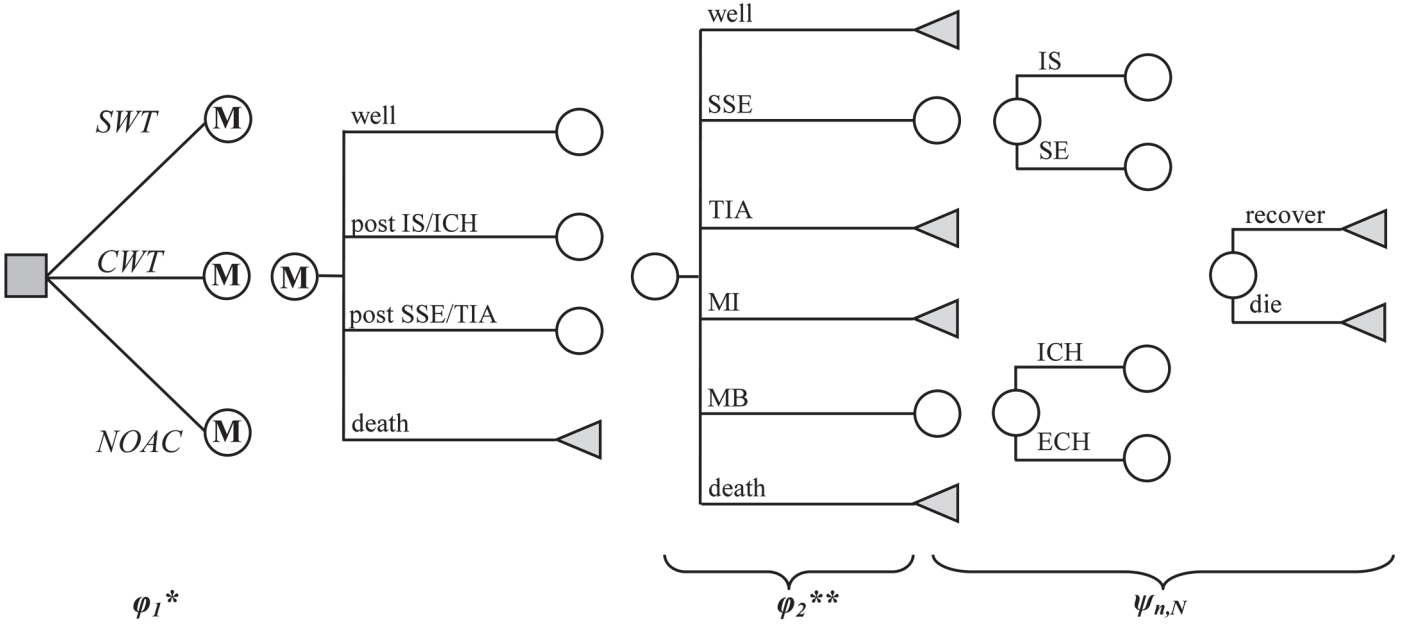



*Figure 5-5. Outline of a hypothetical decision-analytic model with incorporation of the correlation found in study IV to enable evaluation of the impact of quality of therapy on the health economic impact of quality improvement. φ1 may be included in the expected level of quality within the different warfarin therapies. No estimation of correlation for TIA, and MI was estimated and a baseline probability of these events on warfarin therapy would be applied. Transitions not included. CWT: conventionally managed warfarin therapy, ECH: extracranial hemorrhage, ICH: intracranial hemorrhage, IS: ischemic stroke, MB: major bleeding, MI: myocardial infarction, non-vitamin K oral anticoagulant, SE: systemic embolism, SSE: stroke/systemic embolism, SWT: self-managed warfarin therapy, TIA: transient ischemic attack. *uncertainty pertaining to the quality of therapy under the intervention may be included as a distribution of the TTR. **only applicable for warfarin therapy.*

# CHAPTER 6.
# DISCUSSION AND CONCLUDING REMARKS

## 6.1. IMPROVING ORAL ANTICOAGULANT THERAPY IN ATRIAL FIBRILLATION

Evidence on the comparative cost-effectiveness of OAC agents used in AF is vast[36–38]. For this reason, rather than providing confirmations of the results of previous economic evaluations of the different medication types, the clinical aim of the present dissertation was to illustrate the health economic potentials of other approaches to improving OAC in the AF population *as a whole*. Consequently, decision-making on therapy of the AF patient population has been approached from a general, policy decision-oriented level. The approach to decision-making presented here may be used as an aid in organizational considerations and in policy decision processes e.g. on which initiatives to reimburse on a patient population level, based on their mean, expected cost-effectiveness. Economic evaluation of alternative quality improvement may help identify focus areas that could be prioritized in future policy decisions. It is not intended to replace clinical judgement and decision-making in the treatment of individual patients.

As exemplified in study II and III[2,3], more focus areas for quality improvement in OAC exist, where improvement of the treatment of the AF patient population may be cost-effective. There appear to be a potential in ensuring that more patients are treated in accordance with clinical practice guidelines on stroke prophylaxis in AF [3,28,29]. Greater guideline adherence in the treatment of the patient population as a whole might increase equality in care provision with less unsubstantiated nontreatment, while representing a cost-effective treatment strategy[3]. Furthermore, discontinuation of OAC following an ICH is more common than resumption of therapy, although, on average, resumption of therapy appear to improve morbidity, mortality, and entail lower hospitalization costs[2,130,138]. Raising awareness amongst clinicians and patients alike of the net health benefit associated with OAC in AF might increase the proportion of patients resuming OAC following ICH and potentially represent a cost-effective quality improvement[2]. Lastly, therapy may also be improved by increasing the level of quality in the therapy delivered[4]. Although NOACs are increasingly used for OAC in AF, a large proportion of patients

are still in warfarin therapy[30,31]. Different approaches may be employed to improve the quality in warfarin therapy of patient cohorts, e.g. by use of specialized OAC clinics or self-managed warfarin therapy[100]. The cost-effectiveness of using specialized OAC clinics and self-managed warfarin therapy compared to conventional warfarin therapy managed in general practice is currently unknown. The cost-effectiveness is contingent on the effectiveness and safety observed under the different levels of quality of therapy delivered in the different settings. The quality may be evaluated by their mean TTR. However, the occurrence of patient-relevant outcomes is not contingent on the quality of therapy of patient cohorts alone, but is also affected by other factors[4,81,118]. Therefore, relevant characteristics of the patient populations for whom the quality of therapy might be improved should be taken into account when evaluating the cost-effectiveness of improving quality of therapy. The same may apply to the characteristics of the settings in which the therapy is delivered. If this is neglected, the health economic impact of improving quality of therapy may be misestimated[1,4].

## 6.2. EVALUATING THE VALUE FOR MONEY OF QUALITY IMPROVEMENTS – FUNDAMENTAL OBSTACLES

The social objective of economic evaluation is in general agreement with the main purpose of quality improvement in healthcare, i.e. to improve health [13,23]. This substantiates the potentials of applying health economic evaluation to establish the value for money of quality improvements. There are, however, also fundamental differences in the scope of the two scientific disciplines, in what composes viable effect measures, the understanding of evidence, and the execution of research, which should be acknowledged.

As economic evaluation is founded in normative health economics, its methodology is suffused by the core principles of welfare economics. Thus, the extra-welfarist economic evaluation employs the focus of identifying the efficient allocation of resources. In addition, it is distinctly result-oriented, in general accepting only outcomes for analysis that reflect impact on health.[13,45,47] In contrast, quality improvement is often process-oriented, accepting intermediate goals for outcomes, for instance, reflecting changes in processes that might not affect health, per se, and from which ultimate impact on health may occur at an immeasurably slow rate[26,27]. Furthermore, quality in healthcare is composed of multiple constituent elements and consequently quality improvement in healthcare may target multiple foci[62,64].

Thus, quality improvement may target other aspects of quality in healthcare than the efficient allocation of resources, e.g. equal access to healthcare or greater incorporation of patient preferences. Impact on these elements of quality in healthcare might not result in measurable impact on health, but are undoubtedly still of value.[21,43,62] Thus, there may be an inconsistency between the 'value' that the improvement brings and the understanding of 'value' employed in conventional economic evaluation, restricted to include impact on health. Arguably, the application of economic evaluation might then be considered moot when the aim of the intervention is not to improve health per se, as the measure of effect in the evaluation would not reflect the purpose of the intervention. This discrepancy should be acknowledged if economic evaluation is employed to identify the value for money of quality improvements targeting elements of quality in healthcare that do not change either resource consumption or health accumulation. As conventional economic evaluation is only intended to embrace a part of the potential impact that quality improvement may entail, it could therefore be discussed whether economic evaluation, by its current methods, provides appropriate evaluation for such interventions.

The main objective of the healthcare system is to improve the health of the population[43,47]. Efficient use of the resources in the healthcare system supports this objective[13]. If the cost-effectiveness of quality improvements cannot be confirmed, their introduction in the healthcare system should therefore be thought critically through. It might be that the reason why cost-effectiveness cannot be established is that the 'value' that the quality improvements bring is difficult to quantify and include in conventional economic evaluation. If that is the case, it should be considered whether the level of that 'value' is enough to justify the introduction of the quality improvements, although the immediate value, restricted to the extra-welfarist understanding of the term, for money they bring would not be sufficient to support their utilization under an extra-welfaristic approach to decision-making. Satisfaction with care, equity concerns, and other constituents of quality in healthcare are evidently of increasing interest and therefore of value, but these dimensions of healthcare delivery are currently not explicitly included in economic evaluation[20–22].

It could be argued that some elements of quality in healthcare, which are not explicitly included in conventional economic evaluation, such as patient preferences towards care, may be included implicitly through impact on, for instance, patients' compliance and adherence to therapy. This might derive impact on the occurrence of patient-relevant outcomes that consequently would be includible in economic evaluation. It follows, that in some instances it might still be possible and reasonable to perform economic evaluation of interventions targeting these elements of quality in healthcare, as impact on health may eventually be observed. Explication of the purpose of

individual quality improvements may substantiate their importance and defer subjugations of evaluation of their health economic potential by conventional methods, if the interventions are not intended to improve elements influencing efficiency in resource allocation. Likewise, clarifications of the ultimate purpose of quality improvements may aid the identification of interventions for which economic evaluation may be appropriate.

The methodological aim of the present dissertation concerns the application of economic evaluation only of quality improvements that are expected to eventually affect health, but for which evidence on the occurrence of patient-relevant outcomes may not be achievable. The expectation that they eventually influence health would justify the application of economic evaluation to establish their value for money.

## 6.3. PROCURING EVIDENCE ON THE IMPACT OF QUALITY IMPROVEMENTS

More approaches have been proposed for the evaluation of quality improvements, but the proper study design remains yet unresolved. Quality improvement is often obtained through multifaceted, complex interventions, in which more initiatives are launched concomitantly. This could include the introduction of new technology, education of healthcare personnel and patients, and organizational changes[26,139]. When more elements are put into practice simultaneously, it becomes increasingly difficult to command all mechanisms and identify those triggering the potential effect of the intervention[26]. Establishing the effect of the individual components of a complement intervention requires substantial knowledge of potential relationships between the different elements and a considerable amount of data to support analyses. Likewise, with greater ‘distance’ between the intervention and impact on patient-relevant outcomes, it becomes more difficult to establish a clinically and statistically significant relationship between the intervention and final impact on health[23]. Due to the often-high complexity of interactions between the different elements of quality improvements and influencing factors, it has been argued that it is next to impossible to evaluate the impact of quality improvements satisfactorily through the use of the conventional methods of EBM[26,69,139]. It might furthermore be argued that evaluation of quality improvements by currently available methods may not sufficiently capture all impact of interventions on patient-relevant outcomes; especially if impact is achieved gradually over a long time frame or through continuous processes such as dissemination of knowledge or organizational

changes[69,139]. To overcome some of the difficulties in procuring evidence on the effectiveness of quality improvements, QIs may be employed as surrogate for patient-relevant outcomes, as these may be more sensitive to changes caused by interventions. QIs may, however, not reflect impact on health, per se, as requested in economic evaluation.[23,27]

Due to the challenges in applying conventional evaluation for quality improvements, the application of conventional EBM evaluation methods for establishing the correlation between interventions and outcomes is highly disputed. It has even been argued that it should be completely refrained from.[26,69,139] In counter-argument against this stance, it might be reasoned that not establishing evidence for the effectiveness and value for money of interventions entails the risk of shelving the great opportunities that quality improvement holds. This might occur in consequence of implementation of ineffective quality 'improvements' or rejection of quality improvements in favor of other interventions for which the value for money is known. Failing to provide some kind of evidence on the expected impact of quality improvements may thus potentially cause opportunity costs and be regarded equally ill-considered[18].

The fundamental challenge of how to procure evidence on the impact of quality improvements was beyond the scope of the present dissertation. Instead, this dissertation revolves around the potentials of undertaking economic evaluation of quality improvements when evidence on the expected correlation between interventions and QIs ($\varphi_I$) has been acquired[1]. The focus is on the requirements for QIs to be valuable in the context of economic evaluation and on how economic evaluation may be designed to incorporate QIs when these have been used for outcome measure in quality improvements. Prior knowledge on the requirements to a QI is valuable in the design phase for studies intended to establish the impact of quality improvements as it enables the incorporation of these contemplations when the correlation $\varphi_I$ needs to be established.

## 6.4. USING QUALITY INDICATORS IN ECONOMIC EVALUATION

In the present dissertation, it is proposed that when QI are used for effect measure for quality improvements, it might still be possible to generate estimates of expected NB of the interventions to inform decision-making. The presented framework and recommendations for estimation of NB of interventions based on QIs are reasonably consistent with the recommendations on use of clinical, intermediate endpoints to

estimate the NB of interventions[1,53] Modeling expected NB based on surrogate, physiological outcomes remains disputed, but is more or less accepted when appropriate caution is taken[13,53]. The experience with and recommendations on modeling based on clinical, intermediate endpoints may inform future work on the application of QIs to model expected NB of quality improvements. Over time the same acceptance might be achieved for modeling of NB-based QIs.

In 1966 when quality evaluation and improvement was first couched as a scientific discipline, Donabedian queried "*...but how precise do estimates of quality have to be?*"[23]. This question poses the focal point of the present dissertation. The answer may not be straightforward and differs depending on what the intended use of the estimates is. If QIs are used to make inferences about the impact on health and consequently expected cost-effectiveness that interventions supply, a number of requirements may exist for them to be applicable. The correlation between QIs and final impact on health is in all likelihood affected by uncertainty, which may subsequently affect the decision uncertainty and increase the risk of making suboptimal decisions. Estimating NB based on QIs remains the inferior solution, due to the potential risk of misspecifications of the correlation φ2, the additional uncertainty of $\varphi_2$, and consequent erroneous NB estimations. If it is possible, within reason, to acquire evidence on the correlation between interventions and the occurrence of patient-relevant outcomes this would be the recommended approach[1,53]. However, the lack of evidence on impact on patient-relevant outcomes remains a core issue in work with quality improvement. The issue will expectedly only persist, necessitating deliberations of the appropriate handling of it in a health economic context. Use of DAM and the handling of QIs as an extra intermediate link may be a way forward under the expectation that uncertainty and the potential cost of uncertainty are sufficiently highlighted. Analyses may yet be performed and may provide useful information, despite the uncertainty pertaining to the available evidence. The accuracy of economic evaluations is not as important as how the information they convey affects the decision, which it is intended to aid[13].

Through the execution of the studies I-IV[1–4], four steps have been identified that may guide the execution of economic evaluation of quality improvements, when QIs have been used for outcome (Box 5). To enable later economic evaluation of quality improvements, for which QIs are applied for outcome, these steps should be considered even before initiating gathering of evidence.

For quality improvements that are expected to ultimately affect health, only outcome-validated QIs should be used to enable subsequent economic evaluation of the intervention[23,53,67]. Furthermore, it should be considered whether the applied QI embraces all relevant impact on health. Otherwise it might be necessary to employ

more, mutually exclusive QIs to reflect the entirety of the effect of interventions[1]. If non-validated QIs are employed, analyses potentially become uninformative or, worse, misleading. This may eventually result in benefits foregone if either cost-ineffective interventions are used or cost-effective interventions are rejected.

| | |
|---|---|
| I. | Use only outcome-validated quality indicator(s) for outcome measure. The applied quality indicator(s) should be strongly related to patient-relevant outcome(s). |
| II. | Estimate the expected correlation, $\varphi_2$, between the applied quality indicator(s) and patient-relevant outcome(s), if it is not already established. |
| III. | Estimate the expected net benefit of interventions based on the quality indicator(s) by use of decision-analytic modeling. |
| IV. | Estimate and evaluate the expected cost of uncertainty related to the results, particularly the use of the quality indicator(s); the uncertainty parameter, $\varphi_2$. |

*Box 5. Listing of four steps to enable and execute economic evaluation of quality improvements when quality indicators are used for outcome measure.*

If the relationship between QI(s) and patient-relevant outcome(s) is generally accepted, but not yet quantified, this should be done to enable further use of the correlation in DAM. It is critical that the correlation $\varphi_2$ is correctly specified to avoid erroneous NB estimations; especially to avoid overestimation of the benefits involved with the quality improvement as measured via the QI[53]. This may include both the application of an appropriate model for the correlation and appropriate adjustment for factors known to influence the occurrence of the patient-relevant outcomes. As illustrated in study IV[4], adjustment for all relevant factors may prove itself difficult, for instance due to missing information on parameters of interest. The risk and potential consequences of unobserved confounding and bias should be deliberated when the correlation is established as the impact of these would not be reflected in VOI analyses. The more information that is available, and on less aggregate level, the more informative analyses may become and the smaller the risk of establishing confounded correlations. Adjustment for the characteristics of the setting and patient population in which the QI is used may furthermore ensure a better reflection of the expected impact of quality improvement for that specific context and consequently the soundness of analyses[105].

The correlation parameter $\varphi_2$ would be valid for the patient group, for which it is estimated. Ideally, it should be estimated for a reasonable homogeneous patient group to diminish the level of uncertainty pertaining to the parameter. Correlations estimated for highly specific groups would likely entail less uncertainty, leading to less decision

uncertainty. They would, however, potentially lack the transferability to analyses performed for other, dissimilar groups. As a result, the 'global' use of it would be hampered. The balance between specificity and 'globality' of a QI might be struck pragmatically, based on the grouping level relevant for decision-making. That is, struck, based on whether interventions expectedly will be offered to the patient group under investigation or not. Information on subgroups within that group is of less interest, as the decision whether to offer the intervention to the entire group or not would not change. This could be die to a desire of equal access to care within the group. For instance, it appears unlikely that all male AF patients would be offered a certain intervention and female AF patients not, even if the intervention appeared cost-ineffective in the latter group. Estimation of the correlation parameter $\varphi_2$ for the entire AF patient population would thus be more relevant for further use.

When the correlation between QI(s) and patient-relevant outcome(s) has been established, the expected NB of interventions may be estimated by use of DAM[1,13,53]. The decision uncertainty and cost of uncertainty pertaining to the results should be evaluated and it should be considered whether resolving the uncertainty would be relevant to diminish the risk of making the wrong decision, based on the current level of evidence[1,49,61]. Use of the framework presented in this dissertation might consequently ease estimation of the cost-effectiveness of quality improvements where non-health-related endpoints have otherwise been applied, even when incremental costs are positive.

## 6.5. EVALUATING THE VALUE FOR MONEY OF QUALITY IMPROVEMENTS – FUTURE OPPORTUNITIES

In the present dissertation, it is argued that, if done sensibly and cautiously, using QIs in DAM may help inform policymakers on the potential value for money of quality improvements. This may result in more informed and consistent decision-making, as the impact of quality improvements may be evaluated on equal terms with other interventions in the healthcare system.

The present framework may potentially be used for pilot economic evaluation of quality improvements, if evidence on impact on QIs of interventions is made available early during the execution of studies. This may be relevant even when patient-relevant outcomes are expected to eventually occur. Preliminary estimations of the potential health economic impact of interventions may provide early information on their

expected NB. This may enable early termination of studies and discarding of interventions that appear unlikely to provide a positive NB, although evidence on direct impact on health remains unexplored. This would be cost-saving compared to the completion of studies to establish evidence on the (lack of) impact of interventions on patient-relevant outcomes. Nonetheless, as emphasized, estimates of NB estimated through the use of the present framework would be subject to increased uncertainty due to the introduction of the correlation parameter $\varphi_2$, which should be appraised critically when evaluating the validity of results.

The use of the framework is in the present dissertation exemplified by the use of QIs in stroke prophylaxis in AF. The framework should however be transferable to other disease and treatment areas, by incorporation of other outcome-validated QIs relevant for the investigated decision problem. For instance, diabetes patients suffer an increased risk of various late complications, including foot ulcers that further increase the risk of leg amputations. These complications inevitably entail both a high personal as well as economic burden. Increased attention to early symptoms paid by patients and healthcare personnel alike may in part prevent these late complications of diabetes. Diabetes patients are consequently recommended regular foot care and examinations to diminish the risk of developing complication and for early detection of them in order to retain further worsening.[140,141] Accordingly, the proportion of diabetes patients receiving regular foot care is accepted as a process QI[20]. By different initiatives, it might be possible to increase the proportion of patients receiving regular foot care, which would then expectedly decrease the occurrence of diabetic foot ulcers and other complications. If the reduction in risk of late complications achieved by regular foot care were quantified, i.e. $\varphi_2$ were established, it might be possible to estimate the NB of increasing the proportion of diabetes patient receiving regular foot care by use of the present framework. Other quality improvements may be evaluated in the same manner, as long as the necessary correlations can be established and the caveats presented in this dissertation are taken into account.

As patient-centeredness is of increasing importance in healthcare delivery[22,63], the ability to valuate and incorporate e.g. patient preferences, satisfaction, and convenience of care, in evaluations is of increasing interest. The desire to incorporate more elements than cost-effectiveness in evaluation of interventions is not singular for quality improvements. In some therapy fields, the boundary for achievable efficacy and safety is almost reached. Consequently, further benefits of (new) therapies can only be achieved through improvements in other characteristics of therapy, e.g. in the convenience of care. For instance, it could be argued that the major benefit of NOACs compared to warfarin therapy is not an improved safety and effectiveness profile[30,31], but rather that the medication is more convenient for

patients and healthcare personnel alike[113]. To explicitly include patient-relevant elements of care, i.e. preferences towards interventions due to greater convenience etc., the CBA may be a useful approach to the economic evaluation of interventions. Using the CBA would enable the inclusion of more relevant aspects of care than HRQoL, as captured by the CUA, including elements of patient-centeredness and equity, insofar these are appropriately and reliably valuated. As explicated (cf. section 3.3.1), there are however appreciable challenges to the application of CBA within healthcare.

In addition, multi-criteria decision analysis might provide a potential decision support framework that enables the inclusion of more criteria in the decision-making process than is currently enabled in conventional economic evaluation. In multi-criteria decision analysis, the cost-effectiveness of intervention, as determined by conventional methods may constitute one out of more criteria deemed relevant for a specific decision problem. Hence, different aspects of quality in healthcare may be explicitly and systematically included in the decision analysis. The different criteria are subsequently weighted according to stakeholders' preferences towards them.[142,143] Multi-criteria decision analysis thus enables explicit and transparent inclusion of other elements of importance for decision-making, which could include other attributes of quality in healthcare, such as equity concerns.

The present dissertation has demonstrated some of the obstacles for application of health economic evaluation to the area of quality improvements. It also provides a contribution as to the opportunities for using QI as intermediate link in DAM to establish the value for money of quality improvements, when patient-relevant outcomes are not available. The application of economic evaluation to quality improvement requires great interdisciplinary cooperation with participation from diverse evaluation traditions and therefore also the acceptance and inclusion of different epistemologies. Compromises must be made and researchers in both quality improvement and economic evaluation must extend their hands to hopefully improve research. Researchers in the field of quality improvement should consider the purpose and use of QIs early in the design of studies if economic evaluation is under consideration. Researcher in economic evaluation should continue striving to develop the methods for economic evaluation for inclusion of evidence derived from other sources than EBM. The inclusion and merging of more scientific methodologies may improve our understanding of the mechanisms leading to quality improvement in healthcare, benefitting patients and the society as a whole.[26,69,139]

# LITERATURE LIST


[1] Vestergaard AS, Ehlers LH. What is the value of quality indicators for informed decision-making in healthcare? A Bayesian decision theoretical and value of information analysis perspective. 2017.

[2] Vestergaard AS, Skjøth F, Lip GYH, Larsen TB. Effect of Anticoagulation on Hospitalization Costs After Intracranial Hemorrhage in Atrial Fibrillation - A Registry Study. Stroke 2016;47:979–85.

[3] Vestergaard AS, Ehlers. A Health Economic Evaluation of Stroke Prevention in Atrial Fibrillation: Guideline Adherence Versus the Observed Treatment Strategy Prior to 2012 in Denmark. Pharmacoeconomics 2015;33:967–79.

[4] Vestergaard AS, Skjøth F, Larsen TB, Ehlers LH. The importance of mean time in therapeutic range for complication rates in warfarin therapy of atrial fibrillation: a meta-regression analysis. 2017.

[5] Harris RE. Global Epidemiology of Chronic Diseases : The Epidemiologic Transition. Epidemiol. Chronic Dis., Burlington: Jones & Bartlett Publishers; 2013, p. 1–24.

[6] Hvidberg MF. A framework for identifying disease burden and estimating health-related quality of life and prevalence rates for 199 medically defined chronic conditions. 1st ed. Aalborg: Aalborg University Press; 2016.

[7] Organisation for Economic Co-operation and Development (OECD). Health Expenditure and Financing 2016. http://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT (accessed February 14, 2017).

[8] The Ministry of Health. The healthcare system in numbers and facts [Sundhedsvæsnet i tal og fakta] 2016. http://www.sum.dk/Sundhed/Sundhedsvaesnet/Sundhedsoekonomi/Sundhedsvaesenet_i_tal_og_fakta.aspx (accessed February 14, 2017).

[9] Danish Regions [Danske Regioner]. Prioritization [Prioritering] 2017. http://www.regioner.dk/sundhed/medicin/prioritering (accessed February 17, 2017).

[10] The Ministry of Health. What are the costs of the healthcare system? [Hvad koster sundhedsvæsnet?] 2016. http://www.sum.dk/Sundhed/Sundhedsvaesnet/Sundhedsoekonomi/Sundhedsvaesenet_i_tal_og_fakta/Hvad_koster_sundhedsvaesnet.aspx (accessed February 17, 2017).

[11] Pedersen KM. Sundhedsøkonomi. 1st ed. København: Munksgaard; 2013.

[12] Morris S, Devlin N, Parkin D. Economic analysis in health care. 1st ed. Chichester: John Wiley & Sons, Ltd; 2010.

[13] Drummond MF, Schulpher M, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Oxford: Oxford University Press; 2015.

[14] Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty--when is there sufficient evidence? Value Health 2005;8:433–46.

[15] Jensen LW. The growth in medication costs abates [Væksten i medicinudgifterne tager af]. Medicinal & Biotek, MedWatch 2017. http://medwatch.dk/Medicinal___Biotek/article9337013.ece (accessed February 3, 2017).

[16] Øvretveit J, Appleby J. Does improving quality of care save money? vol. 339. London: The Health Foundation; 2009.

[17] Rauh SS, Wadsworth EB, Weeks WB, Weinstein JN. The Savings Illusion — Why Clinical Quality Improvement Fails to Deliver Bottom-Line Results. N Engl J Med 2011;365:e48.

[18] Severens JL. Value for money of changing healthcare services? Economic evaluation of quality improvement. Qual Saf Health Care 2003;12:366–71.

[19] Mainz J, Bartels P, Bek T, Pedersen KM, Krøll V, Rhode P. Kvalitetsudvikling i praksis. 1st ed. København: Munksgaard; 2011.

[20] Mainz J. Defining and classifying clinical indicators for quality improvement. Int J Qual Health Care 2003;15:523–30.

[21] Berger JT. National quality program for the healthcare area 2015-2018 [Nationalt kvalitetsprogram for sundhedsområdet 2015-2018]. 2015.

[22] The Ministry of Health. Eight new national aims for the healthcare system [8 nye mål for sundhedsvæsnet] 2016. http://www.sum.dk/Temaer/Nationale-maal-for-sundhedsvaesenet.aspx (accessed February 17, 2017).

[23] Donabedian A. Evaluating the Quality of Medical Care. Milbank Mem Fund Q 1966;44:166–203.

[24] Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260:1743–8.

[25] Meltzer D. Economic analysis in patient safety : a neglected necessity. BMJ Qual Saf 2012;21:443–5.

[26] Berwick DM. The science of improvement. JAMA 2008;299:1182–4.

[27] Mant J. Process versus outcome indicators in the assessment of quality of health care. Int J Qual Heal Care 2001;13:475–80.

[28] Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 2016;37:2893–962.

[29] The Danish Counsil for the Use of Expensive Hospital Medicines [Rådet for Anvendelse af Dyr Sygehusmedicin]. Background memorandum for oral anticoagulant therapy in non-valvular atrial fibrillation [Baggrundsnotat for oral antikoagulationsbehandling ved non-valvulær atrieflimren] 2016:1–8. http://www.regioner.dk/media/3415/beh-atrieflimmer-juni-2016.pdf (accessed January 13, 2017).

[30] Larsen TB, Skjøth F, Nielsen PB, Kjældgaard JN, Lip GYH. Comparative effectiveness and

safety of non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. BMJ 2016;353:i3189.

[31] Nielsen PB, Skjøth F, Søgaard M, Kjældgaard JN, Lip GYH, Larsen TB. Effectiveness and safety of reduced dose non-vitamin K antagonist oral anticoagulants and warfarin in patients with atrial fibrillation: propensity weighted nationwide cohort study. BMJ 2017;353:j510.

[32] Martinez C, Katholing A, Wallenhorst C, Freedman SB. Therapy persistence in newly diagnosed non-valvular atrial fibrillation treated with warfarin or NOAC: A cohort study. Thromb Haemost 2016;115:31–9.

[33] Lip GYH, Laroche C, Dan G-A, Santini M, Kalarus Z, Rasmussen LH, et al. "Real-World" Antithrombotic Treatment in Atrial Fibrillation: The EORP-AF Pilot Survey. Am J Med 2014;127:519–29.

[34] Ruff CT, Giugliano RP, Braunwald E, Hoffman EB, Deenadayalu N, Ezekowitz MD, et al. Comparison of the efficacy and safety of new oral anticoagulants with warfarin in patients with atrial fibrillation: a meta-analysis of randomised trials. Lancet 2014;383:955–62.

[35] Danish Medicines Agency. medicinpriser.dk 2017. http://www.medicinpriser.dk (accessed February 4, 2017).

[36] Langkilde LK, Bergholdt Asmussen M, Overgaard M. Cost-effectiveness of dabigatran etexilate for stroke prevention in non-valvular atrial fibrillation. Applying RE-LY to clinical practice in Denmark. J Med Econ 2012;15:695–703.

[37] Pink J, Pirmohamed M, Hughes DA. Comparative Effectiveness of Dabigatran, Rivaroxaban, Apixaban, and Warfarin in the Management of Patients With Nonvalvular Atrial Fibrillation. Clin Pharmacol Ther 2013;94:269–76.

[38] Zheng Y, Sorensen S V, Gonschior A-K, Noack H, Heinrich-Nols J, Sunderland T, et al. Comparison of the Cost-effectiveness of New Oral Anticoagulants for the Prevention of Stroke and Systemic Embolism in Atrial Fibrillation in a UK Setting. Clin Ther 2014;36:2015–2028.e2.

[39] Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom JW, Oldgren J, Parekh A, et al. Dabigatran versus warfarin in patients with atrial fibrillation. N Engl J Med 2009;361:1139–51.

[40] Granger CB, Alexander JH, McMurray JJV, Lopes RD, Hylek EM, Hanna M, et al. Apixaban versus warfarin in patients with atrial fibrillation. N Engl J Med 2011;365:981–92.

[41] Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med 2011;365:883–91.

[42] Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL, et al. Edoxaban versus warfarin in patients with atrial fibrillation. N Engl J Med 2013;369:2093–104.

[43] The Ministry of Health. The healthcare system [Sundhedsvæsnet]. Heal Care Act, LBK Nr 1188 from 24/09/2016 (Bekendtgørelse Af Sundhedsloven) 2016:1–2. https://www.retsinformation.dk/Forms/r0710.aspx?id=152710. (accessed March 4, 2017).

[44] World Health Organisation. Preamble to the Constitution of WHO as adopted by the International Health Conference 1946:Official Records of WHO, no. 2, page 100.

[45] Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. Measuring and Valuing Health Benefits for Economic Evaluation. 1st ed. Oxford: Oxford University Press; 2007.

[46] Folland S, Goodman AC, Stano M. The Economics of Health and Health Care. 6th ed. New Jersey: Pearson Educational (International ed.); 2010.

[47] Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. J Health Econ 2008;27:325–38.

[48] McCabe C, Claxton K, Culyer AJ. The NICE cost effectiveness threshold: What it is and what it means. Pharmacoeconomics 2008;2:733–44.

[49] Briggs A, Claxton K, Schulpher M. Decision Modelling for Health Economic Evaluation. New York: Oxford University Press; 2011.

[50] Claxton K. The irrelevance of inference: A decision-making approach to the stochastic evaluation of health care technologies. J Health Econ 1999;18:341–64.

[51] Research Institute at the Hospital for Sick Children. Distributions of economic evalations in PEDE [Paediatric Economic Database Evaluation] by year (n=2972). Trends Econ Eval 2017. http://pede.ccb.sickkids.ca/pede/trends.jsp (accessed May 4, 2017).

[52] Claes N, Moeremans K, Frank B, Jef A, Jos V, Herman VL, et al. Estimating the cost-effectiveness of quality-improving interventions in oral anticoagulation management within general practice. Value Heal 2006;9:369–76.

[53] Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: a review of United Kingdom health technology assessment reports. Int J Technol Assess Health Care 2009;25:6–13.

[54] Karimi M, Brazier J. Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? Pharmacoeconomics 2016;34:1–5.

[55] Greenberg D, Neumann PJ. Cost-Effectiveness Analysis Expands its Reach Worldwide. Value Heal Reg Issues 2016;10:101–2.

[56] Juul S. Epidemiology and evidence [Epidemiologi og evidens]. 1st ed., Munksgaard Danmark; 2004, p. 150–81.

[57] Gallego P, Roldán V, Marin F, Gálvez J, Valdés M, Vicente V, et al. SAMe-TT2R2 score, time in therapeutic range, and outcomes in anticoagulated patients with atrial fibrillation. Am J Med 2014;127:1083–8.

[58] Kristensen FB, Sigmund H. Metodehåndbog for Medicinsk Teknologivurdering. 2nd ed. København: Sundhedsstyrelsen, Enhed for Medicinsk Teknologivurdering; 2007.

[59] Claxton K, Neumann PJ, Araki S, Weinstein MC. BAYESIAN VALUE-OF-INFORMATION ANALYSIS An Application to a Policy Model of Alzheimer's Disease. Int J Technol Assess Health Care 2001;17:38–55.

[60] Raiffa H. Decision Analysis: Introductory lecture on choices under uncertainty. MD Comput an Off J Am Med Inf Assoc 1968;10.

[61] Phillips C V. The economics of “more research is needed”. Int J Epidemiol 2001;30:771–6.

[62] Donabedian A. The seven pillars of quality. Arch Pathol Lab Med 1990;114:1115–8.

[63] Commitee on Quality Health Care in America. Crossing the Quality Chasm. Washington, D.C.: National Academies Press; 2001.

[64] Institute of Medicine. Shaping the Future, newsletter of the Institute of Medicine. Washington, D.C.: 2002.

[65] Agency for Healthcare Research and Quality - U.S Department of Health and Human Services. National Quality Measures Clearinghouse 2017. https://qualitymeasures.ahrq.gov/. (accessed April 6, 2017).

[66] Porter ME, Larsson S, Lee TH. Standardizing Patient Outcomes Measurement. N Engl J Med 2016;374:504–6.

[67] Rose AJ, Berlowitz DR, Frayne SM, Hylek EM. Measuring quality of oral anticoagulation care: extending quality measurement to a new field. Jt Comm J Qual Patient Saf 2009;35:146–55.

[68] Portela MC, Pronovost PJ, Woodcock T, Carter P, Dixon-Woods M. How to study improvement interventions: a brief overview of possible study types. BMJ Qual Saf 2015;24:325–36.

[69] Parry G, Power M. To RCT or not to RCT? The ongoing saga of randomised trials in quality improvement. BMJ Qual Saf 2016;25:221–3.

[70] Heeringa J, Van Der Kuip DAM, Hofman A, Kors JA, Van Herpen G, Stricker BHC, et al. Prevalence, incidence and lifetime risk of atrial fibrillation: The Rotterdam study. Eur Heart J 2006;27:949–53.

[71] Haim M, Hoshen M, Reges O, Rabi Y, Balicer R, Leibowitz M. Prospective National Study of the Prevalence, Incidence, Management and Outcome of a Large Contemporary Cohort of Patients With Incident Non-Valvular Atrial Fibrillation. J Am Hear Assoc 2015;4:e001486.

[72] Lloyd-Jones DM. Lifetime Risk for Development of Atrial Fibrillation: The Framingham Heart Study. Circulation 2004;110:1042–6.

[73] Lane DA, Skjøth F, Lip GYH, Larsen TB, Kotecha D. Temporal Trends in Incidence, Prevalence, and Mortality of Atrial Fibrillation in Primary Care. J Am Heart Assoc 2017;6:1–10.

[74] Frankel DS, Parker SE, Rosenfeld LE, Gorelick PB. HRS/NSA 2014 Survey of Atrial Fibrillation and Stroke: Gaps in Knowledge and Perspective, Opportunities for Improvement. J Stroke Cerebrovasc Dis 2015;24:1691–700.

[75] Stewart S, Murphy NF, Walker A, McGuire A, McMurray JJV. Cost of an emerging epidemic: an economic analysis of atrial fibrillation in the UK. Heart 2004;90:286–92.

[76] Ericson L, Bergfeldt L, Björholt I. Atrial fibrillation: the cost of illness in Sweden. Eur J Heal Econ 2011;12:479–87.

[77] Kim MH, Johnston SS, Chu B-C, Dalal MR, Schulman KL. Estimation of total incremental

health care costs in patients with atrial fibrillation in the United States. Circ Cardiovasc Qual Outcomes 2011;4:313–20.

[78] Ringborg A, Nieuwlaat R, Lindgren P, Jönsson B, Fidan D, Maggioni AP, et al. Costs of atrial fibrillation in five European countries: results from the Euro Heart Survey on atrial fibrillation. Europace 2008;10:403–11.

[79] Jakobsen M, Kolodziejczyk C, Fredslund EK, Poulsen PB, Dybro L, Johnsen SP. Societal Costs of First-Incident Ischemic Stroke in Patients with Atrial Fibrillation - A Danish Nationwide Registry Study. Value Heal 2016;19:413–8.

[80] Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991;22:983–8.

[81] Olesen JB, Lip GYH, Hansen ML, Hansen PR, Tolstrup JS, Lindhardsen J, et al. Validation of risk stratification schemes for predicting stroke nationwide cohort study. BMJ Open 2011;342:1–9.

[82] Marini C, De Santis F, Sacco S, Russo T, Olivieri L, Totaro R, et al. Contribution of Atrial Fibrillation to Incidence and Outcome of Ischemic Stroke: Results From a Population-Based Study. Stroke 2005;36:1115–9.

[83] Lamassa M, Di Carlo A, Pracucci G, Basile AM, Trefoloni G, Vanni P, et al. Characteristics, outcome, and care of stroke associated with atrial fibrillation in Europe: data from a multicenter multinational hospital-based registry (The European Community Stroke Project). Stroke 2001;32:392–8.

[84] Banerjee A, Lane DA, Torp-Pedersen C, Lip GYH. Net clinical benefit of new oral anticoagulants (dabigatran, rivaroxaban, apixaban) versus no treatment in a “real world” atrial fibrillation population: A modelling analysis based on a nationwide cohort study. Thromb Haemost 2011;107:584–9.

[85] Hart RG, Diener H-C, Yang S, Connolly SJ, Wallentin L, Reilly PA, et al. Intracranial Hemorrhage in Atrial Fibrillation Patients During Anticoagulation With Warfarin or Dabigatran: The RE-LY Trial. Stroke 2012;43:1511–7.

[86] Sullivan PW, Arant TW, Ellis SL, Ulrich H. The cost effectiveness of anticoagulation management services for patients with atrial fibrillation and at high risk of stroke in the US. Pharmacoeconomics 2006;24:1021–33.

[87] Fang MC, Go AS, Chang Y, Hylek EM, Henault LE, Jensvold NG, et al. Death and Disability from Warfarin-Associated Intracranial and Extracranial Hemorrhages. Am J Med 2007;120:700–5.

[88] Singer DE, Chang Y, Fang MC, Borowsky LH, Pomernacki NK, Udaltsova N, et al. The net clinical benefit of warfarin anticoagulation in atrial fibrillation. Ann Intern Med 2009;151:297–305.

[89] Friberg L, Rosenqvist M, Lip GYH, Friberg L, Rosenqvist M, Lip GYH. Net Clinical Benefit of Warfarin in Patients With A Report From the Swedish Atrial Fibrillation Cohort Study. Circulation 2012;125:2298–307.

[90] Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med 2007;146:857–67.

[91] Lip GYH, Edwards SJ. Stroke prevention with aspirin, warfarin and ximelagatran in patients with non-valvular atrial fibrillation: A systematic review and meta-analysis. Thromb Res 2006;118:321–33.

[92] Amouyel P, Mismetti P, Langkilde LK, Jasso-Mosqueda G, Nelander K, Lamarque H. INR variability in atrial fibrillation: A risk model for cerebrovascular events. Eur J Intern Med 2009;20:63–9.

[93] Rose A, Berlowitz D, Reisman J, Ash A, Ozonoff A, Hylek E. Percent time in therapeutic INR range (TTR): mean TTR achieved among patients who received prescriptions for warfarin and had sufficient INR values to calculate TTR. Natl Qual Meas Clear 2011. https://qualitymeasures.ahrq.gov/summaries/summary/32739/percent-time-in-therapeutic-inr-range-ttr-mean-ttr-achieved-among-patients-who-received-prescriptions-for-warfarin-and-had-sufficient-inr-values-to-calculate-ttr. (accessed April 6, 2017).

[94] Shalev V, Rogowski O, Shimron O, Sheinberg B, Shapira I, Seligsohn U, et al. The interval between prothrombin time tests and the quality of oral anticoagulants treatment in patients with chronic atrial fibrillation. Thromb Res 2007;120:201–6.

[95] Rosendaal FR, Cannegieter SC, van der Meer FJ, Briët E. A method to determine the optimal intensity of oral anticoagulant therapy. Thromb Haemost 1993;69:236–9.

[96] Connolly SJ, Pogue J, Eikelboom J, Flaker G, Commerford P, Franzosi MG, et al. Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. Circulation 2008;118:2029–37.

[97] Wan Y, Heneghan C, Perera R, Roberts N, Hollowell J, Glasziou P, et al. Anticoagulation control and prediction of adverse events in patients with atrial fibrillation: a systematic review. Circ Cardiovasc Qual Outcomes 2008;1:84–91.

[98] Gallagher AM, Setakis E, Plumb JM, Clemens A, van Staa T-P. Risks of stroke and mortality associated with suboptimal anticoagulation in atrial fibrillation patients. Thromb Haemost 2011;106:968–77.

[99] Senoo K, Lip GYH. Relationship of Age with Stroke and Death in Anticoagulated Patients with Nonvalvular Atrial Fibrillation: AMADEUS Trial. Stroke 2015;46:3202–7.

[100] Mearns ES, White CM, Kohn CG, Hawthorne J, Song J-S, Meng J, et al. Quality of vitamin K antagonist control and outcomes in atrial fibrillation patients: A meta-analysis and meta-regression. Thromb J 2014;12:14.

[101] White CM. INR Stability, Clinical Importance, and Predictors in Patients With Atrial Fibrillation and Venous Thromboembolism Receiving Vitamin K Antagonists. J Pharm Technol 2016;32:253–9.

[102] Holm T, Deutch S, Lassen JF, Jastrup B, Husted SE, Heickendorff L. Prospective evaluation of the quality of oral anticoagulation management in an outpatient clinic and in general practices. Thromb Res 2002;105:103–8.

[103] Wallentin L, Lopes RD, Hanna M, Thomas L, Hellkamp A, Nepal S, et al. Efficacy and safety of apixaban compared with warfarin at different levels of predicted international normalized ratio control for stroke prevention in atrial fibrillation. Circulation 2013;127:2166–76.

[104] Nissen Bonde A, Lip GY, Kamper AL, Staerk L, Gislason G, Torp-Pedersen C, et al. Factors associated with quality of anticoagulation control in patients with atrial fibrillation and risk of stroke and bleeding. Eur Heart J 2016;37:402.

[105] Rose AJ, Hylek EM, Ozonoff A, Ash AS, Reisman JI, Berlowitz DR. Risk-adjusted percent time in therapeutic range as a quality indicator for outpatient oral anticoagulation: results of the Veterans Affairs Study to Improve Anticoagulation (VARIA). Circ Cardiovasc Qual Outcomes 2011;4:22–9.

[106] Heneghan C, Ward A, Perera R, Bankhead C, Fuller A, Stevens R, et al. Self-monitoring of oral anticoagulation: systematic review and meta-analysis of individual patient data. Lancet 2012;379:322–34.

[107] Chan PH, Li WH, Hai JJ, Chan EW, Wong ICK, Tse HF, et al. Time in Therapeutic Range and Percentage of International Normalized Ratio in the Therapeutic Range as a Measure of Quality of Anticoagulation Control in Patients With Atrial Fibrillation. Can J Cardiol 2016;32:1247.e23-1247.e28.

[108] Lind M, Fahlén M, Kosiborod M, Eliasson B, Odén A. Variability of INR and its relationship with mortality, stroke, bleeding and hospitalisations in patients with atrial fibrillation. Thromb Res 2012;129:32–5.

[109] Van Den Ham HA, Klungel OH, Leufkens HGM, Van Staa TP. The patterns of anticoagulation control and the risk of stroke, bleeding and mortality in patients with non-valvular atrial fibrillation. J Thromb Haemost 2013;11:107–15.

[110] Rasmussen LH, Larsen TB, Graungaard T, Skjoth F, Lip GYH. Primary and secondary prevention with new oral anticoagulant drugs for stroke prevention in atrial fibrillation: indirect comparison analysis. BMJ 2012;345:e7097–e7097.

[111] Lip GYH, Larsen TB, Skjøth F, Rasmussen LH. Indirect Comparisons of New Oral Anticoagulant Drugs for Efficacy and Safety When Used for Stroke Prevention in Atrial Fibrillation. J Am Coll Cardiol 2012;60:738–46.

[112] Sjögren V, Grzymala-Lubanski B, Renlund H, Friberg L, Lip GYH, Svensson PJ, et al. Safety and efficacy of well managed warfarin. Thromb Haemost 2015;113:1370–7.

[113] Wilke T, Bauer S, Mueller S, Kohlmann T, Bauersachs R. Patient Preferences for Oral Anticoagulation Therapy in Atrial Fibrillation: A Systematic Literature Review. Patient 2017;10:17–37.

[114] Camm J, Lip GYH, De Caterina R, Savelieva I, Atar D, Hohnloser SH, et al. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. Developed with the special contribution of the European Heart Rhythm Association. Eur Heart J 2012;33:2719–47.

[115] The Danish Counsil for the Use of Expensive Hospital Medicines [Rådet for Anvendelse af Dyr Sygehusmedicin]. Treatment instructions for oral anticoagulant therapy in non-valvular atrial fibrillation [Behandlingsvejledning for oral antikoagulationsbehandling ved non-valvulær atrieflimren] 2013. http://www.regioner.dk/media/3414/arkiv-beh-oral-antikoagulationsbehandling-ved-non-valvulaer-atrieflimren.pdf (accessed June 23, 2016).

[116] Brandes A, Overgaard M, Plauborg L, Dehlendorff C, Lyck F, Peulicke J, et al. Guideline Adherence of Antithrombotic Treatment Initiated by General Practitioners in Patients With

Nonvalvular Atrial Fibrillation: A Danish Survey. Clin Cardiol 2013;36:427–32.

[117] Mikkelsen AP, Hansen ML, Olesen JB, Hvidtfeldt MW, Karasoy D, Husted S, et al. Substantial differences in initiation of oral anticoagulant therapy and clinical outcome among non-valvular atrial fibrillation patients treated in inpatient and outpatient settings. Europace 2016;18:492–500.

[118] Lip GYH, Frison L, Halperin JL, Lane D a. Comparative validation of a novel risk score for predicting bleeding risk in anticoagulated patients with atrial fibrillation: The HAS-BLED (hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile INR, elderly, drug. J Am Coll Cardiol 2011;57:173–80.

[119] Brandes A, Jensen HK, Husted SE. Atrial fibrillation and flutter [Atrieflimren og atrieflagren] 2016. http://nbv.cardio.dk/af#afs15_5_1 (accessed February 28, 2017).

[120] Proietti M, Nobili A, Raparelli V, Napoleone L, Mannucci PM, Lip GYH. Adherence to antithrombotic therapy guidelines improves mortality among elderly patients with atrial fibrillation: insights from the REPOSI study. Clin Res Cardiol 2016;105:912–20.

[121] Díez-Manglano J, Gomes-Martín J, Al-Cheikh-Felices P, Isasi de Isasmendi Pérez S, Díez-Angulo R, Clemente-Sarasa C. Adherence to guidelines and mortality in atrial fibrillation. Int J Cardiol 2014;176:430–6.

[122] Mochalina N, Jöud A, Carlsson M, Sandberg MEC, Själander A, Juhlin T, et al. Antithrombotic therapy in patients with non-valvular atrial fibrillation in Southern Sweden: A population-based cohort study. Thromb Res 2016;140:94–9.

[123] O'Brien EC, Holmes DN, Ansell JE, Allen LA, Hylek E, Kowey PR, et al. Physician practices regarding contraindications to oral anticoagulation in atrial fibrillation: Findings from the Outcomes Registry for Better Informed Treatment of Atrial Fibrillation (ORBIT-AF) registry. Am Heart J 2014;167:601–609.e1.

[124] Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. Circulation 2007;115:2689–96.

[125] Gallagher AM, Rietbrock S, Plumb J, Van Staa TP. Initiation and persistence of warfarin or aspirin in patients with chronic atrial fibrillation in general practice: Do the appropriate patients receive stroke prophylaxis? J Thromb Haemost 2008;6:1500–6.

[126] Gorst-Rasmussen A, Skjøth F, Larsen TB, Rasmussen LH, Lip GYH, Lane DA, et al. Dabigatran adherence in atrial fibrillation patients during the first year after diagnosis: A nationwide cohort study. J Thromb Haemost 2015;13:495–504.

[127] Skeppholm M, Friberg L. Adherence to warfarin treatment among patients with atrial fibrillation. Clin Res Cardiol 2014;103:998–1005.

[128] Rose A, Berlowitz D, Ash A. The Business Case for Quality Improvement Oral Anticoagulation for Atrial Fibrillation. Circ Cardiovasc Qual Outcomes 2011;4:416–24.

[129] Casciano JP, Dotiwala ZJ, Martin BC, Kwong WJ. The costs of warfarin underuse and nonadherence in patients with atrial fibrillation: a commercial insurer perspective. J Manag Care Pharm 2013;19:302–16.

[130] Nielsen PB., Larsen TB, Gorst-Rasmussen A, Skjøth F, Rasmussen LH, Lip GYH. Intracranial hemorrhage and subsequent ischemic stroke in patients with atrial fibrillation: A nationwide cohort study. Chest 2015;147:1651–8.

[131] Nielsen PB, Larsen TB, Skjøth F, Lip GYH. Outcomes Associated With Resuming Warfarin Treatment After Hemorrhagic Stroke or Traumatic Intracranial Hemorrhage in Patients With Atrial Fibrillation. JAMA Intern Med 2017;177:563–70.

[132] Wilson D, Jäger HR, Werring DJ. Anticoagulation for Atrial Fibrillation in Patients with Cerebral Microbleeds. Curr Atheroscler Rep 2015;17:47.

[133] Steiner T. Resumption of oral anticoagulation after warfarin-associated intracerebral hemorrhage yes. Stroke 2011;42:3661–2.

[134] Schulman S. Resumption of Oral Anticoagulation After Warfarin-Associated Intracerebral Hemorrhage: No. Stroke 2011;42:3663–4.

[135] Sandén P, Renlund H, Svensson PJ, Själander A. Warfarin treatment complications do not correlate to cTTR when above 70%. Thromb Res 2015;136:1185–9.

[136] Cancino RS, Hylek EM, Reisman JI, Rose AJ. Comparing patient-level and site-level anticoagulation control as predictors of adverse events. Thromb Res 2014;133:652–6.

[137] Amin A, Stokes M, Makenbaeva D, Wiederkehr D, Wu N, Lawrence JH. Estimated medical cost reductions associated with use of novel oral anticoagulants vs warfarin in a real-world non-valvular atrial fibrillation patient population. J Med Econ 2014;17:771–81.

[138] Nielsen PB, Larsen TB, Skjoth F, Gorst-Rasmussen A, Rasmussen LH, Lip GYH. Restarting Anticoagulant Treatment After Intracranial Haemorrhage in Patients With Atrial Fibrillation and the Impact on Recurrent Stroke, Mortality and Bleeding: A Nationwide Cohort Study. Circulation 2015;132:517–25.

[139] Udsen FW. Health economic evaluation of telehealthcare - can we include “why” and “under what circumstances” telehealthcare is cost-effective in health economic evaluation. Suggested principles for health economic evaluation based on experiences with the Danish "TeleCar North” cluster-randomized trial. 1st ed. Aalborg: Aalborg University Press; 2016.

[140] Ahmad J. The diabetic foot. Diabetes Metab Syndr Clin Res Rev 2016;10:48–60.

[141] Johansen E, Kristensen F. Prevention of diabetic foot ulcers [Forebyggelse af diabetiske fodsår]. Danish J Nurs 2013;6:73–82.

[142] Marsh K, Ijzerman M, Thokala P, Baltussen R, Boysen M, Kaló Z, et al. Multiple Criteria Decision Analysis for Health Care Decision Making—Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. Value Heal 2016;19:125–37.

[143] Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Z, et al. Multiple Criteria Decision Analysis for Health Care Decision Making—An Introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force. Value Heal 2016;19:1–13.

# APPENDICES

# Appendix A. Paper I

## TITLE: WHAT IS THE VALUE OF QUALITY INDICATORS FOR INFORMED DECISION-MAKING IN HEALTHCARE? A BAYESIAN DECISION THEORETICAL AND VALUE OF INFORMATION ANALYSIS PERSPECTIVE

AS Vestergaard, LH Ehlers
*Submitted*[1]

**Abstract:**

The existence of budgetary constraints within healthcare systems is increasingly recognized and prompts additional considerations of cost-effectiveness of interventions before their introduction. A challenge that has arisen within the last decade is the increasing use of quality indicators (QIs) as outcome for interventions that improve quality and safety in healthcare. QIs aim at quantifying the effect of quality improvements and are often not measures that allow for establishment of cost-effectiveness, based on generally acknowledged threshold values, as they do not reflect health gains per se. Opportunities to identify cost-effective quality improvements may be missed if their value for money cannot be established by use of conventional health economic methods. The aim here was to apply Bayesian decision theory and value of information (VOI) analysis to 1) investigate the requirements for QIs to be applicable for economic evaluation and 2) investigate under which circumstances QIs may be used to aid decision-making. Thus, Bayesian decision theory and VOI analysis were used to establish a framework to identify requirements for acceptable QIs and for evaluation of uncertainty when applying QIs for estimation of cost-effectiveness of interventions. Use of QIs introduces an intermediate link in the relationship between interventions and expected net benefit, which increases the total uncertainty pertaining to results. This uncertainty carries a potential cost, interpreted as expected value of perfect parameter information, which should be compared to the expected cost and benefits of resolving that uncertainty. Requirements for an acceptable QI include a correlation between the intervention and the QI and an established correlation between the QI and patient-relevant outcomes. If these correlations are misspecified, the validity of cost-effectiveness results may be compromised. Bayesian decision theory and VOI analysis may comprise a viable framework for evaluation of the cost-effectiveness of quality improvement, allowing for more informed decision-making.

# Appendix B. Paper II

## TITLE: EFFECT OF ANTICOAGULATION ON HOSPITALIZATION COSTS AFTER INTRACRANIAL HEMORRHAGE IN ATRIAL FIBRILLATION – A REGISTRY STUDY

AS Vestergaard, F Skjøth, GYH Lip, TB Larsen
*Stroke (2016) 47(4):979-85*[2]
DOI: 10.1161/STROKEAHA.115.012338

**Abstract:**

*Background and purpose* Intracranial hemorrhage (ICH) is the most feared adverse event with oral anticoagulant therapy (OAC) in patients with atrial fibrillation (AF). The health economic aspects of resuming OAC following ICH are unknown. The aim was to estimate hospitalization costs of thromboembolism and hemorrhage subsequent to ICH in two patient groups with AF surviving the first 90 days post-ICH; 1) patients resuming warfarin therapy within 90 days post-ICH and 2) patients discontinuing therapy.

*Methods* Retrospective data from Danish national registries were linked to identify patients with AF who suffered ICH between 1 January 1997 and 1 April 2011. Study start was 90 days after incident ICH. Mortality was evaluated by use of the Kaplan-Meier estimate. Occurrence of hospitalization-requiring thromboembolism and hemorrhage was used to estimate hospitalization costs by linkage of ICD-10 codes to Danish Diagnosis-Related Group tariffs. The impact of resuming warfarin therapy on average, 3-year hospitalization costs was estimated by use of regression analysis adjusted for between-group differences in baseline characteristics.

*Results* In the inclusion period 2,162 patients suffered ICH; 1,098 survived the first 90 days and were included for analysis and of these 267 resumed warfarin therapy. Therapy resumption reduced the mean 3-year hospitalization cost of hospitalized patients significantly by US$1,588[95%CI:-2,925;-251] and was significantly correlated with fewer hospitalization days per hospitalized patient(-4.6[95%CI:-7.6;-1.6]). The marginal effect of therapy resumption on hospitalization costs per patient was US$-407[95%CI:-815;2].

*Conclusions* Resuming warfarin therapy within 90 days after ICH in patients with AF is associated with a decrease in average hospitalization costs.

# Appendix C. Paper III

## TITLE: A HEALTH ECONOMIC EVALUATION OF STROKE PREVENTION IN ATRIAL FIBRILLATION: GUIDELINE ADHERENCE VERSUS THE OBSERVED TREATMENT STRATEGY PRIOR TO 2012 IN DENMARK

AS Vestergaard, LH Ehlers
*PharmacoEconomics (2015) 33:967–79*[3]
DOI: 10.1007/s40273-015-0281-z

**Abstract:**

*Background* In 2012 the European Society of Cardiology published new guidelines on pharmacological stroke prophylaxis in non-valvular atrial fibrillation. The health economics of adhering to these guidelines in clinical practice remain to be elucidated.

*Objectives* This paper offers a health economic evaluation of two stroke-prophylactic treatment strategies: complete national adherence to the European Society of Cardiology guideline on stroke prophylaxis in atrial fibrillation versus stroke-prophylactic treatment prior to 2012 in Denmark.

*Methods* A cost-utility analysis was performed to compare two treatment strategies. The first strategy reflected national guideline adherence with use of non-vitamin K antagonist oral anticoagulants (i.e. dabigatran etexilate), warfarin, and no treatment. The second strategy reflected observed stroke prophylaxis prior to 2012 with utilization of warfarin, acetylsalicylic acid, and no treatment. A Danish health sector perspective was adopted. A Markov model was designed and populated with information on input parameters from the literature and local cost data. A modeled patient cohort was constructed with a risk profile intended to reflect that of the Danish patient population with atrial fibrillation. The applied outcome was quality-adjusted life-years (QALYs).

*Results* The incremental cost-effectiveness ratio amounted to €3557 per QALY for the guideline-adherent treatment strategy compared with the pre-2012 treatment strategy. This ratio is below a threshold of €25,000 (£20,000) per QALY. Sensitivity analyses revealed that the result was largely robust. All analyses found the guideline-adherent treatment strategy to be cost-effective.

*Conclusions* Guideline adherence is a cost-effective treatment strategy compared with the strategy employed prior to 2012 for pharmacological stroke prophylaxis in atrial fibrillation.

# Appendix D. Paper IV

## TITLE: THE IMPORTANCE OF MEAN TIME IN THERAPEUTIC RANGE FOR COMPLICATION RATES IN WARFARIN THERAPY OF ATRIAL FIBRILLATION: A META-REGRESSION ANALYSIS

AS Vestergaard, F Skjøth, TB Larsen, LH Ehlers
*Submitted*[4]

**Abstract:**

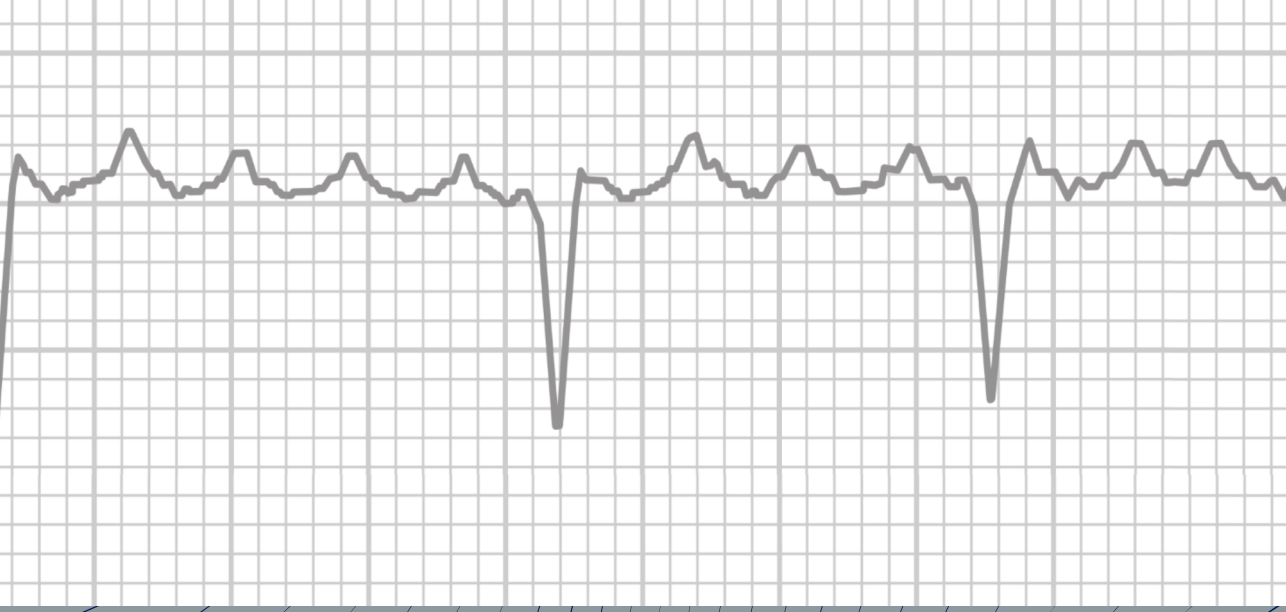
*Background:* Anticoagulation with warfarin is used for stroke prophylaxis in atrial fibrillation (AF) and quality in warfarin therapy is often summarized by the time patients spend within the therapeutic range (percent time in therapeutic range, TTR). The correlation between TTR and occurrence of complications during therapy has been established, but the influence of patient characteristics in that respect remains undetermined.

*Objective:* To examine the association between TTR and complications with adjustment for differences in relevant patient cohort characteristics.

*Methods:* A literature search was conducted in MEDLINE and Embase (2005-2015) to identify studies reporting on use of warfarin therapy of patients with AF and the occurrence of hemorrhage and thromboembolism. The association between mean TTR and major bleeding (MB) and stroke/systemic embolism (SSE) was analyzed by random-effects meta-regression with and without adjustment for relevant clinical cohort characteristics.

*Results*: Of 2169 papers, 35 papers met pre-specified inclusion criteria, holding relevant information on 31 patient cohorts. In univariable meta-regression, increasing mean TTR was significantly associated with a decreased rate of both MB and SSE. However, after adjustment mean TTR was no longer significantly associated with SSE. The proportion of residual variance composed by between-study heterogeneity was substantial for all analyses.

*Conclusions:* Although higher mean TTR was associated with lower complication rates, the strength of the association was decreased when adjusting for differences in relevant clinical characteristics of the patient cohorts. This study suggests that mainly the safety of warfarin therapy increases with higher mean TTR, whereas effectiveness appears not to be substantially improved.



ISSN (online): 2246-1256
ISBN (online): 978-87-7112-964-9

AALBORG UNIVERSITY PRESS