Health Economic Evaluation of Telehealthcare

Udsen, Flemming Witt

DOI (link to publication from Publisher):
10.5278/VBN.PHD.SOCSCI.00064

Publication date:
2016

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

HEALTH ECONOMIC EVALUATION OFTELEHEALTHCARE

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 "TELECARE NORTH" CLUSTER-RANDOMIZED TRIAL

BY
FLEMMING WITT UDSEN

DISSERTATION SUBMITTED 2016

AALBORG UNIVERSITY
DENMARK

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 "TELECARE NORTH" CLUSTER-RANDOMIZED TRIAL

Flemming Witt Udsen

Dissertation submitted as part of the qualification for the Ph.D. degree in Economics at Aalborg University.

AALBORG UNIVERSITY
DENMARK

Dissertation submitted: December 12, 2016

PhD supervisor: Professor Lars Holger Ehlers
Aalborg University, Denmark

Assistant PhD supervisor: Professor Ole Hejlesen
Aalborg University, Denmark

PhD committee: Professor Kjeld Møller Pedersen (chair)
University of Southern Denmark / Aalborg University
Denmark

Professor Marjon van der Pol
University of Aberdeen, Scotland

Professor Jan Sørensen
University of Southern Denmark, Denmark

PhD Series: Faculty of Social Sciences, Aalborg University

ISSN (online): 2246-1256
ISBN (online): 978-87-7112-847-5

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: 118 pages (2,400 characters incl. spaces).

CV

Flemming Witt Udsen had 9 years of previous non-academia experience before entering into a Ph.D. qualification. He has worked as a researcher within the health care sector in the regional administration of Central Denmark Region, where focus was on conducting health technology assessments and specifically evaluating consequences of new technologies on the organization and costs of delivering healthcare. These projects were conducted on behalf of the Danish Health Authority, the State Serum Institute, Danish Regions and Danish hospitals. In the Ph.D. qualification program, Flemming has taught health policy, health economics and non-experimental research designs at Aalborg University.

Flemming is originally educated within the field of organization theory and organizational behavior at Aarhus University and have been interested in quantitative methods for analyzing the effects of structure and behaviors on organizational results. Flemming is still affiliated lecturer and external examiner here.

Flemming holds a Master's degree from 2004 in Management and Economics from Aarhus University.

ENGLISH SUMMARY

One of the most prevalent diseases in the world is chronic obstructive pulmonary disease (COPD). 64 million people are estimated to have COPD leading to 3 million annual deaths. As many as 430,000 Danish citizens could have COPD. The cost of managing COPD is therefore considerable; in Denmark for example, 10 percent of the total annual healthcare budget for citizens older than 40 years can be related to COPD. In 2012, telehealthcare was highlighted as a technology with the potential to increase primarily the health-related quality of life for COPD patients, while also relieving some of the fiscal pressure in the healthcare sector. But this potential had not been demonstrated in large-scale randomized evaluations with high quality health economic evaluation. A pragmatic clinical trial with trial-based economic evaluation, the Danish TeleCare North trial, was therefore designed in order to assess the effectiveness and cost-effectiveness of a particular telehealthcare solution implemented in North Denmark Region and the design of the trial and its results are presented in this thesis. At the same time, the TeleCare North trial was embedded in a "national action plan for the dissemination of telemedicine" meant to test telehealthcare technology, implementation models and potential benefits of telehealthcare across different initiatives and settings. This action plan was used as a foundation for deciding whether or not to disseminate telehealthcare to different patient groups in 2015.

Contrary to expectations, the TeleCare North trial demonstrated no difference in health-related quality of life and the telehealthcare solution was not cost-effective for all included COPD patients. But there was a potential to target the solution to patients with severe COPD, because they were likely to be most cost-effective. The results also indicate that implementation could have a strong impact on cost-effectiveness, more so than health- or socio-demographic factors. The results from the TeleCare North trial were used directly in a decision to implement the telehealthcare solution to patients with severe COPD in Denmark and lead to considerable debate nationally. This debate is an actual account of the usefulness of health economic evaluation for decision making that could also be used to adapt the health economic evaluation approach. It is argued that one lesson from the national decision process is that there is no getting around making decisions that are based on underlying evaluations with a balanced outcome focus and randomization. Another lesson is that telehealthcare interventions are difficult to implement and that effects of implementation can be hard to foresee, quantify and make visible, which challenges the current approach to health economic evaluation.

Furthermore, it is argued that trial-based economic evaluation of telehealthcare in Denmark has gained momentum, which creates the pressing issue of designing even better trial-based economic evaluations that simultaneously address challenges with

isolating cause-and-effect relationships, learning curves and with generalizing results. Based on recent developments in realist evaluation, experiences with conducting the economic evaluation of TeleCare North and participating in the national decision debate, four principles for health economic evaluation of complex telehealthcare interventions is outlined in order to facilitate more informed health economic designs of telehealthcare in the future that should ultimately answer if telehealthcare is cost-effective, for whom, why and under what circumstances.

DANSK RESUME

En af de mest udbredte sygdomme i verden er kronisk obstruktiv lungesygdom (KOL). Alene i Danmark skønnes 430.000 danskere at have KOL, hvilket leder til 3.300 årlige dødsfald. Omkostninger til behandling og pleje af KOL-patienter er derfor også betydelige: I Danmark skønnes 10 procent af det samlede årlige sundhedsbudget for borgere ældre end 40 år at være relateret til KOL. I tiden omkring 2012 blev telemedicin fremhævet som en teknologi, der havde potentiale til at øge livskvaliteten for KOL-patienter, samtidig med, at det kunne reducere KOL-relaterede omkostninger. Men potentialet var ikke blevet påvist i store lodtrækningsforsøg med indlejret sundhedsøkonomisk evaluering. Et stort lodtrækningsforsøg med sundhedsøkonomisk evaluering blev derfor designet i regi af det danske TeleCare Nord initiativ. Formålet var at vurdere effektiviteten og omkostningseffektiviteten af den valgte telemedicinske løsning og dets resultater er præsenteret i denne afhandling. TeleCare Nord initiativet var på samme tid en del af en national handlingsplan for udbredelse af telemedicin, som skulle sikre, at der blev testet telemedicinsk teknologi, implementeringsmodeller og effekter af telemedicin på tværs af forskellige regioner. Handlingsplanen blev brugt som et beslutningsgrundlag i 2015 for, hvorvidt bestemte telemedicinske løsninger kunne bruges af forskellige patientgrupper.

Mod forventning viste TeleCare Nord ingen forskel i livskvalitet for de inkluderede KOL-patienter og den telemedicinske løsning var heller ikke generelt omkostningseffektiv for alle omfattede KOL-patienter. Men der var et potentiale i at målrette løsningen til patienter med svær KOL, fordi sandsynligheden for at dette ville være omkostningseffektivt var særlig stor. Resultaterne viser også, at implementering kan have en stærk indvirkning på omkostningseffektiviteten, måske endda mere end kliniske eller socio-demografiske faktorer. I 2015, blev resultaterne fra TeleCare Nord brugt direkte i en beslutning om at implementere den telemedicinske løsning til patienter med svær KOL i Danmark. Men beslutningen førte også til en betydelig national debat. Denne debat illustrerer et konkret eksempel på brugbarheden af sundhedsøkonomisk evaluering, der kan anvendes til at genoverveje måden, hvorpå sundhedsøkonomisk evaluering foretages. I afhandlingen argumenteres der for eksempel for, at man ikke kommer udenom at træffe beslutninger, der er baseret på underliggende evalueringer, der har et afbalanceret resultat-fokus og lodtrækning. Der argumenteres også for, at telemedicinske interventioner er vanskelige at implementere og at virkningerne af implementering kan være svære at forudse, kvantificere og synliggøre, hvilket udfordrer den nuværende tilgang til sundhedsøkonomisk evaluering.

Endvidere hævdes det, at sundhedsøkonomisk evaluering indlejret i kliniske forsøg af telemedicin i Danmark har fået en vis anerkendelse, men at dette også

nødvendiggør at der designes endnu bedre sundhedsøkonomiske evalueringer. Udover at skulle løse udfordringer med at isolere årsag-virkning sammenhænge, skal sundhedsøkonomisk evaluering også kunne håndtere læringskurver og generalisering af resultater. Baseret på den seneste udvikling i virkningsevaluering, erfaringer med at gennemføre den sundhedsøkonomiske evaluering af TeleCare Nord og som part i den nationale debat om telemedicin, opstilles der fire principper for sundhedsøkonomisk evaluering af komplekse telemedicinske interventioner som er skitseret med henblik på at starte en ny agenda for sundhedsøkonomisk forskning i telemedicin, der i sidste ende skal kunne svare på, om telemedicin er omkostningseffektiv, for hvem, hvorfor og under hvilke forudsætninger.

ACKNOWLEDGEMENTS

This PhD dissertation was carried out during my employment with the Danish Center for Healthcare Improvements (DCHI) and would never have been feasible without the opportunity that was handed to me by the head of the center, Lars Ehlers, who also served as my main supervisor. Over the past years, I have learned a lot about research process from you: your knowledge of how to develop research projects has been undisputable as have your approach to teach young researchers to become independent.

I would also like to thank my co-supervisor, Ole Hejlesen, at the Department of Health Science and Technology. Thanks for sharing your knowledge of telemedicine with me and for trusting me with the primary effectiveness analysis in the trial.

The research in this dissertation would not have been possible without the funding from North Denmark Region and without the cooperation with the TeleCare North secretariat. It has been a privilege to work with such a professional and skilled team responsible for being the first in Denmark to implement a large-scale cross-sectorial telehealthcare initiative. A special thanks to Tina Heide, Kuno Kudajewski, Helen Houmøller Rasmussen and Lone Mylund for your patience in answering my endless questions and for facilitating contact with the many people who have helped with the data collection in the North Denmark Region and the municipalities here.

A thank you also goes to my fellow TeleCare North Ph.D. students, Jannie Kristine Bang Christensen and Pernille Heykendorff Lilholt for cooperation on the project and articles. Similarly, I have also benefitted from many skilled and funny Ph.D. students at DCHI in the past four years. I would like to thank you all for fruitful debates - often marginally related to health economics. I am particularly indebted to Ann Lykkegaard Sørensen and Michael Hvidberg for invaluable encouragement when I lost motivation for the Ph.D. project along the way.

Finally, I would like to dedicate this thesis to my wife, Lise Marie, and to my two daughters, Martha and Ida. Lise; you made it possible for me to go through a Ph.D. qualification rather late in life and it seems that your patience and faith in me is undeservingly endless. At times, it has not been easy for me to juggle this Ph.D. thesis with hopes of being a good husband and father. Especially not towards the end, when I thought that it would never end and I was challenged by reduced cognitive abilities and motivation loss. I am grateful that you care a lot for me and little for qualifying for Ph.D. degrees, for telehealthcare or health economics. Martha and Ida; I am humbled by feeling that being a good dad to you is the best qualification that I could ever hope to aspire to.

TABLE AND FIGURES

LIST OF ABBREVIATIONS

CAT: COPD Assessment test
CE: Conformité Européenne
CEA: Cost-effectiveness analysis
CMO: Context-mechanism-outcome
COPD: Chronic obstructive pulmonary disease
CRT: Cluster-randomized trial
CUA: Cost-utility analysis
DKK: Iso-code for the Danish currency (Danish Crowns)
DREAM: Danish Rational Economic Agents Model
FTE: Full-time equivalent
GDP: Gross domestic product
GP: General practitioner
GSM: Global System for Mobile Communications
HTA: Health technology assessment
ICER: Incremental cost-effectiveness ratio
IT: Information technology
KIH: Klinisk integreret hjemmemonitorering
mMRC: Medical Research Council Dyspnea Modified Scale
MRC: Medical Research Council
NICE: National Institute for Health and Care Excellence
OECD: Organisation for Economic Co-operation and Development
QALY: Quality adjusted life year
QI: Quality improvement
PDSA: Plan-Do-Study-Act
RCT: Randomized controlled trial
RQ: Research question
UK: United Kingdom
US: The United States
WHO: World Health Organization
WSD: Whole system demonstrator project

LIST OF INCLUDED PUBLICATIONS

Paper 1 **A systematic review of the cost and cost-effectiveness of telehealth for patients suffering from chronic obstructive pulmonary disease.**

Flemming Witt Udsen, Ole Hejlesen, Lars Holger Ehlers
Journal of Telemedicine and Telecare 2014; 20(4): 212-220

Paper 2 **Effectiveness and cost-effectiveness of telehealthcare for chronic obstructive pulmonary disease: study protocol for the Danish "TeleCare North" pragmatic cluster-randomized trial.**

Flemming Witt Udsen, Pernille Heyckendorff Lilholt, Ole Hejlesen, Lars Holger Ehlers
Trials 2014; 15(178)

Paper 3 **Telehealthcare for patients suffering from COPD: Effects on health-related quality of life - Results from the Danish "TeleCare North" cluster-randomised trial**

Pernille Heyckendorff Lilholt, Flemming Witt Udsen, Lars Holger Ehlers, Ole Hejlesen
Submitted manuscript

Paper 4 **Cost-effectiveness of telehealthcare to patients with chronic obstructive pulmonary disease: Results from the Danish "TeleCare North" cluster-randomized trial.**

Flemming Witt Udsen, Pernille Heyckendorff Lilholt, Ole Hejlesen, Lars Holger Ehlers
Submitted manuscript

Paper 5 **Heterogeneity analysis of telehealthcare to patients with chronic obstructive pulmonary disease: The case of the Danish "TeleCare North" cluster-randomized trial.**

Flemming Witt Udsen, Pernille Heyckendorff Lilholt, Ole Hejlesen, Lars Holger Ehlers
Manuscript

Additional papers published in Ph.D. qualification period by the author (not included in the dissertation)

Evaluation of a comprehensive EHR based on the DeLone and McLean model for IS success: Approach, results, and success factors

Claus Bossen, Lotte Groth Jensen, Flemming Witt Udsen
International Journal of Medical Informatics 2013; 82(10): 940-953

Boundary-Object Trimming: On the Invisibility of Medical Secretaries' Care of Records in Healthcare Infrastructures

Claus Bossen, Lotte Groth Jensen, Flemming Witt Udsen
Computer Supported Cooperative Work 2013; 23: 75-110

Chapter on the results of the health economic evaluation of TeleCare North
In "TeleCare Nord afslutningsrapport – Et telemedicinsk storskalaforsøg i Region Nordjylland" (In Danish)
Region Nordjylland 2015

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

The “Global Burden of Disease Study” has recently published prevalence and incidences estimates of acute and chronic diseases for 188 countries from 1990 to 2013 (1). One conclusion is that the global prevalence of chronic disease, e.g. cancer, cardiovascular disease, chronic pulmonary disease, diabetes, mental disorders and musculoskeletal disease, is high (1). And as a result of the ageing of the World’s population and better treatment of previously life-threatening diseases, the incidence of chronic diseases has risen (1). Estimates made by the World Health Organization (WHO) make chronic diseases responsible for roughly half of the global burden of disease (46%) (2), numbers that have been echoed by Centers for Disease Control and Prevention that has estimated that roughly half of all American adults have one or more chronic health conditions (3). Based on a prevalence study made in the Capital Region of Denmark (4), the Danish Health Authority have a slightly lower estimate of around one-third of all Danish citizens (5).

One of the most prevalent chronic diseases is chronic obstructive pulmonary disease (COPD) (1,6,7). The WHO projects that 64 million people have COPD worldwide which account for more than 3 million annual deaths (8). COPD has been shown to be a major health problem in the European Union with between 4-10% of citizens diagnosed with COPD within member states (9). In Denmark, the Danish Lung Association estimates that as many as 430,000 Danish citizens could suffer from COPD (10) with a scientific study setting the prevalence to 9% (11). According to the latest analysis on the disease burden in Denmark, around 3,300 Danish citizens die due to COPD each year (6% of total number of deaths) (12). Additionally 2,500 die each year of other causes related to COPD (13). This makes COPD the fourth largest cause of death in Denmark (13).

Because of the high prevalence, the costs of treating and caring for COPD patients are considerable. There is no systematic data collection on the global costs of treating and caring for COPD patients, but a survey conducted in 2003 in seven countries (Canada, France, Italy, the Netherlands, Spain, the UK and the US) demonstrates “significant” healthcare sector costs (14). In Denmark, previous studies have concluded that as much as 10 percent of the total annual healthcare budget for citizens older than 40 years is related to COPD and the recent disease burden study estimated costs in the vicinity of £250 million per year, primarily due to hospital admissions and productivity loss (12).

This creates a burning platform, because many countries have at the same time been struggling to recover from a global recession since 2008 (15). In Denmark, the Danish Ministry of Finance downgraded the economic outlook for both 2011 and 2012 to an increase in Gross Domestic Product (GDP) of only 1% in December

2011 (from 1.3 and 1.8%, respectively) (16). According to the Organisation for Economic Co-operation and Development (OECD) economic survey for Denmark published in early 2012, public expenditure had risen in Denmark from 51% of GDP in 2007 to 58% in 2010 (17). This financial context combined with projections of the demographic development, the prevalence of disease and a potential for future shortages in recruiting healthcare personnel was seen as constituting a threat to the Danish publically funded healthcare system by challenging its fiscal sustainability (18). However, another OECD analysis had argued that achieving better health outcomes at lower costs in Denmark were possible (19). It was argued, that this end could be achieved by focusing more on patient pathways across state, regions and municipalities thereby reducing waste by duplication of tasks, having more control over spending and better incentives for providing cost-effective services across sector-boundaries (19). Given the structure and funding of the Danish health care system, the government had ample opportunity to encourage and potentially put political and financial pressure on the municipalities - and in particularly the regions - for containing public expenditure e.g. by focusing more on cross-sectorial initiatives and -technologies (20).

Telehealthcare was highlighted as one of different solutions that could be implemented as an integrated part of patient pathways across sectors in order to meet some of the threats facing the healthcare system (20). It was argued that telehealthcare had the potential to address challenges with an ageing population and a rising number of patients with chronic disease thereby also relieving some of the fiscal pressure in the healthcare sector (20,21). There had previously been some experience with telehealthcare in Denmark, also for COPD patients (22–24), but mostly mono-sectorial small-scale demonstration projects, with only 40% of them in current operation and without having formally evaluated effects (25). Internationally though, telehealthcare had been systematically reviewed for COPD patients and this evidence suggested a potential for increased health-related quality of life and reduced hospital contacts (26,27). But reviews also demonstrated a shortage of randomized studies with cost and cost-effectiveness evaluation and that those that did exist were small feasibility or pilot studies, meaning that the evidence base for the effects and costs of telehealthcare were relatively limited (26,27). These results were not unique at the time. Systematic reviews had been conducted with similar results for other chronic diseases or as syntheses of several other chronic diseases (e.g. asthma, cardiovascular disease, chronic heart failure, diabetes and hypertension) (28–32).

In 2012, there was therefore a scientific need for more and larger randomized studies with embedded cost-effectiveness analyses that could be used to assess whether or not telehealthcare could in fact deliver on its promises in settings closer to routine practice. A pragmatic clinical trial, the Danish TeleCare North trial, was designed in order to assess the effectiveness and cost-effectiveness of a particular

telehealthcare solution implemented in North Denmark Region and the results from this trial are presented in this thesis.

At the same time, the TeleCare North trial had a connection to a Danish "national action plan for the dissemination of telemedicine" that was published in 2012 (33). The intention with the action plan was to ensure that new and larger telehealthcare initiatives were tested and evaluated in Denmark from 2013-2015 and that they included cost-evaluations (33). In addition to the TeleCare North trial, the action plan also funded a second large-scale telehealthcare initiative with formal - but differently designed – economic evaluation of its telehealthcare solution (33). Based on the generated evidence from the action plan, a decision was taken in late 2015, to implement telehealthcare to patients with COPD.

This decision-process is interesting to describe. In many countries there are an agreed form and process for conducting economic evaluation that are used in national health care resource allocation. In the United Kingdom (UK) for instance, the National Institute for Health and Care Excellence (NICE) systematically assess new medical technologies in order to estimate both their effectiveness and cost-effectiveness (34). Their program is comprehensive and has since year 2000 covered 566 Health Technology Assessments (HTAs) with cost-effectiveness evaluation of pharmaceuticals, medical devices, diagnostics and clinical procedures (34,35). Healthcare resource allocation by some form of HTA approach including a standard methodology for including economic analyses is also implemented in many other countries such as New Zealand, Australia, Korea, Thailand, Canada, Scotland, Ireland, Germany, the Netherlands, Norway and Sweden. But this is not the case in Denmark. In fact, broad and systematic evaluation has in general been neglected since the late 1990s and never really pursued whole-heartedly. Responsibility for conducting HTAs was originally placed at the Danish Health Authority but funds had been reduced over several years (36,37). In 2008, the Danish Health Authority decided to downsize most of its activities on HTAs (it continued to assess vaccines for the national child vaccination program and other pre-planned HTAs) (36,37). Some HTAs with health economic evaluation are instead conducted by the five Danish regions in collaboration, but no more than one or two are published each year and the process for choosing technologies for assessment is more ad hoc than systematic.

Most Danish national politicians have historically been reluctant to use health economic evaluations, especially if these evaluations take the form of cost-utility analyses; presumably because it ultimately involves putting an explicit price on the lives or health of its citizens. Even though new guidelines for economic evaluation used for reimbursement of medicine was published in 1997 (38), a broad and systematic approach to decision-making that uses economic evidence that "smells" like standard cost-effectiveness evaluation remains a politically hot topic to this day. Up until now, the Danish regions have instead made decisions on the use and

funding based on recommendations made by two advisory expert panels ("Koordineringsrådet for ibrugtagning af sygehusmedicin (KRIS)" (39) and "Råd for anvendelse af dyr sygehusmedicin (RADS)" (40) (no English translations). But these panels only evaluate hospital medicine and are not allowed to include cost in their recommendation calculus. Although the Danish Regions in the spring of 2016 decided to establish a new advisory organ called "Medicinrådet" (no English translation), that should make decisions on whether or not to buy expensive hospital medicine by including costs in the calculus, it is still unclear if this organ would later include recommendations for other technologies than hospital medicine (the very last paragraph in the document describing the overall principles for "Medicinrådet" hints that the model could be expanded to other technologies in the future (41)) or what type of economic evidence should be included to make these decisions.

There are two main implications of this. One is that no consensus exists on whether or not health economic evidence should be included at all in healthcare decision-making on technologies other than pharmaceuticals; what characteristics this health economic evidence should have or the best way to include and combine economic evidence. Another implication is that health economic research is usually decoupled from national prioritization. But what happens if results from a large-scale cost-utility evaluation alongside a clinical trial were to be used in a national action plan combining results from other initiatives meant to inform a decision on whether or not to disseminate telehealthcare to different patient groups? How would economic evidence be applied by decision-makers?

This actual account of the usefulness of health economic evaluation in a national decision-process is combined with health economic theory, experiences with conducting the TeleCare North evaluation and topics from realistic evaluation in order to suggest a new research agenda for an adapted approach to trial-based economic evaluations of complex telehealthcare interventions. Four normative principles is outlined in order to facilitate more informed health economic designs of telehealthcare in the future that should ultimately answer if telehealthcare is cost-effective, for whom, why and under what circumstances.

1.1. RESEARCH QUESTIONS

1.1.1. RESEARCH QUESTION ONE (RQ1): THE WITHIN-TRIAL EFFECTIVENESS AND COST-EFFECTIVENESS OF THE TELECARE NORTH TRIAL

As described above, the results and conclusions from the evaluation of the TeleCare North trial are important and interesting in themselves, because larger comparative studies with more rigorous health economic evaluation have been requested

scientifically for telehealthcare in several systematic reviews. The first research question for this thesis therefore becomes:

> "*Is the chosen telehealthcare solution for COPD patients adopted in the TeleCare North initiative effective and cost-effective?"*

1.1.2. RESEARCH QUESTION TWO (RQ2): APPLICATION OF EVIDENCE IN A NATIONAL DECISION

But the results are also interesting because Danish healthcare decision stakeholders would use, combine or at least relate to the various new economic evidence that were provided within the action plan to the decision-making process. How would that unfold? What would they focus on? The research question therefore becomes:

> "*How was evidence applied in a national Danish decision-making context of whether or not to adopt telehealthcare?"*

1.2. OUTLINE OF THE THESIS

The thesis outline is presented in Figure 1. The thesis sets out by describing relevant aspects of the theoretical foundation used (*chapter 2*) and the methodology applied (*chapter 3*). *Article one* (the systematic review in *appendix 1*) is essentially trying to answer how much evidence of health economic benefits of telehealthcare to COPD patients there existed in 2013 and how sound this evidence was. The existing evidence was informed a detailed design of a health economic evaluation nested in the cluster-randomized trial in TeleCare North. *Article two* therefore describes a design that would have a higher quality by including more recommended characteristics in a sound health trial-based health economic evaluation (see *appendix 2*).

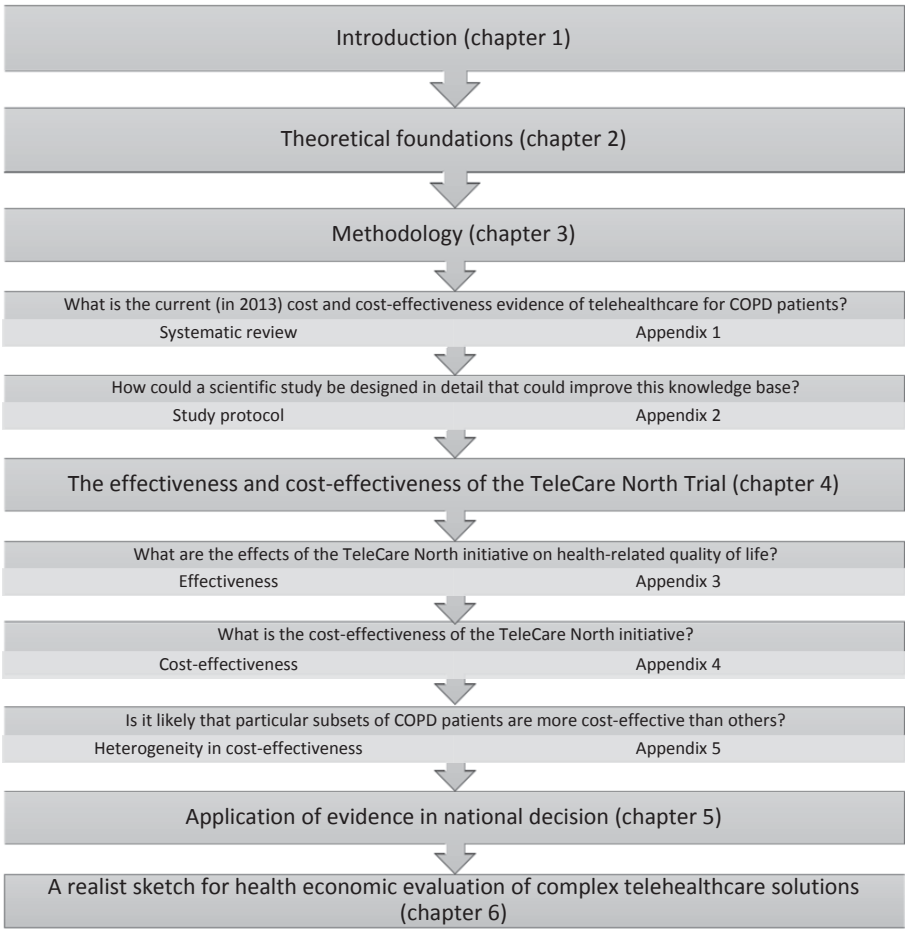
In order to answer **RQ1**, a series of empirical publications were planned and are divided in further three publications. *Article three* is the main effectiveness results from the TeleCare North trial that focuses on assessing gains in health-related quality of life (see *appendix 3*). Then the main results from the economic evaluation of TeleCare North are presented in *article four* (see *appendix 4*) for all of the included patients in accordance with the trial protocol. Finally, knowledge of potential sources of heterogeneity in cost-effectiveness is presented in *article five* (see *appendix 5*). This study seeks out potential subgroups of patients that are more or less likely to be cost-effective. A short summary of these articles is presented in *chapter 4*.

To answer **RQ2**, a document analysis is performed that analyzes how evidence both from initiatives within the national action plan and from abroad was incorporated

into the national decision of dissemination of telehealthcare in Denmark that took place in 2015-2016 (*chapter 5*).

Finally, a discussion is presented in *Chapter 6*, which seeks to make sense of and build additional theory for trial-based economic evaluation that are based on the theory presented in chapter 2, the experiences gained from the design process and results from the cost-effectiveness evaluation of TeleCare North as well as the analysis of the national decision. Based on this sense-making process, principles for "a realist sketch for health economic evaluation of complex telehealthcare interventions" is outlined. This is meant as a small theoretical contribution in order to facilitate more informed health economic designs of telehealthcare in the future.

Figure 1: Thesis outline. Source: Own contribution.



1.3. CONTRIBUTION(S) OF THE THESIS

In addition to providing new scientific evidence on the effectiveness; cost structure and cost-effectiveness of telehealthcare to COPD patients in RQ1, this thesis has further two main contributions.

Firstly, by changing unit of analysis and applying a qualitative research methodology, the thesis seeks to place the TeleCare North trial in a health policy resource allocation context by describing and commenting on the national decision in late 2015 and early 2016. This is meant as an illustration of "what went on" nationally in the resource allocation for telehealthcare in order to provide an actual account, as opposed to a theoretical examination, of the usefulness of health economic evaluation for decision-making. The intention is also to draw out themes that can used to build additional theory on economic evaluation of telehealthcare that a presented in chapter 6.

Secondly, by briefly attempting to switch philosophy of science standpoint from positivism to critical realism and investigate a particular evaluation methodology called "realistic evaluation" often applied in sociology, principles for an integrated evaluation process called a "realist sketch for health economic evaluation of complex telehealthcare interventions" is proposed that seek to transcend a methodological discussion of which existing evaluation approach is best suited for telehealthcare research. This is meant as a suggestion for a new research agenda for health economic evaluation of telehealthcare that, in addition to seeking to answer whether or not telehealthcare interventions work and for whom, also begins to focus on why and under what circumstances.

Conducting the evaluation for this Ph.D.-thesis has been an unusual task.

- The results from particularly the economic evaluation have been used directly in a decision to disseminate the telehealthcare solution throughout Denmark. This gave rise to a national debate on the methodology of economic evaluations and the strengths and weaknesses of trial-based economic evaluation and other types of economic evaluations. I was part of this debate in the media and as invited speaker and panel member on conferences. It also entailed that I have used two periods throughout the Ph.D.-study - one in 2015 and another in the spring of 2016 - where I provided additional analyses and counseling for the Danish government in order for them to calculate and revise a national business case for telehealthcare.

- The reporting of the health economic evaluation was coordinated with the effectiveness evaluation. This meant that the empirical economic publications waited almost a year for submission. It also meant that I became involved in the effectiveness evaluation of the TeleCare North trial by conducting most

data management and all quantitative analyses used in the scientific reporting of main and subgroup outcomes described in the trial protocol and by conducting further analyses of secondary effectiveness outcomes for North Denmark Region and the Danish Government of particularly mortality for the included patients.

- The amount of attention the economic evaluation received was overwhelming; in part because a summary report of the main conclusions from the economic evaluation were publically available, while the scientific articles remained unpublished. For example in the weeks after the main results from the economic evaluation was presented in November 2015, I received more than 200 e-mails and phone calls from people in Denmark and abroad making enquiries of further details from the study or requesting comments on the national decision process.

- The clinical trial entailed a large amount of practical data collection and data management work for me, which is unusual for authors of economic evaluations. By ultimately including 1,225 patients, 2,200 paper-based questionnaires each with more than 80 variables were sent out and entered manually. 200 patients who did not respond to questionnaires were contacted by telephone and a much used phone hotline for patients, practitioners and institutions were established and manned. Similarly, the register data task was large by including not only national register data on all healthcare contact patterns, but also specific costs categories from all municipalities in North Denmark Region. Including municipalities in economic evaluations of healthcare initiatives has not been done often in Denmark and elsewhere and ultimately ended in 120 datasets from the included 10 municipalities alone.

CHAPTER 2. THEORY

2.1. CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)

As professionals in the field of health economics, we risk forgetting the people behind the COPD diagnosis and treatment alternatives. I would therefore like to begin this theoretical chapter by checking in with reality and give the reader a quick sense of life with COPD. Besides having a disease to fight, life with COPD can be hard (42). Feelings of helplessness can make it difficult to be a COPD patient and there is a constant fear of suffocating:

> *"Anyone who has tried not to be able to breathe, know how scary and unpleasant an experience it is. As a COPD patient you may experience that even the smallest things can cause shortness of breath - to eat, talk on the phone or just to get nervous. Therefore, you tend to keep to yourself and isolate yourself in your home […]. So you can easily feel lonely and helpless, which can reinforce your anxiety. It can be difficult for family, colleagues and others to understand how you feel, or get them to take your condition seriously. They may think you just have to "pull yourself together"* (translated from Danish) (43).

It is actually possible to feel the effects of breathlessness with a small exercise (you should try it before you read further!): Take ten deep in- and exhalations and then put a straw to your mouth while holding your nose. Then go up and down a set of stairs while breathing only through the straw and see how long you can keep doing that. Imagine everyday life under these conditions…

2.1.1. CHARACTERIZATION OF COPD

Clinically COPD is defined as "*a disease characterized by persistent airflow limitation that is usually progressive and associated with an enhanced chronic inflammatory response in the airways and the lung to noxious particles or gases*" (44, p2). COPD cannot be cured and the clinical course is a gradually deteriorating lung function without patients necessarily being aware of it, which often leads to a late diagnosis at an advanced severity stage of the disease (45). The main symptoms of COPD are dyspnea, decreased exercise tolerance, wheezing, recurrent lung infections, "smoker's cough" and abnormal sputum (46). COPD patients can experience some variations in their usual symptoms over time, e.g. during the day or from day to day (47). Yet, it is also normal that patients (depending on the severity of COPD) occasionally experience a number of events (exacerbations), where symptoms become more severe than normal variations, which requires

changes to the usual management of their disease (44,48). A number of other conditions are also associated with COPD, e.g. cardiovascular disease, metabolic syndromes such as diabetes, osteoporosis and mental illness (49,50). Smoking (or even second-hand smoking) is by far the main reason for developing COPD, but air pollution, childhood lung infections and genetics can also play a role (7,51).

2.1.2. DIAGNOSIS AND ASSESSMENT OF COPD

Diagnosis of COPD can be based on the history of symptoms, patient questionnaires, physical examination, pulmonary function tests, blood-gas analyses and chest radiography (46). Part of the goal with diagnosis is to determine the severity of the disease that is used to determine prognosis and treatment strategy (44). Pulmonary function test by spirometry is the most widely accepted instrument in diagnosing airflow limitation and in assessing COPD severity (44). A spirometer expresses airflow limitation by two variables: Forced Expiratory Volume in one second (FEV1) and Forced Vital Capacity (FVC). When FEV1/FVC < 0.70, it constitutes a COPD diagnosis (44). The severity of airflow limitation can be assessed with the FEV1% predicted (measured FEV in percentage of the expected value given the age, gender, height and race of the patient) (44). Table 1 presents this severity stratification, which was also applied in article 3 and 5 to analyze if there could be a systematic variation in effectiveness and cost-effectiveness across COPD severity.

Table 1: Spirometric classification of COPD severity. Source: (44).

Severity	FEV1 value in % compared to predicted FEV1 value
Mild	FEV1 ≥ 80 %
Moderate	50 % ≤ FEV1 < 80 %
Severe	30 % ≤ FEV1 < 50 %
Very severe	FEV1 < 30 %

The patients' subjective experiences with dyspnea can be used to complement spirometry and allow for a more complete picture of the COPD condition and especially two questionnaires are widely used. The Medical Research Council Dyspnea Modified (mMRC) Scale is a questionnaire that are comprised of five statements that describe experiences with breathlessness from no experience (category one) to almost invalidating breathlessness (category 5) (52). The COPD Assessment Test (CAT) consists of 8 questions measuring health status that can result in a score from 0-40 (53). For both questionnaires, a higher score is representative of more severe experiences with COPD. Both questionnaires can be

self-administered or filled in by an interviewer by asking patients to indicate the phrase that best describes their experience.

In 2013, around the time of the initiation of TeleCare North, updated guidelines for assessing COPD severity was published that in addition to lung function tests also account for subjective experiences and the risk of exacerbations (44). A more representative picture of severity is now a combined risk assessment that are divided into four categories designated by A, B, C and D, as demonstrated in Table 2 (reproduced from (44)). This classification was unfortunately not incorporated in the trial.

Table 2: Combined COPD assessment classification of severity. Source: (44).

A = Low risk, low symptom burden
Low symptom burden (mMRC of 0-1 OR CAT score < 10) AND
FEV1 of 50% or greater (old GOLD 1-2) AND low exacerbation rate (0-1/year)
B = Low risk, higher symptom burden
Higher symptom burden (mMRC of 2 or more OR CAT of 10 or more) AND
FEV1 of 50% or greater (old GOLD 1-2) AND low exacerbation rate (0-1/year)
C = High risk, low symptom burden
Low symptom burden (mMRC of 0-1 OR CAT score < 10) AND
FEV1 < 50% (old GOLD 3-4) AND/OR high exacerbation rate (2 or more/year)
D = High risk, higher symptom burden
Higher symptom burden (mMRC of 2 or more OR CAT of 10 or more) AND
FEV1 < 50% (old GOLD 3-4) AND/OR high exacerbation rate (2 or more/year)

2.1.3. THE DANISH HEALTH CARE DELIVERY MODEL

In order to provide a context for the evaluation, but also to give the reader a sense of which perspectives can be chosen for health economic evaluation of technologies in Denmark, it is important to understand how the health care system is organized and funded and which sectors are currently involved in managing COPD.

Denmark has a tax financed health care system for all Danish citizens. It covers treatment and care and subsidizes medicine and medical devices (54). The public sector in Denmark has the main obligation for providing health care; an obligation which is divided between the Danish state, five regions and 98 municipalities (55). In this division, the regions play the largest role by having the responsibility for both psychiatric and somatic hospital services and primary care (55). The role of the 98 Danish municipalities in health care is largely confined to disease

prevention, home care and rehabilitation. The Danish state is the governing-body, advising and regulating the healthcare sector (55). A private sector exist with a few private hospitals and the entire operation of primary care is contracted out to independent general practitioners (55,56). But provision of healthcare is generally a public responsibility with relatively decentralized delivery (55).

Since healthcare funding stems from general taxation at the state and municipality level, the regions play a reduced political role (54). Healthcare activities in regions are financed by block grants from the state and payment based on regional activity (remuneration based on Diagnose-Related Groups (DRG)). Municipalities are co-funding the operation of regions by an activity-based payment. Municipalities receive a block grant from the state but otherwise collect taxes directly from citizens. However, the state sets a yearly ceiling on municipality taxes (54,55). Private co-payment or user fees exist (around 17% of the total healthcare budget), mainly for medicine, dental care, physiotherapy, chiropractors and for stay at nursing homes (54,56).

2.1.4. USUAL PRACTICE FOR MANAGING COPD IN DENMARK

Since 2002, COPD has been given priority as one of eight diseases in the Danish public health policy (57). The Danish Health Authority has subsequently strengthened systematic efforts to prevent these diseases and to promote health within these diseases in the Danish health care system (58,59). According to guidelines, a typical management strategy for COPD patients in Denmark consist of a combination of monitoring, treatment and rehabilitation (60) and concrete management activities are as follows:

Several training and counseling opportunities exist (60). All COPD patients that are current smokers are offered motivational counseling and help to stop smoking (e.g. by medicine) and are informed about the risks of continued smoking (60). All patients are encouraged to exercise and patients with at least moderate COPD can be referred to individually tailored programs for physical training (60). Patients with overweight or patients that experiences unexpected weight loss can be referred to nutrition counseling and –control (60). All COPD patients are offered education in the characteristics of the disease and the importance of adhering to suggested prevention and rehabilitation activities suggested for them (60). Finally, psychosocial aspects such as social isolation, anxiety or job status of each COPD patient's is assessed in order to initiate relevant support initiatives that can increase quality of life and a well-functional daily life (60).

Pharmacological treatment is a requirement for minimizing COPD symptoms and to prevent exacerbations and the concrete treatment varies depending on disease severity and the acuteness of illness (61). All patients can receive short-term inhaled airway-expanding medicine as needed to relieve symptoms; patients with

moderate, severe and very severe COPD can be treated with long-acting airway expanding medicine potentially combined with steroid inhalants if acute exacerbations are experienced (61). Antibiotics are frequently used during exacerbations to reduce inflammation as are pharmaceuticals used to relieve anxiety (61). Medicine for nicotine replacement is also used to support smoking cessation (61).

Patients should be monitored “regularly” in order to assess disease progression and to support current training and counseling or to initiate new ones (60). This could take the form of regularly assessing COPD severity by spirometry or subjective experiences by MRC or CAT-scoring, but also by having different vital signs indicative of health status checked, such as weight, blood pressure, heart rate and oxygen saturation.

Oxygen therapy are needed when lungs cannot ventilate the blood sufficiently on their own anymore (61). Some patients only require short-term oxygen therapy due to acute exacerbations while they are admitted to hospitals or in the weeks following discharge (61). Other patients might need long-term oxygen therapy (defined as at least 15 hours/day) which is often the case for very severe COPD patients (61). Eventually lungs can be so deteriorated that mechanical ventilation is necessary in order to prevent instant respiratory failure and death.

A patient’s general practitioner (GP) takes the central role in diagnosing, treating and monitoring COPD patients (62). Citizens are not screened for COPD, but according to recommendations (60), smokers, ex-smokers or people employed in high risk jobs over 35 years who have pulmonary symptoms, should have a consultation with their GP (63). After a diagnosis of COPD and the assessment of COPD severity, regular monitoring by the GP is initiated in order to keep the COPD stable, e.g. by preventing exacerbations as much as possible (63). Depending on the certainty of diagnosis, severity of COPD or existence of comorbidities, COPD patients can also be treated and monitored at hospitals (64). Acute hospitals admissions are also normal during exacerbations and mechanical ventilation is usually also initiated at hospitals for COPD patients that are close to respiratory failure. Danish municipalities are, based on a clinical evaluation of a patient’s needs, involved in practical help, home nursing care and pulmonary rehabilitation of COPD patients (63,65). Community care (practical help and nursing care) is provided at regular intervals and training courses and counseling is offered in smoking cessation, exercise planning and diet (63). Community care personnel are not necessarily certified nurses and not specialized in COPD.

This usual practice and typical management strategy for COPD implicate that many healthcare providers are involved in treating and caring for COPD patients and contacts to providers are frequent. In Denmark, COPD patients have recently been estimated to account for as many as 17,000 annual hospital admissions (2% of all

hospital admissions), 65,000 outpatient visits and up to 490,000 visits to general practitioners (12). This management strategy and the organization of Danish health care means that all three sectors have economic incentives for effective and cost-effective management of COPD.

2.2. TELEHEALTHCARE

2.2.1. DOES TELEHEALTHCARE WORK AS A MANAGEMENT STRATEGY FOR COPD?

Based on different syntheses of the available high quality evidence, the short answer to this question was “a hesitant yes” in 2011-2012. Systematic reviews of telehealthcare suggested that telehealthcare could lead to more effective treatment and possible even lower costs, e.g. in the Cochrane review from 2011, that focused on the effectiveness of telehealthcare compared to face-to-face care for COPD patients (26). The review found ten studies who met their inclusion criteria and concluded that telehealthcare was associated with no difference in mortality, higher health-related quality-of-life and a decrease in visits to emergency departments as well as reduced admissions to hospitals (26). This conclusion is largely a confirmation of a previously conducted effectiveness review (66) and in line with similar reviews being published around the same time (67,68). Although costs were secondary outcomes in the conducted reviews, they also concluded that a few studies also demonstrated a potential for telehealthcare to cut costs in the management of COPD (67,68).

But several reviews have also warned that the evidence base for the effectiveness and especially the cost and cost-effectiveness of telehealthcare to COPD patients were weak (28,30,31,67–70). The reviews only found studies with few participants, short follow-up and across studies the participants had different health- and socio-demographic characteristics and they applied different outcomes measures and several classes of telehealthcare solutions. Therefore, any systematic differences across disease type, type of telehealthcare solution or patient characteristics were hard to ascertain and it was therefore difficult to assess what worked for whom (28). Based on these reviews, more evidence of economic effects of larger initiatives in future telehealthcare research were requested that could allow investigators to come a little closer to telehealthcare solutions that worked for particular subgroups of patients (71,72).

2.2.2. DEFINING TELEHEALTHCARE

Tele-related technologies applied in healthcare is not just one solution with the same (proportion of) ingredients. A quick Google search for related terms used to describe the type of communication and networking technology employed in

healthcare (i.e. that uses eHealth, telehealthcare, telehealth, telecare and telemedicine as terms) yields over 20,500 hits and more than 8,500 scientific papers with this exact combination of keywords. There have been a few attempts to review the scientific literature for definitions of some of these terms. For example, Oh and colleagues reviewed definitions of "eHealth" in 2005, and found 51 unique but overlapping definitions. In 2007, Sood and colleagues found 104 peer-reviewed definitions of "telemedicine" (73). So it is safe to say, that the applied terminology used to define and describe different technologies enabling some form of treatment, support, monitoring and care for patients at a distance are partly overlapping, have complex attributes and that no commonly accepted definition exist (73–77).

Instead, attempts to describe frameworks for various definitions of tele-related technologies have been developed in order to compare and contrast different attributes of the applied terminologies. Two of these frameworks are developed by Doughty and colleagues in 2007 (75). The first framework distinguishes between assistive technologies and telehealthcare. Assistive technologies consist of fixed mechanical devices (e.g. walking aids, ramps, grab rails, lifts) and electronic systems (orientation aids, reminder systems, environmental controls) installed at the patients' homes. The focus of assistive technologies is to help citizens conduct daily activities at home without much involvement from institutions (i.e. healthcare or social care providers). Telehealthcare is used as a general term for any telecare, e-care and telemedicine technology that support patient independence but have a stronger emphasis on geographical distance and involvement of institutions involved in protecting and promoting health. Telehealthcare therefore covers technologies that trigger alarms associated with health status that health professionals can respond to, technologies that offers lifestyle and behavior advice, enables proactive contact to healthcare personnel and that transmits various measurements indicative of a person's wellbeing. They could also be technologies enabling health professionals to diagnose or even treat patients remotely from a medical center or hospital. The second framework view assistive technologies only as house adaptations and telemedicine only as hospital services with telecare pretty much as everything else in between (75).

Goodwin (74) defines the differences between terms in another way: *Telehealth* is concerned with the electronic transfer of physiological data between a patient at home and health professionals to assist in diagnosis and monitoring of a disease. *Telecare* is about automatic and remotely monitoring patients to manage risks associated with independent living in their own home. *Telemedicine* is more concerned with online treatment. Regardless of framework chosen, it should be obvious from above that terms that describe tele-technologies employed in the health- and social care sector have slightly different uses and are partly overlapping.

As in a 2012 review of telehealthcare for COPD (78), I prefer the term "*telehealthcare*" as an umbrella term to cover the definitions of telehealth, telecare

and telemedicine described above. This means that "telehealthcare" is not defined in this thesis per se, but are argued to have three important features: First, it is a technology that contains data from a patient – e.g. written responses or statements, measured physiological data, audio or video. Secondly, there is an electronic transfer of this data over a physical distance and thirdly, a healthcare professional exercise their judgment in providing personalized feedback to the patient (79).

Defining telehealthcare by distinguishing between terms can risk implying that particular health professions are delivering healthcare (telemedicine is often associated with curative medicine, i.e. doctors, and telecare with nursing). Since the focus in this thesis is on attaining evidence of cost-effectiveness of the tele-technology and not on professional boundaries, the distinction of terms seems less relevant.

2.2.3. TYPICAL ELEMENTS IN TELEHEALTHCARE SOLUTIONS

Many "modern" telehealthcare solutions combines some of the same ingredients (80,81). The main ingredients are a device and a number of peripheral equipment that can communicate with the device. The device and peripheral equipment enables patients to transmit data indicative of their health status to health professionals who are situated at some other location. Health professionals are then able to provide feedback (e.g. treatment and care advice and/or emotional support) to the patient based on the transmitted data and the patients can receive the feedback at a place where it is convenient for them. These core elements of telehealthcare are often supplemented by a deliberate delegation of monitoring responsibility from one health profession to another (typically from physicians or general practitioners to nurses). Usually, the solution also involves technology- and disease-specific education or training to both health professionals and patients.

But telehealthcare solutions differ in how these ingredients are mixed. Hardware installations are either fixed (telephone or pc) or mobile (laptop, smartphone or tablet) and allow for different data to be transmitted (e.g. verbal communication, questionnaires, weight measurements, blood oxygen levels, blood pressure, pulse, expiratory flow measures etc.) that give a more or less complete picture of the disease that are monitored and allow for different levels of standardized responses to these data (from automated response of judgment to the exercising of judgment in each case). The technological solutions can facilitate data and/or feedback to be delivered either synchronously (in real time) or asynchronously (i.e. store and forward data and response). Patients can also track their own historical data and feedback to a varying degree and the monitoring data are more or less integrated within existing IT-systems applied by health professionals. Finally, the type, amount and intensity of training received by both patients and monitoring staff can vary as can the health professions responsible for monitoring patients.

So, the ingredients of telehealthcare solutions may be the same, but a telehealthcare intervention can be configured in a multitude of ways meaning that several aspects or ingredients may have different effects on the patients receiving it. And these effects may also differ depending on the characteristics of the target population. This means that the effectiveness and costs is likely to vary depending on the characteristics of the technology-configuration and/or the targeted patient population.

2.3. HEALTH ECONOMIC EVIDENCE

The intention with this section is not to describe and review the health economic literature in its technical details, but rather to give an overview of what science in health economics in general mean by "economic evidence" and the role it has in health care decision-making.

Drummond and colleagues have published the most cited and used book on health economic evaluation (82). In it, they define health economic evaluation as "*a comparative analysis of two or more courses of action in terms of both their costs and consequences*" (82, p4). A course of action is a technology in its broadest sense, e.g. a set of activities included in health promotion, diagnosis, treatment or rehabilitation of disease (83). The intended audience for economic evaluations is decision-makers responsible for assigning funding to particular health technologies (84). The role of health economic evaluation in decision-making may be described in the following way:

> *"Whatever the context or specific decision, a common question is posed: are we satisfied that the additional health care resources […] should be spent in this way rather than in some other ways? The other ways these resources could be used might include providing health care for other patients with different conditions, reducing the tax burden of collectively funded health care, or reducing the costs of social or private insurance premiums"*(82, p3).

From the quote, two important features can be elicited. First, health economic evaluation is concerned with *choices* by focusing on the costs and health outcomes of different courses of action, i.e. the inputs and outputs that these technologies would apply and result in, thereby seeking to make criteria for decision-making explicit (82). Second, the theoretical assumption behind health economic evaluation is that of *opportunity costs*: Given a fixed budget, every time money is spent on one course of action, it displaces another course of action that could have been funded instead (82). So, decision-makers are "in principle" better off funding those courses of action that give the highest health-effect for their money (technical efficiency). However, health economic evidence is not the sole source of information used in health care resource allocation, because there can be other concerns than technical

efficiency. In countries that systematically assess new technologies, health economic evaluation is one element in health technology assessments (HTAs). An HTA is a systematic and multidisciplinary evaluation of health effects, economic effects, organizational impacts and relevant social or ethical issues (83). This means that a health technology can end up being prioritized/not prioritized for other reasons than technical efficiency (e.g. equity issues).

There are different approaches to conducting health economic evaluation which includes decision modeling; trial-based economic evaluation and observational studies. Decision modeling are usually the preferred approach for economic evaluation used in health care resource allocation, e.g. as part of HTAs, since economic modeling provides a more comprehensive framework for decisions under uncertainty (82). Trial-based economic evaluation and observational studies is not recommended as a single data source for economic evaluation used to make health care decisions, but results from trial-based economic evaluations are central as input to health economic decision models by "feeding" these with values on resource use, unit costs and health effects (85). In this process, observational studies can also be used (85).

Decision modeling for economic evaluation

Economic evaluation conducted as decision models incorporates secondary data retrospectively from multiple published sources (85). The strengths of decision models are that they provide a structure to the decision problem by reflecting the costs and outcomes in clinical pathways or health states that patients might undergo and how different clinical processes or interventions may influence these pathways or health states (82,85). Decision models can also incorporate all relevant evidence into the economic analyses and compare all relevant options for courses of action (82,85). Decision models translates this evidence into estimates of total expected costs and outcomes of alternative courses of action thereby identifying which alternative is the best -with the available evidence- via a pre-specified decision rule such as a willingness-to-pay threshold for additional quality-adjusted life years (QALY) (82,85).

Decision models also allow for reflection of how uncertainty in the underlying evidence translates into uncertainty in health care decision-making (82,85). Through this assessment of uncertainty, it can point to the value of more information from future research, i.e. instead of being forced to accept or reject a course of action now, the decision can be postponed until further evidence has been generated on model input parameters or the clinical course of the disease (86).

Technically, a decision model uses statistical probabilities to relate a range of plausible consequences that are likely results of alternative courses of action under study (85,87). Each consequence has a cost and an outcome. Based on input to the model, the likelihood of each consequence is presented with transition-probabilities

(85,87). For each given course of action, it is possible to calculate the expected total costs and outcome by “rolling back” the model or by “sending” a cohort of hypothetical patients through the pathways (calculating the total weighted expected costs and outcomes) (85).

Gathering evidence for decision modeling can be a challenging task. Single studies rarely compare all relevant options and there is therefore a need to build a decision model from many sources (82). Furthermore, a proportion of these sources may be generated outside the decision-jurisdiction (e.g. another country with a different health care system) (88). This makes some form of evidence synthesis and/or translation of results necessary first in order to arrive at input values for a decision model (82,88). In addition, the outcome measures and time horizon applied in clinical studies of treatment options will limit their usage for economic evaluation, since evidence on economic end-outcomes are usually missing (i.e. end-outcomes such as resource use, costs, health-related quality of life) and the follow-up period may be too short to capture all relevant costs and outcomes relevant for decision-making (85). Ways to link immediate outcomes with final outcomes and extrapolation of costs and outcomes may therefore be necessary (85). These assumptions and steps taken in order to apply a model to a specific decision may seem speculative, but to use a famous quote from the statistician George Box on the use of empirical models: "*Essentially, all models are wrong, but some are useful*" (89).

Trial-based economic evaluation

In health research, experimental evaluation or randomized evaluation are considered the “gold standard” for evaluating the effects of health technologies (90). The most used variant, the randomized controlled trial (RCT), randomly assigns individual patients to one or more treatment alternatives and a control group in which patients does not receive treatment (90). Patients are thereafter followed in parallel using the same methods in order to compare outcomes between them over a certain time period (90). The randomization should maximize the likelihood that patients in the alternatives are comparable in every way except the treatment under consideration, which leaves only the different treatments as explanations for observed differences in outcomes (90). To achieve credibility in the cause-and-effect conclusions drawn (a high internal validity), this design requires that the investigators follow a number of commonly accepted procedures in order to avoid biasing or confounding the results (91).

Trial-based economic evaluation or “economic evaluation alongside clinical trials” is designed to collect primary data prospectively from individual patients within these randomized trials (92). The focus is on estimating incremental costs and incremental health outcomes of health technologies and the uncertainty surrounding these estimates within this single study (82,92). The treatment options compared and time horizon for the economic analysis is determined by the design of the

randomized trial (93). Data on relevant parameters used in economic evaluation (e.g. resource use, mortality, health-related quality of life) are collected more or less simultaneously with the clinical parameters from all individual patients or a subset of patients within the trial (93) and an incremental analysis of the costs per outcome measure is conducted (e.g. *additional* costs per QALYs *gained*) (93).

The advantages of trial-based economic evaluation is the opportunity to exploit the trial's high internal validity (82). The marginal costs of collecting additional data in the trial is also relatively low and it can be time-saving compared to collecting economic evidence retrospectively (82). The downsides of using trial-based economic evaluation for decision-making - in addition to the challenges mentioned in the section on decision-modeling (inappropriate comparisons, outcomes and time horizon) - is generally problems with generalizability (external validity) (82).

Trial-based economic evaluation mainly embedded in trials designed for clinical purposes (called "piggy-backing studies") (94) are often explanatory evaluations concerned with estimating the efficacy of courses of action, i.e. estimate the outcomes and costs of technologies under ideal circumstances (82,92). These studies often have a very high internal validity, but the implications could be that effects are overestimated. Patients could be selected based on a set of strict criteria leading to a sample of patients that may not have the same prognostic characteristics as "normal patients" would have in routine practice (e.g. patients selected in the trial may only be those with fewer comorbidities, that are the most motivated for treatment etc.). Furthermore, the activities undertaken in efficacy trials may not be representative of the activities that are feasible in routine practice, because activities in trials are more frequent or invasive, standardized, detailed or health professionals used in the trial may be better trained. All of these problems can make it unlikely to replicate the costs and health outcomes in routine practice thereby reducing the potential for generalization (82).

A clinical trial can also be more specifically designed to reflect routine practice and serve as a better carrier of economic evaluation (82,95). These studies seek to estimate effects of courses of action in a setting closer to routine practice and are often called "effectiveness trials" or "pragmatic trials". The idea is to balance the objectives of a high internal validity with larger opportunities for generalizing the results (93). This is done by maintaining the randomization to different courses of action while imposing fewer restrictions on the included patients and how they were followed in the duration of the trial. This would often lower the internal validity of the conclusions drawn, but increase the generalizability of the results. Pragmatic designs are actually the current recommended approach when conducting trial-based economic evaluation (93).

Efficacy trials and pragmatic trials are often viewed as extremes or end points on an internal-external validity continuum, where efficacy trials have an explanatory

purpose used in clinical practice (i.e. a high internal validity is therefore most important) and pragmatic trials are used in order to make implementation or funding decisions (i.e. generalizability is important) (82).

Observational studies

An economic evaluation could also be conducted alongside an observational design such as case-control designs, cohorts or cross-section study (82). In observational studies, patients receive treatment based on routine decisions made by healthcare professionals and no or very few restrictions are made on e.g. selecting patients or the activities conducted (82). These studies are essentially real world studies investigating "what happens" (or "what happened") and many applies registers either prospectively or retrospectively.

Observational studies can be used when experimental evaluation is unethical (e.g. randomization to abort/not abort is unethical!), impractical (e.g. events occur very rarely) or when the evaluators cannot influence how courses of action can be allocated for some reason (e.g. lacking the managerial influence to impose experiments) (96). Like experimental evaluation, observational studies can compare different alternative courses of action (e.g. in case-control designs). Or it can simply follow a group of participants without having a control group (single cohort designs) (96).

The main disadvantage of observational studies is the lack of randomization, which make it more likely that estimated treatment-effects could be biased or confounded (96). At a minimum, statistical adjustment for systematic differences in prognostic factors relevant for the treatment-effect must be applied (97); however, this is only feasible for observable differences. What randomization ensures is that any observable *and* unobservable imbalances between patients in alternative courses of action only exist by chance (96), so this design will have a much higher internal validity. If no control group exist at all (e.g. in a single cohort with before/after evaluation of costs and outcomes), it is generally considered very difficult to isolate a treatment-effect from what would have happened anyway (96).

2.3.1. CLASSIFICATION OF HEALTH ECONOMIC EVALUATIONS

Several research designs of economic evaluations could be applied in order to evaluate the clinical outcomes and/or costs of healthcare interventions (82). Drummond and colleagues distinguishes the characteristics of evaluations used in health care decision-making along two dimensions: Firstly, whether the evaluation includes a comparison of one or more alternative courses of action and secondly, whether the evaluation directly relates costs with a clinical outcome (82). Cost descriptions or cost analyses such as cost-of-illness studies (98) and traditional cash flow forecasting used in financial management or investment theory (99) are partial evaluations, since they either do not include a comparator and/or does not directly

relate costs with a clinical outcome (82). Likewise, clinical trials without economic analyses are also partial evaluations, since they do not include the costs of technologies or relate costs with health outcomes. Full economic evaluation includes both a comparison of one or more alternative courses of action and directly relates costs with a clinical outcome (82,100).

There are three basic forms of full economic evaluation, which differs primarily on how health-effects are defined and measured (82). In a *cost-effectiveness analysis*, health-effects are measured in natural units, such as years of life gained, number of averted incidents or improvement in a pain score (84). The health economic benefit of a technology from a cost-effectiveness analysis is therefore expressed as costs per unit of effect, e.g. costs per avoided death (82). A limitation of cost-effectiveness analyses is that they cannot be used to compare health-effects across health technologies if outcome measures for assessing the effectiveness of alternative technologies are different, e.g. orthopedic surgery in a knee is usually not to avoid death (82). So, cost-effectiveness analyses are most useful for prioritization within disease groups or within homogenous healthcare programs. The *cost-utility analysis*, in contrast, uses a composite outcome measure for health-effect in order to compare the effects of technologies across disease groups or healthcare programs. Usually, health-effect is measured as QALYs gained or some variation of it (e.g. disability-adjusted life years) which is a way of adjusting life expectancy for quality of life during the years lived. Results from cost-utility analyses are presented as costs per QALYs gained (82). Cost-utility, therefore, have a broader appeal and are often viewed as more useful for health care decision-makers. *Cost-benefit analyses* values health-effects in monetary units and can therefore be used to prioritize technologies across sectors of an economy such as between healthcare, transportation and education (82). Cost-benefit analyses also allow for the valuation of negative or positive externalities of a health technology (82). However, cost-benefit analysis is rarely used due to the absence of market prices.

2.3.2. IMPORTANT DESIGN FEATURES OF HEALTH ECONOMIC EVALUATIONS

Several guidelines for designing, conducting and reporting health economic evaluations have been published and can be general recommendations in text books, e.g. Drummond and colleagues (82) or general guidelines from societies for health economic evaluation (101). There are also specific additional requirements depending on the adopted approach, i.e. for decision modeling in health economic evaluation (102), trial-based economic evaluation (93) or economic evaluation in observational studies (103,104). Looking across these recommendations and guidelines several important aspects are important in sound health economic evaluations.

It is important to describe a well-defined research question that defines the target population, the health technologies to be assessed and the objectives of the economic evaluation (82,93,101–103). A perspective for the analysis must also be described and justified (82,93,101–103). The perspective is the viewpoint from which the relevant costs and health-effects are collected (82). The perspective can be comprehensive (e.g. a societal perspective or the perspective of the healthcare and social sector) or more narrowly scoped (e.g. a specific healthcare provider) (82). Choice of courses of action that are compared must be described and justified (82,93,101–103). Comparators are the health technology under scrutiny and one or more "usual practices" or most effective, least costly or frequent alternative health technologies relevant for treatment and care for the target population (82). For transparency and to address the potential for generalization, a detailed description of all alternatives should be made (82,93,101–103).

If mortality and health-related quality of life are important outcomes, QALYs should be used as outcome (82,93,101–103) and validated instruments should be applied (82,93,101–103). Relevant costs categories should reflect the chosen perspective as well as measured and valued appropriately (82,93,101–103). The first step is to identify resource consumption that can vary between alternative health technologies (82). Secondly, the most credible sources for collecting resource consumption must be chosen and combined, e.g. from case record forms, accounts, cost diaries, questionnaires, registers, time studies, other studies, expert opinions etc. (82). The third step is valuation of the resource consumption (82). Ideally, the evaluation should be made based on opportunity costs or market values. These would often not exist, so unit valuation is often based on tariffs, average salaries or remuneration. From information on resource consumption and unit costs, the total costs for each health technology can be calculated (82).

An appropriate time horizon for the analysis should be chosen (82,93,101–103). The choice should ensure that all relevant differences in health effect and costs are observable within the evaluation period (82). Health effects and costs occurring throughout this period should be discounted in order to reflect a time preference assumption and to make interventions with different flows of outcomes and costs directly comparable (82,93,101–103).

An incremental analysis must also be presented, since it is important to examine the additional costs and health outcomes that a technology imposes over other technologies (82,93,101–103). In health economic evaluation, this is often reflected in the incremental cost-effectiveness ratio (ICER), which is the difference in total costs between two technologies divided by the difference in health outcome between the same two technologies (82).

Inevitably, any evaluation will contain some form of uncertainty relating to the data or the assumption in the evaluation (105). The stochastic uncertainty reflects the

variability in costs and health-outcomes demonstrated for seemingly identical patients (105). Heterogeneity is concerned with the observed differences in health-effects and costs between patients that can be explained by observable characteristics of those patients (105). Parameter uncertainty reflects assumptions made about valuation of specific parameters in the evaluation and structural uncertainty is concerned with the analytical choices made (in- and exclusion of variables or analysis model) (105).

The features included in health economic evaluation are therefore manifold, which leaves room for many choices of detailed design. Even though recommendations for transparent reporting of results from economic evaluations exist, it can still be challenging to compare economic evaluations. As a result, some would argue for a "reference case" or a preferred choice in the design of an economic evaluation based on the features mentioned above in order to make comparisons more reliable and transparent (106). According to Gold and colleagues, the societal perspective should be chosen, costs should include both cost of health care, patient- and caregiver time and productivity loss and a 3% discount rate should be used for discounting costs (106). Effectiveness estimates should incorporate both benefits and harms, mortality and morbidity should be combined in QALYs and existing practice should be a control group (106). Finally, sensitivity analyses should be made (at least one-way analyses) and the ICER should be compared with other relevant interventions (106).

CHAPTER 3. METHODOLOGY

3.1. A DESCRIPTION OF THE TELECARE NORTH INITIATIVE

3.1.1. INTERVENTION

The chosen telehealthcare intervention was supposed to complement or be offered in addition to usual practice for monitoring and caring for COPD patients (107). The intervention was designed independently of the evaluators.

The technical part of the telehealthcare solution was a package or "Telekit" (see Figure 2) consisting of a standard Samsung Galaxy tablet with an inserted SIM card and peripherals capable of measuring and transmitting blood pressure, pulse, blood oxygen saturation, and weight. The package also contained a tablet-pen and a user manual (108). Each TeleKit was technically prepared in advance. During installation at the patient's home, the Global System for Mobile communications (GSM) coverage was checked and in the event of poor coverage, it was possible to switch to a local wireless network if this was installed in the home (108).

Figure 2: A Telekit that consists of a tablet, tablet pen, pulse oximeter, blood pressure monitor and a scale. Source: (108)*.*

The tablet had two applications (apps) installed: First, a measuring and monitoring application containing health-related questions that also could collect and transmit the measurements made by the peripherals. From this application, the patient could also create an overview over historical measurement values and gain access to a message function that allowed for an asynchronously dialogue with a healthcare professional. Second, a training-application containing a digital version of the manual, several videos showing how the various parts should be used and a training video targeted the patient (108). The device was portable but could only be used inside Denmark (108). The provided tablet could be updated and managed

centrally. Should any of the electronic items in the Telekit display errors, the entire TeleKit was usually replaced (108).

On the receiving end of the measurements conducted via the tablet, health professionals from municipalities (usually nurses) could view and evaluate the submitted measurements in a web interface (OpenTele). They could only access data on patients who were associated with the municipality district they were employed in (108). Conducted measurements and answers to disease-specific questions were automatically color-coded based on threshold values made in advance by the general practitioner for individual patients (108). To facilitate COPD patients taking charge over their disease and allow for municipal evaluation of monitoring data both groups received disease specific training and training in using the device (108).

The telehealthcare solution offered built upon a previous telehealthcare solution in North Denmark Region (109), which required several changes. The new technical platform required different software solutions to be developed (e.g. the two applications mentioned above) and the measurements were stored and integrated with existing information technology (IT) systems throughout the healthcare sector (i.e. electronic patient record at the hospitals, care journals in municipalities and medical practices throughout the region) to allow for data exchange (108). A change of monitoring responsibility from physicians at hospitals and GPs to health professionals (primarily nurses) in municipalities was included in TeleCare North. This entailed the development of organizational models for monitoring patients and implementation of new transitions procedures of patients between sectors (110). Particularly health professionals in municipalities would find new types of tasks related to overseeing and evaluating monitoring data. In addition, new tasks related to the support and instruction in the use of the telehealthcare equipment in the patient's home was expected. New tasks for GPs included the formulation, revision and reporting of threshold values for various physical indicators of COPD. GPs would still be the patients' primary contact and portal to the healthcare system. Health professionals at hospitals would be expected to receive fewer patients for admission and outpatient controls thereby reducing their involvement in managing COPD patients (110). The COPD patient would be expected to take on an active role in managing their own disease and thus responsible for conducting daily measurements of physical indicators at home according to a predetermined agreement (110).

The intended target group originally included all COPD patients in North Denmark Region "who could potentially benefit from the telehealthcare" (107), which meant that patients with all COPD severities could potentially be included in the implementation. This also meant that several thousand patients could potentially be included. Later on the inclusion criteria was narrowed down a bit to the description made in the trial protocol (111).

3.1.2. ORGANIZATION OF TELECARE NORTH

There is no question that TeleCare North was an important initiative for the involved stakeholders. It was developed in the geographical area of North Denmark Region with the participation of all sectors involved in the management of COPD patients. There was a high level of political and management attention and the organization surrounding the initiative was large.

TeleCare North was led by a steering committee consisting of top officials or senior members from all of the involved organizations (e.g. the chairwoman of the steering committee was the CEO of North Denmark Region and the vice-chairwoman was a CEO from Aalborg Municipality, the largest municipality in the region) (107). The connection to the national action plan meant that the group of involved stakeholders quickly expanded by members of the Danish government (i.e. the Danish Agency for Digitalization) and representatives from the Danish Lung Association (107). An executive committee was also formed with members from the region, municipalities and general practitioners (107). These two committees had three "reference groups" that followed the initiative (107). This meant that the project had the attention and support of all levels of North Denmark Region, all 11 municipality councils and the local branch of the Danish Medical Association (107).

The task of completing the development and implementation of the telehealthcare intervention was given to a newly formed region-based secretariat solely dedicated to the initiative (107). This secretariat reported directly to the steering committee and the CEO of North Denmark Region (107). The secretariat became responsible for revising the IT-solution, revising how data was stored and evaluated, changing the responsibility for monitoring and the frequency of monitoring COPD patients and adjusting the content of work for particularly healthcare professionals in municipalities (108). The secretariat split these tasks into four main "tracks" each with its own project organization (108): An *IT track* focused on the technical elements and integration and coordination with national stakeholders who focused on the future telehealthcare technical infrastructure. An *organization track* focused on establishing coherent cross-sectorial work procedures and managed potential legal challenges. A *clinical track* focused on clarification of roles in delivering healthcare services, the content of the clinical work and patient pathways. Finally, an *implementation track* coordinated and supported the local implementation across all municipalities and hospitals (107).

The responsibility for evaluating the TeleCare North initiative was divided between three professors from three different departments at Aalborg University and they all had a seat in the steering committee, where they decided on the overall research design for the evaluation. The task of designing a detailed evaluation and for completing those evaluations was then given to three more or less separate Ph.D.-stipends under the supervision of the professors. The general purpose of evaluation

was to generate quantitative and qualitative evidence of the effects of the initiative and for future implementation of telehealthcare in Denmark in accordance with the national action plan (112). A Ph.D.-student from the Department of Health Science and Technology focused on "patient-related" outcomes such as health-related quality of life, mortality, lung function, blood pressure and pulse. (112). A Ph.D.-student from the Department of Sociology and Social Work concentrated on the management- and organizational implications in preparing for and operating a large-scale implementation of telehealthcare (112). Finally, a Ph.D.-student from the Department of Business and Management (the author of this PhD thesis), were given the responsibility of conducting a health economic evaluation of the TeleCare North initiative (112).

3.1.3. THE DANISH "NATIONAL ACTION PLAN FOR DISSEMINATION OF TELEMEDICINE"

TeleCare North had a connection to a national action plan meant to ensure an evidence-based foundation for a later decision on whether or not to implement certain telehealthcare initiatives to particular patient groups in Denmark (107). This decision took place in late 2015 (113).

In June 2012, the Danish government had published "The national action plan for dissemination of telemedicine" (33), which was also agreed upon by "Local Government Denmark" and "Danish Regions", interest organizations for the 98 Danish municipalities and the five Danish regions, respectively. The action plan was meant to ensure that new and more ambitious telehealthcare initiatives were tested and that already effective telehealthcare initiatives would be applied to a greater extent (33). The action plan also entailed conducting evaluations, which was important in order to facilitate a later national decision of which telehealthcare solutions should be disseminated and to which patient-groups (33). The action plan was complemented by £8 million dedicated to co-fund five telehealthcare projects across the country (33). The five telehealthcare initiatives were selected based on the scale of the initiative, i.e. the number of patients or number of regions/municipalities included, a positive economic business case, that the initiatives supported integrative patient pathways and that the initiatives had strong local support (33). Two initiatives were small demonstration projects on previously uninvestigated disease areas (mental illness) and one initiative was a further dissemination of an already effective telehealthcare solution (for diabetes) (33).

Most relevant for this thesis, the action plan also consisted of two large-scale implementation projects with evaluations of the effects of these implementations – one of which was TeleCare North. The two initiatives should in general seek to coordinate and establish a generic IT infrastructure that supported home monitoring and video consultations nationally that could also be opened up to other patient-groups in the longer run.

The total budget for implementing TeleCare North was roughly £5 million and in addition to the funds made available in "the national action plan for dissemination of telemedicine" also included funds from the Obel Family Foundation, the European Social Fund as well as local partners in North Denmark Region (107). The quantifiable objectives of TeleCare North was to increase the health-related quality of life of COPD patients thereby increasing their quality-adjusted life years (QALYs), to reduce the number of hospital contacts and the amount of time spent in municipalities on home care (33,114). Furthermore, several "qualitative" effects were expected, e.g. better coordinated patient pathways and increasing self-management skills for patients (33,114).

The second large-scale initiative part of the action plan was "*Clinically Integrated Home Monitoring*"; Danish acronym "*KIH*") (33). It was implemented in two Danish regions (the Capital Region and Central Denmark Region) and would use a slightly different technical solution and organizational setup to monitor an estimated 2,000 patients (actual inclusion app. 1,200 patients (115)) from five different patient groups (diabetes, COPD, pregnant women with or without complications and patients with gastro-intestinal inflammation) in their own home (33). The objectives were across initiatives in KIH to reduce the number and duration of hospital admissions and outpatient control visits. In addition, a planned task-shifting from physicians to nurses should free up specialized resources. Finally, a number of "qualitative" effects were expected (e.g. correct and consistent treatment across sectors while strengthening patients' skills and empowerment) (33). The total budget of KIH was £6.6 million (116).

The five initiatives in KIH consisted of different evaluation designs; two were based on randomized trials and three were cohort designs. The economic analysis of resource consumption in the health care and social sector for all initiatives were based on before/after time studies of health professions involved in outpatient visits, admissions and municipality based social services *or* their professional estimates of time consumption (115). This time consumption was valued based on annual average salaries for the involved professions (108). The sub-project focusing on telehealthcare for COPD patients that are most relevant in comparison with TeleCare North were one of the two randomized trial with 281 participants (140 received telehealthcare) with severe COPD (108). This study focused specifically on time consumption for telehealthcare, hospital admissions, outpatient visits and patient transport time for hospital visits. Time consumption for personal care at home and home nursing care were also included, but only as a sample over a three to four week period (108).

3.2. RESEARCH DESIGN FOR RQ1: THE EFFECTIVENESS AND COST-EFFECTIVENESS OF TELECARE NORTH

The research design is described in the trial protocol (article 2 in appendix 2), so this paragraph therefore consists of a summary of this article and the conclusions from the systematic review (article 1 in appendix 1) that lead to the design of particularly the trial-based economic evaluation of TeleCare North.

3.2.1. SUMMARY OF THE SYSTEMATIC REVIEW (ARTICLE 1)

Title: A systematic review of the cost and cost-effectiveness of telehealthcare to patients suffering from chronic obstructive pulmonary disease (117).

Objectives: To conduct a systematic review of the evidence on the costs and cost-effectiveness of telehealthcare for patients with COPD. This could be used to assess the number and quality of the current evidence of costs and cost-effectiveness of telehealthcare and to inform the design of future evaluations.

Methods: A systematic literature search was conducted within 8 scientific databases complemented by a search for grey literature in Google Scholar. Studies should include primary data on cost or cost-effectiveness specifically for COPD patients, have applied technologies similar to the definition of the latest Cochrane review on effectiveness of telehealthcare to COPD patients, and only one study from the same trial could be included. No exclusion was made based on language and publication date. Studies fulfilling these inclusion and exclusion criteria were then critically appraised by applying the Consensus Health Economic Criteria list for assessing the methodological quality of economic evaluations.

Results: Only six studies were identified with the inclusion- and exclusion criteria used in the review. These studies were conducted in Europe and North America and involved a total of 559 COPD patients whereof 281 had received telehealthcare. Although all included studies demonstrate a potential for cost savings for primarily patients with severe and very severe COPD, only one of these studies is a cost-effectiveness study whereas the other five are cost studies only. The telehealthcare interventions are also very heterogeneous and the activities in their comparator "usual care" are difficult to identify. Furthermore, the studies have relatively narrow cost-scopes or perspectives which mean that the identified cost savings in hospitals are at risk of merely being transferred to other healthcare delivery sectors (e.g. general practitioners or communities). Finally, the studies were efficacy trials that included a small number of patients and have relatively short follow-up period. The quality of the evidence was poor. Five of these studies were deemed as having "low quality" and one study was assessed as "moderate quality" in its economic evaluation design.

Discussion: The review demonstrated a potential for cost savings mainly to hospitals and/or the healthcare sector when telehealthcare is applied to COPD patients. However, we argued that decision-makers should be skeptical of the current evidence of cost-effectiveness in telehealthcare research. Several improvements to future research were deemed necessary, e.g. more cost-utility analyses that relate costs to QALYs should be conducted; these studies should also have broader cost-scopes and longer follow-up periods and finally larger studies should be conducted to allow for heterogeneity analyses.

3.2.2. SUMMARY OF THE TRIAL PROTOCOL (ARTICLE 2)

Title: Effectiveness and cost-effectiveness of telehealthcare for chronic obstructive pulmonary disease: study protocol for a cluster randomized controlled trial (111).

Building on the call for more research in previously conducted systematic reviews of telehealthcare for COPD patients and the systematic review from article 1, two departments collaborated on designing a clinical trial together in order to estimate the effectiveness and cost-effectiveness of the TeleCare North initiative.

Objectives: To describe the design of a pragmatic cluster-randomized trial with nested economic evaluation. The study is meant to assess the effectiveness and the cost-effectiveness of the Danish TeleCare North telehealthcare solution for patients suffering from COPD compared to usual practice implemented throughout North Denmark Region in 2014-2015.

Methods: North Denmark Region and 10 out of 11 municipalities participated in the trial and all general practitioners were invited to participate. Spirometry-diagnosed COPD patients were eligible for inclusion if they received treatment according to GOLD-guidelines, if COPD was a primary disease and if they met a set of additional in- and exclusion criteria (MRC and CAT score, residence, language and GSM coverage). Patients were recruited by general practitioners. 26 municipality districts (13 pairs) defined randomization units to reduce the risk of contamination. Assessment was divided into an effectiveness theme and a cost-effectiveness theme. The primary effectiveness outcome was differences in health-related quality of life measured by Short-Form 36 (SF-36) from baseline to 12 months follow-up. Secondary effectiveness outcomes were changes in mortality and physiological parameters from baseline to the follow-up at 12 months. The cost-effectiveness outcome was the incremental cost-effectiveness ratio measured as the cost per QALY gained from baseline to follow-up at 12 months.

Discussion: The study was designed to answer the call for more comparative research with nested economic evaluation of large-scale implementation of telehealthcare to COPD patients. It does so by including an estimated 1,200 patients and has a relatively broad cost-scope. The duration of 12 months might be too short to capture all relevant evidence of costs and effects; however, for practical and

ethical reasons, the involved stakeholders decided that all eligible patients should be able to option for the telehealthcare intervention no later than 12 months after randomization.

3.3. RESEARCH DESIGN FOR RQ2: APPLICATION OF EVIDENCE IN THE NATIONAL DECISION

The research design for RQ2 is rooted in the qualitative research paradigm, which could be defined as "*any kind of research that produces findings not arrived at by means of statistical procedures or other means of quantification*" (118, p17). The paradigm seeks to shed light on, understand or to create theoretical or methodological lessons for future reference in similar situations, which is very different from quantitative research aimed at prediction and generalization of results (119). Qualitative research is an investigative process where a researcher seeks to make sense of a phenomenon (whether cognitive or social) (120) and has strong foundations in fields such as sociology and political science (120).

Qualitative research has several important characteristics (119): it is conducted in a "natural setting" outside the control of the researcher (e.g. where actual events take place) (120). The researcher is the primary instrument for interpreting data in a process of selecting, examining and analyzing the applied data (119,120). Usually, several data sources and -methods are applied such as documents, observations or interviews (119,120). Results are reported as descriptive data in quotations or pictures (119,120). Furthermore, qualitative research focuses on the events and outcomes as the occurred (119,120). Finally, it is important that the researcher reflects on the role he or she has in the interpretation of the data and conclusions drawn (119,120).

This research design is meant as a carrier to allow for investigation of how a national decision on telehealthcare unfolded. As an inductive research paradigm, qualitative research could also facilitate theory-building that could be incorporated in future research on cost-effectiveness methods for telehealthcare. As such, this investigation seeks to illuminate how a national resource allocation unfolded and seek to fertilize the ground for an analytical generalization of the lessons learned that could be used to increase the usefulness of future evaluation of the economic effects of telehealthcare.

3.3.1. DATA COLLECTION STRATEGY

Research question two was defined *ex post* as part of an after-reflection of the national decision-making context of telehealthcare in Denmark, i.e. after the primary data collection for TeleCare North had ended, data for the trial were analyzed, results presented and the Ph.D thesis initiated. Therefore, feasible

research designs to choose from in answering this question was limited. The choice fell on document analysis, which according to Bowen can be defined as "*a procedure for reviewing or evaluating documents*" (121, p27), such as agendas, letters, press releases, institutional reports, event programs, web pages, newspaper clippings and –articles (121).

Document analysis can also be used as a "stand-alone" method for specific purposes (121). Examples includes tracking changes and development in the formulation of reports or bills or to analyze changes in descriptions of initiatives or projects over time (121) as well as the examination of document usage by certain people or organizations (121). Document analysis can also be used to generate new questions that needs to be addressed in research (121). Document analysis is generally thought of as an efficient research method for these purposes, since it involves data selection and not data collection and can be a preferred or necessary choice when collection of new data is unfeasible (121).

3.3.2. BOUNDING THE STUDY

Selection of documents is not based on random sampling; rather documents are *purposefully selected* in order to help the researcher understand the problem (119). This selection is based on four criteria (122): 1) *the setting*, i.e. where research could take place 2) *the actors*, i.e. from whom data could be collected, 3) *the events*, i.e. what pre-planned situations data could be collected from 4) and *the process*, i.e. how events actually unfolded (120).

The chosen setting is publically available documents from researchers and institutions involved in creating evidence on the effectiveness and cost-effectiveness of telehealthcare solutions (e.g. research articles and reports published within the national action plan). In addition, publically available documents from decision-stakeholders participating in the national decision or debate on whether or not to disseminate telehealthcare solutions to certain patients based on the national action plan (e.g. web pages and newspaper clippings).

Relevant actors are the Danish Ministry of Health, the Danish Regions and Local Government Denmark, the Danish Agency for Digitalization, PA Consulting Group and researchers from Denmark affiliated with TeleCare North and KIH as well as international researchers generating new evidence since the formulation of the action plan in 2012. The Danish Ministry of Health, Danish Regions and Local Government Denmark was responsible for agreeing to create and fund a national action plan and for making the decision to disseminate telehealthcare to certain patient groups. A national steering committee with representatives from the five Danish regions, 98 municipalities, the Ministry of Health and the Danish Agency for Digitalisation has since been responsible for implementing and communicating this national dissemination. Researchers have since 2012 and up to the decision in

autumn 2015 published evidence on the effectiveness, costs and cost-effectiveness of telehealthcare solutions both outside and within the national action plan especially two widely cited international studies (Telescot in Scotland and Whole System Demonstrator in England). PA Consulting Group was responsible for conducting an initial and revised national business case based on available evidence. Prominent health economist and other researchers within telehealthcare took part in the debate as it unfolded.

The "events" include the content of the Danish "national action plan for the dissemination of telemedicine" including the data and data-sources it applied. But also the new evidence that was generated internationally and nationally from January 2013 up to the national decision that could potentially be part of the evidence used in the national decision in autumn 2015. The events also include the actual decision taken and the premises for this decision. This is complemented by the process as it unfolded, particularly a debate that occurred in the media in April and May 2016 on business cases and economic methodology used in the decision.

For this research question, the following central documents are therefore selected (Table 3).

Table 3: Overview of documents selected for analysis of evidence application in national decision. Source: Own contribution.

Document	**Rationale**
"The national action plan for the dissemination of telemedicine" (33)	Contains the initial expectations from the Danish government, Danish regions and municipalities for the effects and costs of telehealthcare and the types of evidence used nationally. But it also gives a sense on how evidence synthesis methodologically was created nationally.
Initial business case for TeleCare North (114)	Contains the initial expectations from North Denmark Region and are used to formulate the "national action for dissemination of telemedicine"
International scientific studies on telehealthcare for COPD patients published after the initiation of the action plan, especially • Whole system demonstrator project (WSD) in England (123–126) • Telescot in Scotland (127,128)	Important new evidence and methodology that could in principle be incorporated into the decision
Empirical results from initiatives funded by the "The national action plan for the dissemination of telemedicine" • Systematic review made by CIMT	Contains new evidence directly associated with the decision to implement telehealthcare. It also includes an assessment of the quality of existing evidence that could have provided background to

(25) • Final report from "Clinically Integrated home monitoring (KIH)" (115) • Final report from TeleCare North (108) • Research articles from TeleCare North (117,129–131)	the formulation of the action plan.
National business cases on the economic effects of disseminating telehealthcare in Denmark • Initial business case (132)	Important analyses to assess the expected impact the different sectors involved in managing COPD had in Denmark over time. Results from the business cases would have a huge impact on the Government, regions' and municipalities' desire to implement the telehealthcare solution.
Web pages of the Danish Agency for Digitalization (113,133–138)	Communication platform for the dissemination of telehealthcare in Denmark
Altinget.dk panel debate on the initiatives in "The national action plan for the dissemination of telemedicine" (139–157)	Became an important source for assessing the relevance, balance and wording of different stakeholders on the results from the action plan as events unfolded. Also source of input to how stakeholders perceived the research designs used.
Magazine clippings on the national decision (158,159)	Became an important source for assessing the relevance, balance and wording of different stakeholders on the results from the action plan as events unfolded. Also source of input to how stakeholders perceived the research designs used.

3.3.3. PROCEDURE FOR THE ANALYSIS

As with most qualitative methods, the analytical strategy in document analysis is an iterative process of reading, categorization topics, formulating themes and interpretation (119,120). There are several analytical steps to undertake and they could look slightly different depending on the employed text-book, but the general procedure is this (119,120): First, the data for analysis are organized and prepared, e.g. by printing or transcribing documents and then sorted depending on the type and source of the document (119–121). Second, all documents are read and for each document, the researcher elicit the overall meaning or message of the document (119–121). Third, for each document, the researcher start coding the text (119–121). Fourth, this coding process are used to synthesize categories into a smaller and more abstract set of themes that could later serve as the main findings from the analysis (119–121). The themes should look across documents and be supported by multiple sources (119–121). Fifth, the researcher applies his or her personal interpretation from an understanding that is brought to the study from past

experiences or the literature (119–121). These steps are iterative in the sense that data will be continuously reviewed and repeatedly coded until "data saturation", where no more major themes can be elicited from the texts by the researcher. Finally, the interpretations made from the analysis should undergo a process of verification, ideally based on several strategies, such as triangulation of methods, a hearing process with the involved decision-stakeholders or a prolonged data collection with repeated analyses, participating in and checking all steps in the analysis process and by reflecting over and stating the role of the researcher (119–121).

The documents selected for this analysis were printed out on paper and organized into the groupings made in Table 3 and the researcher's initial thoughts on the meaning of each document were noted on the top of document. The documents in each grouping were arranged in piles. Then hand-coding commenced for each document by highlighting certain passages and writing a representative word or category for this passage in the margin. When all documents for each grouping were read, major themes were created on post-it notes and attached to their respective piles. More than one theme could be created per pile. All of this was combined by the researcher's personal experience in the past, the theoretical literature on resource allocation in health policy and employed methodology used within health economic science. Finally, the researcher went back to each document to look for additional categories or themes inspired by the generated themes within each grouping of documents as well as the literature and personal experience.

3.3.4. RESEARCHER'S ROLE

To allow the reader to evaluate the credibility in the accounts made, a verification is made based on a single strategy, where the researcher declare any position and worldview that could affect the description and interpretation of the results conducted (119). Seeking to become a health economist and being the corresponding author of the trial-based economic evaluation of TeleCare North, I am obviously not a neutral bystander in the national decision. I am prone to accept and recommend the methodology and reporting standards that are developed for health care resource allocation and health economic evaluation. Therefore, I can recognize and agree on much of the criticism that there has been of the national decision process which came from health economists in particular. But I also have some insights into some of the reasoning behind these decisions that were unavailable to those critics.

I have no position on the level of treatment and care for COPD patients in Denmark or if telehealthcare as opposed to other technologies should be adopted. I would prefer that only cost-effective technologies are disseminated, but recognize that this would ultimately depend on other rationales than technical efficiency, how certain decision-makers would like to be and how much they are willing to spend to

achieve effects. I also acknowledge that no evaluation can answer all questions that decision-makers might have.

I did not know any of the stakeholders personally prior to the TeleCare North evaluation, but have had a close contact with representatives from North Denmark Region, the municipalities based here, government and the health economists that have criticized the national decision process throughout the study. I also have past experience with the same sort of research problem and stakeholders by being part of the Danish HTA evaluation milieu prior to becoming a Ph.D. student and have some experience with how decision processes are and can be in Denmark.

I have some experience with conducting and reporting scientific mixed methods studies in health informatics, e.g. (160,161) and have respect for the merits of qualitative research methodology, although conducting qualitative studies is not my main skill. In being part of mixed methods research, I have often been struck by how dogmatic both quantitative and qualitative researchers can be in insisting that their approach are the only right way to answer a research question instead of recognizing the relative strengths and weaknesses of each methodology.

My intention with describing and commenting the national decision process is not to put anyone on display or pass judgment. Rather it is to become an “active listener” in order to gain insight into decision-making in order to inform future economic research in telehealthcare.

CHAPTER 4. THE EFFECTIVENESS AND COST-EFFECTIVENESS OF THE TELECARE NORTH TRIAL

The research question posed was: "Is the chosen telehealthcare solution for COPD patients adopted in the TeleCare North trial effective and cost-effective?" The intention was to answer a call for more comparative research on the effectiveness and particularly the cost-effectiveness of telehealthcare to COPD patients. This research agenda is sought answered by three main publications presented in appendices 3-5. A short summary of the contents of the articles is provided in this chapter.

4.1. SUMMARY OF THE EFFECTIVENESS EVALUATION (ARTICLE 3)

Title: Telehealthcare for patients suffering from COPD: Effects on health-related quality of life - Results from the Danish "TeleCare North" cluster-randomised trial (131).

Objectives: To assess the within-trial effectiveness of telehealthcare compared with usual practice for patients with COPD.

Methods: The main outcome of telehealthcare was assessed by following 1,225 COPD patients (578 in intervention group; 647 in control group) for 12 months and having their health-related quality of life measured and compared by the Short-Form 36 (SF-36). The study applied an intention-to-treat principle (162), multiple imputation, good practices for reporting clinical trials extended to CRTs (91,163) and multilevel methods in accordance with methodology guidelines for cluster-randomized trials (164,165). Data for this analysis stemmed from paper-questionnaires sent to patients (health-related quality of life and socio-demographic characteristics) and measurements of physiological indicators made within the office of the patients' GP. Statistical analysis was performed by two separate linear mixed models for continuous outcomes for the mean differences in PCS and MCS scores between groups adjusting for treatment arm, respective baseline PDC or MCS scores, age, gender, baseline forced expiratory volume in one second (FEV1%), marital status, diabetes status, cancer status, and clustering.

Results: Contrary to expectations, the TeleCare North cluster-randomized trial demonstrated no difference in health-related quality of life measured by SF-36. The adjusted mean differences in HRQoL from baseline to the 12-month follow-up were

for PCS: 0.1399 (95% CI, -1.3689; 1.6496) and for MCS: 0.3603 (95% CI, -1.6788; 2.3994). Although, a tendency for a slower deterioration in COPD patients' health-related quality of life in general and for some subgroups were detected, these were not statistically significant either.

Discussion: This study is one of the largest effectiveness studies on telehealthcare in the world and the first large-scale evaluation in Denmark. Based on previous experiences with telehealthcare in North Denmark Region, this study sought to strengthen implementation by ensuring that patients had user-friendly telehealthcare equipment and by gearing the delivery organizations more to respond to reported physical indicators to prevent COPD exacerbations. Still no statistically significant effects on health-related quality of life were detected. Potential reasons include the Hawthorne effect that could have influence the control group; that we have somehow missed important covariates explaining differences in health-related quality of life; that the instrument SF-36 was not sufficiently sensitive to detect outcome differences in patients with COPD or if health-related quality of life is the most relevant effectiveness measure in a hierarchy of outcome measures. Future experimental research should therefore be directed to identifying important mechanisms leading to specific outcomes with the help of qualitative research.

4.2. SUMMARY OF THE MAIN COST-EFFECTIVENESS RESULTS (ARTICLE 4)

Title: Cost-effectiveness of telehealthcare to patients with chronic obstructive pulmonary disease: Results from the Danish "TeleCare North" cluster-randomized trial (129).

Objectives: To report the within-trial cost-effectiveness of a telehealthcare solution in addition to usual care compared with usual care for all participants included in the trial.

Methods: The trial-based economic evaluation was designed as a 12 months cost-utility analysis alongside a CRT with a healthcare- and social sector perspective. Incremental total cost per QALY gained was the main outcome. The study applied an intention-to-treat principle (162), multiple imputation, good practices for reporting economic evaluations alongside clinical trials (166) and multilevel methods in accordance with methodology guidelines for cluster-randomized trials (164,165) and for economic evaluation of cluster-randomized trials (167). Data for the economic evaluation came from national registers (resource consumption in hospitals, GPs offices and medicine), care journals (resource consumption in municipalities) as well as questionnaires sent to patients (health-related quality of life and socio-demographic characteristics) and measurements of physiological indicators made within the office of the patients' GP at baseline (111,129,130). Estimation of treatment effects was based on two separate linear mixed effects

models to allow for different covariates. Total costs were adjusted for treatment arm, baseline EQ5D score, baseline costs, age, baseline FEV1%, presence of musculoskeletal disease and clustering. QALYs gained were adjusted for treatment group, baseline EQ5D score, age, gender, baseline FEV1%, marital status, presence of diabetes, presence of cancer and clustering. A deterministic ICER-estimate was calculated via the treatment effects outputs from both models. Uncertainty surrounding cost-effectiveness was quantified by exporting the outputs from the multilevel models along with Cholesky's decomposition matrix to a Monte Carlo simulation program developed in Microsoft Excel. Redrawing 5,000 new parameter estimates from the estimated treatment-effects with its standard errors, new estimates of incremental QALYs and incremental total costs were calculated to construct cost-effectiveness acceptability curves. Three one-way sensitivity analyses was also conducted: One focused on assessing the effects of a different definition of what constituted a relevant hospital admission; a second on effects of reduced procurement prices and scale of the trial; a third on learning curve effects in monitoring, i.e. reduced monitoring time.

Results: The adjusted mean difference in total costs between telehealthcare and usual care was €728 (95% CI: -754; 2211) from a health- and social sector perspective. The adjusted mean difference in QALYs gained was 0.0132 (95% CI: -0.0083; 0.0346). This lead to a deterministic ICER estimate of €55,327 per QALY gained. Without a willingness-to-pay (WTP) threshold value, decision-makers should be willing to pay more than €55,000 to achieve a probability of cost-effectiveness greater than 50%. This conclusion is robust to changes in the definition of hospital contacts and reduced intervention costs. Only in the most optimistic scenario combining the effects of all sensitivity analyses, does the incremental cost-effectiveness ratio fall below UK thresholds values (€21,068 per quality-adjusted life-year). The telehealthcare solution is therefore unlikely to be cost-effective for all included COPD patients.

Discussion: This study was a large-scale pragmatic randomized trial with embedded cost-utility analysis that has been requested by systematic reviewers in order to improve the current evidence base on the cost-effectiveness of telehealthcare. A relatively broad health care and social sector perspective was chosen and the cost-analyses of resource use are based on register data. The way telehealthcare was implemented may have affected cost-effectiveness, e.g. because involved organizations and healthcare professionals had to find new ways of working together after implementation of the telehealthcare solution and adapt to new work procedures which could have affected resource consumption. The target COPD population might also have proven too broad, e.g. included participants presumably had stable COPD across all severities at recruitment and it is unknown if inclusion of patients with more acute COPD would have improved cost-effectiveness.

4.3. SUMMARY OF THE HETEROGENEITY ANALYSIS OF COST-EFFECTIVENESS (ARTICLE 5)

Title: Heterogeneity analysis of telehealthcare to patients with chronic obstructive pulmonary disease: The case of the Danish "TeleCare North" cluster-randomized trial (130).

Objectives: To explore sources of heterogeneity in cost-effectiveness for patients with COPD.

Methods: The analysis utilized the same design, data sources, reporting standards and methodology as the main cost-utility analysis. The only difference is that it focused on treatment-by-covariate interactions in order to make conclusions about certain patient subgroups. Six subgroups were analyzed post hoc which focused on differences in cost-effectiveness across the severities of COPD, differences in the presence of a set of comorbidities, and differences in cost-effectiveness across age-groups, gender, resource patterns (resource use in social care sector prior to randomization) and monitoring site.

Results: Results indicate that existing resource patterns of patients and variations in delivery site practices might have a strong influence on cost-effectiveness, possibly stronger than the included health or socio-demographic sources of heterogeneity. Across COPD severities measured as GOLD1-4, the highest probability of cost-effectiveness was achieved for severe COPD (GOLD3) (68% at a WTP threshold of €25,000 and 70% at €40,000). At the same time, the probabilities that other COPD severities are cost-effective remain low (between 33-44%). This result seems to indicate that previous cost-effectiveness research identified in the systematic review (article 1) have (inadvertently) chosen the most cost-effective patients for cost-analysis which can explain why a potential for cost-effectiveness and lower costs have been demonstrated for telehealthcare previously. This would not generally be the case across COPD patients. Similarly, telehealthcare to patients younger than 60 years might be more likely to be cost-effective than older COPD patients. No firm conclusions could be made of the included comorbidities and gender.

Discussion: There is almost no existing knowledge of sources of heterogeneity in cost-effectiveness for COPD patients in telehealthcare research. This study has sought post hoc to nuance the available evidence by presenting incremental costs, incremental QALYs and the uncertainty around these estimates in a set of subgroups from a trial-based economic evaluation. A limitation of the study is that the statistical power is insufficient to conclude that the differences found is no more than random noise in the data, but as long as the analyses are exploratory, they could be used to inform future research designs. It is recommended that much more focus should be on how context and implementation affects the cost-effectiveness of telehealthcare, e.g. by integrating formative evaluation into economic evaluation.

CHAPTER 5. APPLICATION OF EVIDENCE IN THE NATIONAL DECISION

The research question posed was: "How was evidence applied in a national Danish decision-making context of whether or not to adopt telehealthcare?". The intention is to place the results from TeleCare North into a national resource allocation context by providing a first-hand empirical account, as opposed to yet another theoretical examination, of the usefulness of economic evaluation in decision-making. Along with a theoretical standpoint in health economics and the experiences gained with the design and conduct of TeleCare North, the intention is to illustrate dilemmas with the current approach to health economic evaluation and also draw out themes that can be used to build additional theory on economic evaluation of telehealthcare.

5.1. THE DECISION

Although not published scientifically yet, the results from the economic evaluation of TeleCare North (results from article 4 and 5) were presented to the steering committee of TeleCare North in August 2015. In October 2015, these results became part of the annual financial agreement for 2016 between the Danish Government, "Danish Regions" and "Local Government Denmark" whereby telehealthcare should be disseminated to "relevant" COPD patients (originally described as patients with severe COPD, GOLD 3 (135)) throughout Denmark (168). Six months later, in March 2016, the Danish Agency for Digitalization presented an implementation plan (134), in which it became up to each region and the municipalities in the regions' geographical area to locally implement the solution (134). It was also decided to continue developing the IT-infrastructure from KIH and TeleCare North and to test the telehealthcare intervention on new patient groups (113).

5.2. METHODOLOGICAL APPROACH

The methodology used to synthesize the available evidence and support the national decision followed the principles and guidelines of the Danish Government's business case model (132,136) designed for government IT-projects that are subject to state accounting rules. In general, business case methodology rests on investment theory, but little scientific literature specific to the design and conduct of business cases exist (i.e. few scientific journals publishes business case results and

methodology). Instead, some books and other material can be found from consultancy agencies and other institutions that agitate for and describe business cases to be used for corporate decision-support, e.g. (169,170). In practice, a business case become associated with investment decisions, where the net present value of expected future cash flows (income and expenses) is analyzed for a given investment and potentially put into a strategic market context for the company considering making the investment. The Business Case Guide for example (170) is an attempt to describe the content of business cases and divides these into five sections: An introduction containing a summary of the business case, criteria for evaluation (net present value or some other criterion) and time horizon for the analysis (170). A methodology section should contain the boundaries for the study (geographical, organizational or other), major assumptions underlying inputs to the analysis as well as the model for calculating income and expenses (170). A results section should include a detailed description of the elements used in the cash flow calculation and potential not-quantifiable effects can be described and justified as well (170). A risk assessment including sensitivity analysis typically involving changed assumptions of input parameters or changing strategic assumptions (e.g. market growth, competitor or supplier behavior, inflation) can be carried out using simple techniques or more advanced simulations (170). Finally, the main conclusions should be stated and often several specific recommendations are described (170).

Most of these business case principles can also be found in the Danish Government's business case model, that also applies to single studies (136). In addition to a similar build-up of the report, three principles are described in the guidelines for designing an reporting the business case model (136). First, the business case model generally relies on expense-analyses based on before and after studies with help of a "program model" (a program theory describing as-is and to-be work flows and organization) (136). The current operating situation must be described and an assessment made of current expenses involved in performing the tasks that the new program seeks to influence (as-is situation). Thereafter, a quantification of estimated development/implementation expenses is required. Two future operational situations (i.e. to-be scenarios) must be described and estimated: A "0-scenario", describing what would likely happen if the program is not implemented along with estimated expenses associated with this scenario. A "1-scenario" describing the program and likely effects if the program is implemented should also be included. For all as-is and to-be situations, effects should be quantifiable, but there is room for reporting "non-economic" effects separately (136). Economic effects can only be included if they can be accrued and realized by public institutions, i.e. be factored into a specific public account (136). Second, cash flows should be described in detail as should data sources and input values. Third, any risks must be described as should at least two sensitivity analyses (a worst case and best case scenario) (136). After evaluation initiation, the business case can be updated in order to manage and regulate expenses (136).

The Government's business case model was used as a funding format for the initiatives in the action plan and it was initially required that all initiatives should continue with this format in their subsequent economic evaluations in order to update the business cases with empirical results from each initiative. The KIH initiative continued with the business case format and chose to focus on a labor saving variant of it (115) by applying a program model to describe current and future work flows for health professionals in municipalities and particularly for health professionals involved in hospital admissions and outpatient visits (115). The evaluation of TeleCare North was the only initiative where results should be published scientifically. Building on the call for more comparative research of telehealthcare and a trial-based economic evaluation of a pilot trial, it was decided to seek to continue with this evaluation format for TeleCare North. So, the changed design for the economic evaluation was a particular focus prior to the implementation of TeleCare North and after some debate, it was accepted that TeleCare North deviated from the business case format. The compromise became that the results from the trial-based economic evaluation should be "translated" to the business case format by the Ph.D.-student and handed over to a consultancy agency in aggregated form in order for them to calculate and recalculate a national business case.

Three rounds of business cases based on the initiatives within the action plan were conducted. First, the ex-ante expectations of expense-reductions were formulated in the action plan and were presented for each of the initiatives based on the business cases used to seek funding (33). Second, in the spring of 2015, an initial national business case was published (132) based on preliminary unadjusted results from KIH and TeleCare North for all included COPD patients. Third, in the spring of 2017, the final business case is expected to be published based on the heterogeneity analysis of COPD severity in TeleCare North.

Relying almost entirely on business cases for decision-making became the subject for a panel debate in the media from April to May 2016 following the decision to implement telehealthcare in Denmark. The debate was initiated with a critical recap (150) of the economic potential described by business cases in the action plan (33) and the initial national business case that were based on preliminary results from both KIH and TeleCare North retrieved in the spring of 2015 (i.e. the first and second rounds of business cases) (132). This recap demonstrated that the national cost-saving expectations was gone from 1,600 full-time equivalents (FTEs) to a negative profitability for all the KIH projects (DKK 0.1 mill.) and from cost-savings around 700 FTEs to a negative profitability (DKK 5 mill.) for all 1,225 patients in TeleCare North. This lead the author to conclude that particularly ex ante business cases were manifestations of "fever fantasies" (150) and that other evaluation approaches, such as HTAs, should be applied in the future instead (159).

Indeed, the evidence mentioned in the formulation of the "national action plan for dissemination of telemedicine" (33) were based on a questionable evidence synthesis methodology that lead to high ex ante expectations. In addition to the business cases used to seek funding (where a "positive" business case was a selection criteria…), the following argument were used illustrating this point:

> *"…there are great expectations for the economic potential of telehealthcare solutions - partly on the basis of small pilot projects in Denmark, and partly on the basis of major implementation projects abroad, e.g. Scotland and the UK. The British Ministry of Health have in December 2011 announced the preliminary results of a large-scale demonstration project with the following conclusions: The early indications show that if used correctly, telehealth can deliver a 15 percent reduction in emergency visits, a 20 percent reduction in admissions, a 14 percent reduction in bed days and an 8 percent reduction in tariff costs. More strikingly they also demonstrate a 45 percent reduction in mortality rates*" (33, p6) (translated from Danish).

It turned out that the expected outcomes of telehealthcare referred from the press release were based on 6,191 patients from the Whole System Demonstrator Project (WSD) that were never published scientifically. Critics within the UK and abroad did also heavily criticize the British Department of Health for the press release saying that these results were "*sensationalist and cherry-picked*" from unpublished data (171). It is also unclear which "small pilot projects in Denmark" that are referred to, but most notable studies had been pilot efficacy trials completed in North Denmark Region (22,172), the Capital Region of Denmark (23) and in the Region of Southern Denmark (24). The North Denmark Region study included a trial-based economic evaluation and concluded that telehealthcare was cost-effective after 6 months for the included 111 patients (60 in intervention group) with severe and very severe COPD (172). The Capital Region study did not include an economic evaluation, but concluded that no statistically significant differences were found in health-related quality of life at 3 months (23). The Region of Southern Denmark study found no statistically significant difference in mortality or readmissions at either 4, 8, 12 and 26 weeks for 266 patients with severe COPD (132 in telehealthcare group) (24).

Both Danish Regions and the Danish Agency for Digitalization has defended the use of business cases by stating that they are the "*best solution*" (149) and "*the only*" general standard for prioritization within and between sectors (142). But they also recognize that business cases must include "qualitative" evidence as well and be continually updated (142). Indeed, the national business cases were updated with input from the initiatives in the action plan. Answering only for the TeleCare North trial; the changing cost-conclusion reflect a move from an efficacy trial (the

efficacy trial in North Denmark Region that was the foundation for the business cases calculated here), to an effectiveness trial, where outcomes can be smaller than under ideal circumstances (an efficacy vs. effectiveness issue). It also reflect that the efficacy trial in North Denmark Region only included patients with severe and very severe COPD, where patients with severe COPD later turned out to be the only severity-group that were likely to lead to cost-savings in TeleCare North (a "selection bias" issue) (130). But health economists would not agree that business cases are the best, nor the only, option for prioritization within and between sectors, as described in the theory-chapter (82). Making such strong claims must be because full economic evaluation is unfamiliar or undesired by decision-makers.

That there is a lack of recognized requirements for the quality of documentation of evidence in business cases is a challenge, which was also pointed out (143):

> *"…it is important to note that decisions should not be based on an IT manager, two project nurses or fifteen specially selected patients that find the effect of telehealthcare good, but that there exist a high level of evidence"* (143) (translated from Danish).

The KIH business case results were published first in August 2015 (108). There were differences in the profitability of the five included initiatives, but there were a negative profitability overall across initiatives. The results from the COPD initiative targeted patients with severe COPD demonstrate a 21% negative profitability primarily due to no difference in the number of and duration of hospital admissions and the intervention costs (108). But the KIH evaluation was subject to harsh criticism by health economists for its business case design and applied methods, e.g. for not including all relevant intervention costs, for including only labor saving costs, for being an efficacy trials with questionable scalability of the conclusions and for including expert opinions instead of actual data on resource use in many cases (147,158) leading one author to state:

> *"If it was one of my economics students who had handed in the KIH report, I would have flunked this student"* (158) (translated from Danish).

In many ways the business case format deviates from the design of standard health economic evaluation and can in some ways probably be better described as what health economists calls *budget impact analyses* that can be used to complement health economic evaluation (173). In addition to the evaluation design, the main difference can be found in objectives; while health economic evaluation evaluates costs and health outcomes for their *technical efficiency* for resource allocation decisions, budget impact analyses addresses cash flows related to development and implementation in order to assess *affordability* for the involved stakeholders (173). This could explain why the methodological approach was seen as "the only" and

"best" solution, if stakeholders are used to evaluations based on affordability. But then evaluation skips a step, where affordability becomes a more important criterion for resource allocation than efficiency. And the use of business cases can only be an approximation to budget impact analyses, since business cases are not based on a logic of evidence hierarchies (budget impact analyses are) and the epidemiology and/or etiology of the disease under investigation are seldom used in business cases (budget impact analyses do). In contrast to health science, where an "evidence hierarchy" is one of the foundations for assessing study quality and designing studies with a high validity, there is no stringent methodology for designing business cases or for choosing underlying studies here. If business cases apply studies at the lower levels in the evidence hierarchy, there is a risk that they will become misleading in interpreting a causal link between the telehealthcare technology and observed outcomes. In this regard, it is interesting that the action plan initially suggested an HTA approach to evaluation (33), which was not conducted in practice. The advantage of an HTA approach would have been that the methodology applied here *is* based more on a health science tradition of grading evidence and reflecting the investigated disease.

5.3. STUDIES USED

The economic evaluation of TeleCare North was presented to the public three months after KIH in November 2015. The main conclusion for all 1,225 patients (article 4) was presented along with a heterogeneity analysis of especially differences in cost-effectiveness across patients with different COPD severities (article 5). The heterogeneity analysis in particular comes to the opposite conclusion than in KIH by concluding that the probability of cost-effectiveness is higher for patients with severe COPD.

That the economic evaluation in TeleCare North was allowed to deviate from the business case format became essential to the decision and it has been used actively by highlighting that it was a health economic evaluation from TeleCare North that were used as input to the national business case (142). The unpublished and not peer reviewed results from the economic evaluation of the TeleCare North trial therefore became the main evidence used in the decision to implement telehealthcare in Denmark (138). Although, the KIH project is mentioned by the Danish Agency for Digitalization, the emphasis is officially made on describing the experiences from TeleCare North and in contrast to TeleCare North, no links are provided to KIH on the web-page informing about the dissemination of telehealthcare in Denmark (138).

> *"Good experience from two large scale telehealthcare projects Telecare North and KIH, has paved the way for the political agreement […]. Lessons from especially Telecare North shows that telehealthcare increases the quality of life for people with COPD,*

> *because the citizen experience increased security, freedom and being able to act for themselves during exacerbations […]. Read more about the experience of Telecare North"* (translated from Danish) (138).

The same critical voices that questioned the validity of the KIH study have criticized this single-study strategy (144):

> *"Based on, at best, a single analysis documenting cost savings - other studies are "forgotten" - is an economic agreement for 2016 between the Danish Government, Danish Regions and Local Government Denmark agreed that involves nationwide dissemination of telemedicine for COPD patients by the end of 2019"* (translated from Danish) (144).

There are no publicly available documents and information on national decision-makers motives for choosing to focus on the TeleCare North evaluation, so there is no way of knowing whether it was because of the criticism raised of the KIH evaluation or for some other reason. But there were other options, than deciding to disseminate telehealthcare to COPD patients based on a single study.

One option was to apply evidence generated internationally that used a comparable design (i.e. a cost-utility analysis) and some studies did exist, primarily from the UK: The WSD in England had been published as a series of publications from 2012 to 2015. The overall results are not specific to COPD (they include patients with diabetes and heart failure as well), but still interesting, since they provide the first large-scale evidence on the effects of telehealthcare in the world and as such are hard to ignore. The cost-effectiveness study by Henderson and colleagues concluded that the telehealthcare solution was unlikely to be cost-effective after 12 months (ICER=£92,000; probability of cost-effectiveness=11% at £30,000) (125) and previous results had already demonstrated a reduction in mortality and overall admission proportion and bed days, while differences in other hospital use (elective admissions, outpatient visits, emergency department visits) were non-significant (125). In addition, there were no improvement in health-related quality of life or any of the psychological outcome measured (126). Rixon and colleagues published patient-related outcomes for the COPD subgroup within the WSD in 2015 (123) and concluded that the 447 participants showed that telehealthcare had a "minimal effect" on health-related quality of life and psychological distress and only emotional functioning and mastery were significantly higher with telehealthcare (123). In Scotland, the Telescot program was also focusing on telehealthcare for various patients with chronic diseases (174). A 12 months clinical trial from Telescot testing the effectiveness of a telemonitoring solution specifically to COPD patients concluded in 2013 that telehealthcare was not effective in improving health-related quality of life, postponing hospital admissions or reducing the

number and duration of readmissions (128). This was later followed by the results from the economic evaluation alongside the trial, which revealed that the telemonitoring intervention gave rise to higher costs, which meant that the intervention was unlikely to be cost-effective (ICER=£137,277 per QALY; probability of cost-effectiveness=15% at £30,000) (127).

These international results were acknowledged by decision-makers and are described directly in an overview of experiences that was reported as part of the implementation of the action plan (25, pp39-41). But international studies are excluded from the synthesis in the initial national business case conducted by PA Consulting group (132, p1).

A second option would simply have been to postpone a decision until further (Danish) evidence existed. Both options could have been founded on a decision-model with a value of information analysis as described in the theory-chapter.

5.4. HIGHLIGHTED RESULTS

The economic evaluation of TeleCare North was designed as a full economic evaluation. In this regard, it is interesting that none of the identified sources, other than the scientific studies, actively use the ICER (additional cost per QALY gained) that directly relates costs with an outcome. Instead results from TeleCare North are transformed into results used in partial evaluations with much emphasis on cost-savings in both the national business cases and in the communication about the results.

The results from the subgroup of patients with severe COPD in the economic evaluation quickly came into focus with the cost-savings that were likely here.

> *"[Heading: Economic gains depend on the target group]. TeleCare North has shown that the economic benefits of telemedicine are greatest for people with severe COPD. TeleCare North estimates that for this group annual savings are DKK7,000 annually per patient to the municipalities and regions. This cost-saving is mainly due to fewer hospitalizations. If telemedicine is offered to all patients with COPD regardless of COPD severity it is associated with additional costs of approximately DKK5,400 per person per year. The reason is that telemedicine cannot reduce the number of admissions to the same extent for the entire target group as with citizens with severe COPD"* (133) (translated from Danish).

In health research, it can be difficult for analyses that investigates subgroup effectiveness to be published, because these results are explanatory, meant to inform clinical practice and there is a risk that effects are found by chance (175).

However, it is entirely appropriate to conduct subgroup or heterogeneity analysis in cost-effectiveness research used for decision-making, because these are concerned with whether or not to fund certain courses of action, not treatment (176,177). However, there are certain procedures and ways to report findings from heterogeneity analysis that must be met (176,177). In health economic evaluation this is done by pre-specifying potential subgroups prior to analysis and these should be likely to be operationalized in routine practice and give clinical or economic meaning (176,177). But contrary to health research traditions, reporting results from subgroup analyses are not based on differences in treatment-by-covariate effects being statistically significant (176,177). Rather, the uncertainty surrounding the cost-effectiveness estimate in subgroups are quantified, mostly by a technique called “probabilistic sensitivity analysis” (176,177).

In the national debate following the dissemination of telehealthcare, some argued that the focus was too much on the subgroup of severe COPD (144). I have responded that the reporting of subgroup results from the economic evaluation of TeleCare North in the decision-context has been mostly fair and balanced (153). A focus on the subgroup of severe COPD could also reflect a desire from decision-makers to communicate why this decision was made and that the overall results were always presented along with the subgroup results (153).

But some of the results from the economic evaluation of TeleCare North were used without some of the reservations or precision that must be made. For example, it is stated that there is an effect on quality adjusted life-years both in the overall trial results and in selected subgroups:

> *“The health economic research shows that telemedicine has an effect measured in quality-adjusted year of life for the vast majority of patients. The effect is very similar to the results from the Whole System Demonstrator project in England. The research also shows that telemedicine for patients with COPD should be targeted […]. When telemedicine is targeted to patients who have severe COPD, GOLD 3, it is possible to achieve an effect in terms of quality-adjusted life years while saving costs. The cost-savings are around DKK7,000 per citizen"* (108) (translated from Danish).

It is true insofar that health economists are interested in the absolute differences in outcomes and costs in the calculation of an ICER (178). But if health outcomes are reported separately from costs, it is still necessary to focus on statistical significance in order to inform clinical practice (178). Instead, the results from probabilistic sensitivity analyses should have been much more in focus, but are only mentioned by researchers. More precision in communicating the results could be warranted, but it is a complex message to understand and communicate when no health economic tradition is routinely applied in decision-making.

Some have critiqued that health or patient-related outcomes were missing from the national decision (140,141,146). This is partly because few effectiveness studies had been conducted at the time of the decision (e.g. mortality or physiological indicators were secondary outcomes in the trial and not analyzed at the time of the decision). But the effectiveness study from article 3, reporting the results of the main trial outcome (no statistical significant difference in health-related quality of life measured by SF-36 on all 1,225 patients or any subgroup) were presented alongside the economic evaluation, but this result is not highlighted anywhere. Patient-related outcomes that were published first while the clinical trial was being implemented, such as initial positive user experiences with the Telekits and the increased sense of security and control it lead to for them, were used to support the decision made, e.g. in the initial national business case (132) and in the communication by the Danish Agency for Digitalization (133). However, these results could also have been used with more precision: A representative from Danish Regions summarize the results from TeleCare North in the following symptomatic way (156):

> *"[Heading: Clear benefits]. The benefits of the upcoming nationwide dissemination of telemedicine are clear. In the middle of November, the parties behind one of Europe's large-scale projects in telemedicine, Telecare North, published the results of nearly four years of investigation by almost 1,400 COPD patients from North Jutland. The conclusions were:*
> *• 6 in 10 people feel that they can better control their own illness with telemedicine.*
> *• 7 out of 10 feel safer in everyday life with telemedicine.*
> *• 9 out of 10 find the telemedical equipment user-friendly"* (156).

But these data were published in a conference proceedings journal and only based on 60 patients from the intervention group before the trial was completed (179) and the representativeness of these patients to the decision made are uncertain (e.g. all 60 patients did not have severe COPD).

5.5. HIGHLIGHTED DECISION-UNCERTAINTIES

In addition to mentioning that the organization of telehealthcare monitoring may be different in each region (108), when telehealthcare is eventually disseminated in Denmark and the potential for targeting the peripherals in the Telekits in the national business case (132), little is communicated to the public about decision-uncertainty on the national arena. One reason was that the scientific papers from TeleCare North, with the reservations described herein, were in peer review and could not be disclosed publically.

But in the national debate of the strengths and weaknesses of business cases, a representative from Danish Regions revealed that there was decision-uncertainty, but also that it is important to take risks (149):

> *"It takes courage to drive public sector development. It would be easy to simply looking at everything that is going well and not take any further action. We have in Danish Regions chosen a different strategy which implies that we sometimes have to take a chance and take risks in order to grow and to create better solutions"* (149).

The Danish Agency for Digitalization do make a reference to decision-uncertainty (134) by linking to a report published by North Denmark Region in November 2015 that contains their experience with the implementation of the TeleCare North initiative (108). In it, a set of major circumstances are highlighted that they feel could potentially have affected the outcomes of the TeleCare North trial, which are not accounted for in the sensitivity analyses made in the economic evaluation.

First, the initiation of TeleCare North coincided with a national labor dispute between the Danish Medical Association and Danish Regions in 2013 (108). Although, there was positive support for TeleCare North by the Danish Medical Association and most GPs in North Denmark Region, some GPs were assessed to influence the initiative negatively by affecting both cooperation and dialogue. This made it challenging for municipalities to have "*a standardized and close dialogue with general practice for all their citizens*" (108).

Second, TeleCare North used existing national solutions and contributed to the establishment of a new national IT-infrastructure where IT-solutions were missing (108). A central national IT-solution to be developed was a dataset ("Kronikerdatasættet", no English translation) that should allow for cross-sectorial data-sharing of home monitoring data between GPs, municipalities and hospitals and integrated into existing local IT-systems. The work with this dataset was eventually disbanded. Instead, a web-based portal solution was made for TeleCare North ("OpenTele") to allow for data-sharing between municipalities and hospitals and an existing portal ("www.sundhed.dk") was used by GPs (108). This meant that the cross-sectorial coordination could probably have been even better throughout the trial period. The national decision therefore also acknowledged that a prerequisite for implementation is an increased maturation of the IT-infrastructure to further support telehealthcare monitoring (113).

Third, introducing telehealthcare required new skills to be learned especially for nurses, since they were given the task of training and monitoring patients (108). This was a new task involving training in order to be able to read and evaluate the conducted measurements, to recognize any relevant patterns or change in the development of vital values in patients and to be able to act proactively (108).

Combined with challenges in making the hardware and software configuration running smoothly, this could have led to a "*significant overuse*" of resources in the beginning of the trial (108):

> *"Hospitals, municipalities and general practice have entered into the project without any experience. This meant that an organizational maturation occurred while data was collected from the control and the inclusion group for the evaluation […]. Our assessment is that we were far into the research efforts before well-established routines and skills associated with the telehealthcare effort was established"* (108, p40).

A comprehensive search for heterogeneity and the effects of changed assumptions was also conducted post hoc in order to come closer to demonstrating *why* the telehealthcare solution was not more cost-effective than demonstrated. Much of this work is unpublished, but central points have been included in the sensitivity analyses in the main cost-effectiveness analysis in article 4 and the heterogeneity analysis in article 5. The overall results were relatively robust to changed assumptions, but heterogeneity analyses demonstrated higher probabilities of cost-effectiveness for "system-variables", than for any health- or socio-demographic characteristics, i.e. there were variations from 0% to 100% in the probability of cost-effectiveness across municipalities and the probability of cost-effectiveness was much higher for patients having existing resource use in the municipalities at evaluation outset than patients who did not (89% vs. 18%) (130). Of course targeting patients based on contacts with healthcare providers would be considered unethical, but it does indicate that "implementation" could be important for achieving cost-effectiveness, which is also argued in article 5 (130).

This point was also made by fellow qualitative researchers from TeleCare North in the national debate (151,152). Based on the qualitative research conducted in TeleCare North, it was e.g. argued that the TeleCare North initiative is about organizational development and change requiring constructive cooperation on all organizational levels, conflict resolution techniques, local adaption and continuous attention and adjustments that might not even be fully in place to this day:

> *"To implement a cross-sectorial telehealthcare solution is not something you do with a snap of the fingers. But if it can be done, the potential of telehealthcare can prove to be even greater than the TeleCare North evaluation has demonstrated, because the concept is still evolving"* (151)

That implementation is "crucial" to the effects of telehealthcare was also seen as implicating that quantitative evaluation methods needs to evolve in order to achieve a higher external validity (152). Trial-based evaluation methods developed in health

science and utilized in health economic evaluation does not sufficiently take implementation into account and should be complemented by "a conscious approach" to implementation (152).

5.6. INTERPRETATION

In Denmark, we have a more or less ad hoc national setup for assessing the effects of technologies such as telehealthcare. No single authority have the formal responsibility for ensuring that a systematic synthesis of evidence is conducted and to ensure that applied methods are suitable for the technology to be assessed. This is often seen as a disadvantage, since this implicates that the applied methodology and criteria for sound evaluations come to reflect the purview of the national institution(s), where the decision process is placed each time. Rather old Danish guidelines for economic evaluation of pharmaceuticals exist and the European Medicines Agency can be relied on approval of pharmaceuticals for the Danish market as well. But the only requirement for telehealthcare to be approved is in principle a "Conformité Européenne" (CE) marking on the equipment, which does not guarantee that the equipment actually leads to desired benefits. This will not change with the new "Medicinråd" established by Danish Regions in 2016 that in the near future only covers hospital medicine.

On the other hand, the Danish "national action plan for dissemination of telemedicine" *was* an attempt at systematizing the testing and evaluation of telehealthcare in Denmark. A lot of good intentions were put into the development of the plan, e.g. a willingness to try a standardized evaluation mindset, attempting to coordinate the included initiatives, using large-scale cross-sectorial initiatives to test the implementability and cost-saving potential demonstrated in some efficacy-trials (essentially a phased development mindset as with pharmaceuticals). But there were also significant lessons to be learned from this action plan and the evaluation approaches it contained as events unfolded, which are described below (relating to the themes "Methodological approach", "Highlighted results" and "Highlighted decision-uncertainties").

In my opinion, the first lesson is that for telehealthcare there is no getting around making decisions that are based on underlying evaluations with a balanced outcome focus and randomization. These evaluations might be challenging, time consuming and expensive to apply alongside implementation, but we need these types of evaluations and it would also be difficult to engage healthcare professionals and researchers and convince them of the soundness in decisions, if studies with low internal validity and a single expense-outcome is used. The greatest risk of using business cases as a methodology for resource allocation, is that "qualitative effects" become secondary outcomes at best and that no criteria of what constitutes good evaluation designs is an inherent tradition here. I agree with the criticism that emerged from health economists in arguing that health economic evaluation should

be the foundation for resource allocation decisions, because these designs are designed for efficiency decisions and minimize the risk of bias and confounding, and because reporting of health economic evaluation are based on relatively scientifically well-defined good practices that are both transparent, standardized and stringent.

However, we should remember that health economic evidence *was* useful for the decision and incremental costs was applied in the national decision-making process, albeit possibly by coincidence or foresight by stakeholders in North Denmark Region, and mostly translated into a budget impact form. The decision to implement telehealthcare in Denmark relied on a single-study strategy that was part of the evidence foundation for the national action plan (the TeleCare North trial). This can be viewed as a start, even though a single-study approach is not recommended for health care decision-making and TeleCare North was not a "reference case" trial-based economic evaluation either. Other options were available such as including other evidence (e.g. from KIH or international evidence generated from 2013-2015) or choosing to postpone a decision. In that way, more data for calculating the expected costs and/or benefits of telehealthcare could have been included in decision-making. But that could have been a challenging argument to sell to decision-makers given the effort and expense that were spent in the action plan.

A second lesson is that telehealthcare interventions are difficult to implement and that effects of implementation can be hard to foresee, quantify and make visible in cost-effectiveness evaluation. Several factors could in theory have affected cost-effectiveness outcomes positively and negatively in TeleCare North and some concrete factors were mentioned. This underlines that a health economic evaluation cannot be seen in isolation, without the circumstances it was conducted in and that implementation can have an impact on the evaluation result. With trial-based economic evaluations, we usually have no direct way of demonstrating if we have a telehealthcare solution that has released its full cost-effectiveness potential or if this potential has been hampered by a "not optimal" implementation or other contextual factors to a certain degree.

Combining these two "lessons" creates a pressing issue. It would probably be too far-fetched to claim, that national decision-makers are ready for making decisions based on traditional health economic evaluations in the future. Politically or administratively, there might be an increased recognition of the strengths of health economic evaluation of telehealthcare in answering if telehealthcare works and for whom compared to other research designs, e.g. more initiatives are now tested in Denmark and North Denmark Region has received national funding for a new initiative, "TeleCare North 2" to chronic heart failure patients that includes a new trial-based economic evaluation to be used for national resource allocation decisions (137). If such an increased recognition of trial-based economic evaluation

has gained momentum, then we also need better trial-based economic evaluations that address challenges that are threatening both internal and external validity. As long as the usefulness of single cost-effectiveness studies could easily be called into question, it becomes less relevant to debate which approach for evidence syntheses should be chosen in resource allocation decisions. From both a scientific and decision-maker standpoint, there is a need for developing trial-based economic evaluation of telehealthcare by incorporating design elements that lead to higher decision-certainty. When decisions has to be made of whether or not to disseminate telehealthcare solutions from one region to the entire country, is it enough to answer questions such as “was telehealthcare cost-effective and for whom”, when there is an increased need to answer questions such as “how can the intervention be improved or adapted to be more cost-effective?” and “how much more cost-effective can telehealthcare be expected to be?”.

Failing to address these types of questions can result in “trench warfare” every time trial-based economic evaluations of telehealthcare are published. For example, in interpreting the results from recent trial-based economic evaluations, one position could be that these recent studies of telehealthcare are larger, have broader cost perspectives and are more pragmatic, so telehealthcare is not cost-effective and earlier studies suggesting the opposite conclusion are questionable laboratory experiments with poorly designed cost-studies. The opposite position would be that telehealthcare is cost-effective and failure to demonstrate this in recent studies is because newer studies have a lower internal validity and are not strict replications of older studies, so something in the intervention or evaluation have gone “wrong”. If decisions to implement telehealthcare or not depends on a dominant coalition of stakeholders belonging to one or the other of these positions, we have really not demonstrated genuine interest in bringing research and better decision-making forward. The usefulness of trial-based economic evaluation in healthcare resource allocation would therefore be open to easy criticism and we fail to utilize the momentum that could have been generated or we risk implementing a system of “opinion-based evidence” - which is a contradiction in terms of course. In this regard, that we do not have a systematic approach for including economic evidence in healthcare resource allocation in Denmark, could actually also be seen as an opportunity to investigate what health economic evaluators can do differently or better in order to achieve an even broader foundation for decision-making.

CHAPTER 6. REFLECTIONS ON HEALTH ECONOMIC EVALUATION OF COMPLEX TELEHEALTHCARE INTERVENTIONS

This chapter is based on reflections initiated by decision-makers and other researchers across the country that have asked me each time I was invited as speaker on conferences or meetings throughout 2016: “Why was TeleCare North not more effective and cost-effective?” and “Could you point to what could be done differently in future implementation or evaluation of telehealthcare?”. Based on the TeleCare North evaluation, no firm answers to these sorts of questions can be made, but they sparked the initiation of a search for theories, experiences and evaluation approaches that could inform future trial-based economic evaluations. The topics covered in this chapter are grounded in health economic evaluation theory, experiences with the evaluation design of TeleCare North and its implementation as well as themes from the national debate of evidence. These topics are combined with different evaluation paradigms and theoretical approaches to evaluation that could be relevant for trial-based economic evaluation of telehealthcare. In this respect, this chapter wants to create a discussion of how to develop trial-based economic evaluation of telehealthcare further by inductively attempting to add theory to existing trial-based economic evaluation.

6.1. ACKNOWLEDGED CHALLENGES IN THE EXISTING COST-EFFECTIVENESS EVALUATION APPROACH OF TELEHEALTHCARE

In general, a strong case has been argued for conducting trial-based economic evaluations of medical devices in general and by implication also telehealthcare (180–182). These technologies are also incorporated into the program by NICE (183,184). Conducting trial-based economic evaluations is considered a strength since it borrows the internal validity of randomization from the clinical trial. Randomization should ensure that patients in the alternatives are comparable in every way except the treatment under consideration, which leaves only the different treatments as explanations for observed differences in outcomes (96). In other words, if research questions explicitly involve cost-effectiveness objectives (i.e. determining whether or not there is an *extra* effect or costs *of an intervention* and the *magnitude* of these extra effects), then experimental evaluation must be at the top of the “evidence hierarchy” (90).

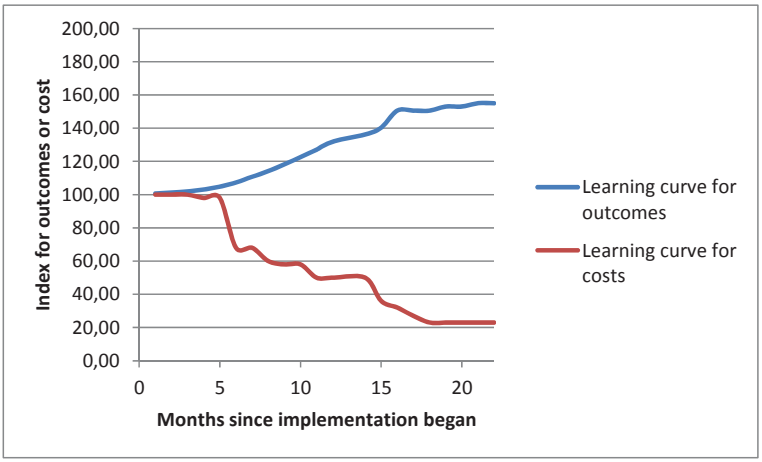
But, it is increasingly recognized within health economic research that medical devices such as telehealthcare are different than the pharmaceuticals that economic evaluation were developed for assessing and that this poses unique conceptual and practical challenges that are difficult to handle in existing trial-based economic evaluation (181,182,185–188). These challenges are primarily seen to constitute threats to internal validity: First, constraints in study design primarily related to problems with blinding telehealthcare initiatives in the underlying trial have several implications. Involving both a device, training and monitoring, it cannot be avoided that it becomes visible who receives treatment and who does not in telehealthcare trials (182,185,186). This implicates that patients or healthcare personnel risk affecting trial outcomes, e.g. by having or giving different expectations to treatment alternatives, by affecting trial drop-out or by implementing heterogeneous administration of interventions (100). A real risk of contaminating treatment alternatives are also present, since intervention patients or healthcare professionals can inadvertently use intervention training (e.g. disease specific training) on control patients (165). Patients may also refuse to enter the clinical trial if they are afraid of using or are unfamiliar with the device and disappointed patients can choose to drop out of the trial after being randomized to control treatment (165). Second, there are many more explanatory factors for telehealthcare cost-effectiveness than for pharmaceuticals (185,188); telehealthcare cost-effectiveness can e.g. vary with the user friendliness of the device and peripherals, many different characteristics of patients, the skills of health professionals and the organizational setting (188). Third, learning curve effects are plausible, since baseline skill and experience of patients and health professionals in using telehealthcare will change over time. This means that cause-and-effect assessment entails comparing familiarity with one alternative compared to unfamiliarity with the intervention, adding a dynamic element to cost-effectiveness evaluation (181,182,185,186) Fourth, wider organizational implications of adopting telehealthcare is likely (181,182,186), e.g. implementing telehealthcare often entails reorganization of services and work flows.

One way to meet some of the challenges with internal validity of trial-based economic evaluation of telehealthcare (primarily the first and second challenges from above) have been to focus on even more sophisticated designs, more complicated statistical analyses and elaborate data collection, such as the CRT design (189,190). This could minimize the risk of contamination and allow for inclusion of organizational-level covariates. But designing and implementing CRTs has a “price”: Firstly, CRTs are less statistically efficient than RCTs (165). Depending on how many patients that are in each cluster and how correlated outcomes are within these clusters, the required sample size necessary for achieving sufficient statistical confidence in the results can be significantly larger than for RCTs. There is generally a need for larger studies in telehealthcare research, but conducting CRTs reinforce this need. And large studies are expensive and time consuming. Secondly, selection bias is a greater concern for CRTs and can be

difficult to avoid in practice (165). Because CRTs applies a different randomization unit (i.e. the general practitioners in WSD or municipalities in TeleCare North), preparation of delivery organizations become challenging. In order to conceal treatment allocation, delivery organizations (and patients) must be recruited before randomization. But why should delivery organizations e.g. send healthcare professionals to courses in disease-specific training or in using the telehealthcare equipment, if they end up being allocated to control treatment? Preparing delivery organizations after randomization could in theory be done. But in large-scale trials this would usually result in a much delayed implementation of telehealthcare at the patients' homes leaving a much shorter period for the intervention to have an effect in the trial period. And this strategy can also be unethical if it leads to study drop-outs (increasingly ill patients will likely have a reduced cognitive or emotional energy over time and risk dying!). Thirdly, compared to randomizing individuals, imbalances between treatment alternatives have a higher chance of occurring in CRTs, particularly if the numbers of clusters are small (165). Methods for avoiding and adjusting these imbalances should therefore be used to a larger extent (e.g. using cluster-covariates in addition to patient covariates or by matching clusters). However, methods for trial-based economic evaluation are not as well-developed as cost-effectiveness methods alongside RCTs. Although, some authors have pointed to potential methods for analyzing cost-effectiveness data in CRTs (167,191), it is still a field very much in development. Depending on the concrete design of the CRT (the number and balancing of clusters and number of participants within clusters) as well as the "behavior" of the data, there are several statistical methods that can be applied. None are optimal and all of them would probably not yield different cost-effectiveness conclusions, but certainly different sizes of incremental cost and incremental effects. Fifthly, variations in local treatment by delivery organizations within and across clusters, e.g. in how training was actually conducted, who and how measurements were conducted or monitored, are much more difficult in CRTs, because activities are dispersed over different locations. A way to handle this is to demand that delivery organizations adhere to strict treatment protocols, but this may be hard to accept by them and adherence in large-scale evaluations is very difficult to control as a researcher.

The third challenge (learning curves) also implicates that patients and delivery organizations undergo a "shakedown phase" after implementation, i.e. because it takes time to learn new skills and to adapt new daily routines and work flows (192). Figure 3 illustrates (probably an exaggerated point) that evaluating cost-effectiveness in this time period may not reflect the likely cost-effectiveness that could be expected when the telehealthcare technology finds its steady-state operation over time for patients, healthcare professionals and the involved organizations.

Figure 3: Two different learning curves illustrating that incurred costs and outcomes can be a function of time. Source: Own contribution



Practical solutions include allowing patients and healthcare professionals to gain familiarity with the telehealthcare intervention before beginning the evaluation. But identification of a steady-state operation can be difficult (e.g. if curves do not flatten as they do in Figure 3) and it would lead to a shorter trial-duration. Instead, some learning curve effects was incorporated in TeleCare North post hoc by analyzing the impact on cost-effectiveness of reducing the average time spent per patient on monitoring physical measurements to reflect the monitoring time used in the end of the trial period (129). Another solution is to use expert users, but this would come into conflict with the "effectiveness" objective of the economic evaluation, where the interest is on the likely effects as close to real practice as possible. To address organizational implications (the fourth challenge), an unpublished sensitivity analysis focused on the failure to reorganize workflows surrounding outpatient visits by eliminating certain visits from the cost-effectiveness calculation and analyzing the impact (some patients were requested to show up for an outpatient visit where physical measurements were taken instead of using the measurements conducted at home, since hospitals in some cases forgot or could not see that patients were included in the TeleCare North trial).

So, some learning curve effects and wider organizational implications can be handled by traditional sensitivity analyses in trial-based economic evaluation, especially post hoc, but in general they are a challenge in experimental evaluation (i.e. it can be difficult to identify a priori where learning curves and organizational implications would take effect before evaluation). Should experience-based requests for modifications of technology or work flows after evaluation initiation arise (e.g. if it is discovered that more optimal work flows or technology features would lead to higher effectiveness or lower costs), an adaptation of the intervention

in the evaluation period is generally rejected in clinical trials, since this could introduce confounding. This was also an issue in TeleCare North, where modifications were requested but halted by evaluators in order to maintain internal validity in the trial. An "easy" theoretical solution to this challenge is to conduct faster evaluations, but a requirement for economic evaluation is that the duration of the study should be as long as possible in order to capture all relevant benefits and costs (82). This duration is ideally sometimes several years, which means that the results from a cost-effectiveness analysis of telehealthcare risk being obsolete by the time of publication.

In summary, conducting "traditional" trial-based economic evaluation of telehealthcare can therefore be a huge and challenging undertaking that requires significant time, funding and data. If sophisticated designs and analyses are applied and if practical dilemmas are addressed to the advantage of evaluation, some of the issues surrounding internal validity are solved. But the scale and complexity of the evaluation approach makes it a project that cannot realistically be conducted routinely. And it would still be a task permeated with analytical dilemmas especially related to a priori identification of learning curve effects and wider organizational implications of implementing telehealthcare. This has lead a review from 2015 to conclude that the implications of the unique characteristics of medical devices are "*either not addressed or insufficiently addressed*" in economic evaluation (186). However, few concrete solutions are provided and if they do, they usually take the form of analytical adaptations to account for problems with internal validity (181,182,185–188).

6.2. REALISTIC EVALUATION AS AN ALTERNATIVE EVALUATION CONCEPT

Instead of focusing on internal validity in answering "what worked and for whom?", a few health economists have argued that trial-based economic evaluation should instead focus more on addressing implications of the clinical logic underlying health economics in order to move health economic evaluation forward (193–195). In order to achieve higher explanatory power and usefulness for decision-makers, it is necessary to focus more on external validity to account for the context in which interventions are embedded and the agency of participants in producing those outcomes (193–195). Health economic evaluations are conducted as "*largely a-theoretical exercises in measuring or estimating the inputs (resources and costs) and outcomes*" as a health economist recently put it (196, p328).

Theory of agency and context are core evaluation elements found in "realistic evaluation" (197). Originating from a completely different philosophy of science paradigm, critical realism, the realistic evaluation tradition is one of the harshest explicit critics of experimental evaluation. Originally termed by Pawson and Tilley in 1997, realistic evaluation completely discards experimental evaluation as a tool

for evaluation in the beginning of their book (chapter 2 “out with the old”), since experimental evaluation does not provide any real lessons by only answering if an intervention worked or not and sometimes for whom (197).

It is argued that experimental evaluation has misunderstood scientific methodology by assuming that knowledge can be accumulated based on theoretically uninformed observations of outcome differences and descriptions of interventions (197). Experimental evaluation is said not to understand what really happens in the “change engine” of interventions, meaning that this design cannot explain *why and under which circumstances* benefits are (un)likely to occur (197). Much attention in experimental evaluation has been paid on concrete intervention activities and too little on change mechanisms and how they interact with each other and the context in which they must operate. Mechanisms are hidden and unobservable, but central underlying causal processes leading to outcomes that are triggered by activities in interventions and these mechanisms are highly sensitive of contexts (198,199). By only analyzing empirically how intervention activities lead to outcomes, critical realists would state that experimental evaluation relies on statistical association and not a theoretical foundation of causation (200). Bonell and colleagues highlight the implications of this lack of theoretical approach in the following way:

> “*There is concern that RCTs designed primarily to identify whether or not a specific intervention is effective have focused too much on the internal validity of the trial, addressing the question of efficacy rather than broader questions of reach, effectiveness, adoption, implementation and maintenance. This has led to an evidence base that is dominated by high quality RCTs of poorly theorized interventions, with effects that are poorly understood and unlikely to be universally replicated in translation studies or real world implementation*” (200, p2300).

Using an (in)famous analogy from Pawson and Tilley (197) to further illustrate the point of missing causation theories, imagine firing a gun with help of gunpowder. Gunpowder consists of several ingredients and if one or more is missing or if the ingredients are mixed using a wrong recipe, then the gun will not fire. Furthermore, if you forget to pull the trigger, forget to put a bullet in the cartridge or if the gunpowder is wet, then although the recipe for gunpowder was correct, the gun still will not fire or fire blanks. So, the firing of a gun (outcome) is not depending on an “active ingredient”, but rather on the composition and interaction of ingredients (mechanisms), and the circumstances under which these ingredients must operate (context).

In order to explain outcomes, realistic evaluators would therefore focus on identifying change mechanisms operating within interventions instead of focusing on the concrete activities applied and seek to explain or understand how these

mechanisms are influenced by factors in the historical-, social-, institutional- or policy context in which the intervention is embedded. Indeed, some critical realists would even argue that "*context is everything*" when trying to explain outcomes (201). To explain cost-effectiveness of telehealthcare, it is therefore necessary to focus on which mechanisms that is responsible for leading to higher health-related quality of life, lower mortality and costs. It is also necessary to understand the contextual circumstances that (in)activate well-functioning mechanisms. The full cost-effectiveness potential of telehealthcare can only be expected if all mechanisms leading to reduced costs and better outcomes were activated and if the context in which these mechanisms operated all worked and functioned as "a coherent whole" (202).

Having a control group is the reaction to the problem of explaining "context" in experimental evaluation, i.e. that it is possible to strip away other explanations that could contribute to explaining outcomes. But realistic evaluators would stress that contexts cannot be "*bracketed out*" in this way and controlled for:

> "*Researchers who wish to understand how improvement works, and why and when it fails, will never succeed if they regard context as experimental noise and the control of context as a useful design principle*" (201).

Another criticism, is that experimental evaluation does not contain information that come close enough to the actual mechanisms of change and the contexts in which they operate in order to explain what works for whom under what circumstances (197). According to Van Belle and colleagues, the defining nature of mechanisms is that they are "*a response to resources*" offered (e.g. information, skills, materials, support) (198). In other words, what is relevant in order to understand outcomes is how participants react to intervention activities (i.e. changed their logic, values, attitudes, actions) (198).

This means that explanations for outcomes such as cost-effectiveness are much more subtle than can be elicited from data that are usually included in experimental evaluation. Patients are not just defined by socio-demographics, disease severity or lifestyle, but have values, concerns and aspirations as well as cognitive and emotional ways of reasoning, which are not characteristics *per se*, but rather a mindset that could also depend on the actions of health professionals or the institutional context in which they all operate. Likewise, the outcomes of concrete intervention activities are not just dependent on characteristics such as *if* the activities were performed and their "dosage", but concrete events and experiences that these activities lead to, such as *how* activities were introduced, performed and adapted. And health care systems are not just defined by characteristics e.g. their size, available funding or the number of health professionals employed, but also if

organizational cultures, incentive systems or management were supportive and enabling or not (202).

The implication of a missing thorough theoretical understanding of mechanisms and associated contexts in experimental evaluation can only result in a body of empirical knowledge that is filled with mixed evidence (which is exactly the case with experimental telehealthcare research today):

> "*The overall result is a multitude of trial teams […] testing various intervention packages of varying degrees of similarity and difference. Some of these interventions are reported as effective, some as ineffective and a few as harmful, and it is generally difficult if not impossible for either primary RCTs or systematic reviews to make firm conclusions about which intervention components are likely to have the most potential and which combinations of these will produce the greatest effects*" (200, p2301)

Critical realists will argue that "prediction" of outcomes *is* possible, but in order to understand what works for whom, why and under what circumstances, critical realists will seek to build and refine theory over time and increase confidence in predictions by accumulating lessons across different contexts (200). A process which they call continuously refining "context-mechanism-outcome (CMO) configurations" (197). What is initially needed is a literature-driven theory that incorporates contexts, mechanisms and outcomes. From this theory; several hypotheses could be made and analyzed empirically that explicitly relates CMOs together. These are not quantitative hypotheses to be tested; in fact the realistic evaluation framework would rely on methodological pluralism, i.e. a combination of qualitative and quantitative data to investigate empirically how mechanisms interact with contexts in order to produce outcomes. Finally, empirical investigations undergo a process of induction, where "findings" are not generalized to populations, but rather increasingly nuanced, specified and published in order to refine theories of causation that can be used as predictions across other interventions. This last step is perhaps the most distinguishing feature of realistic evaluation.

To illustrate the evaluation approach applied in realistic evaluation and especially the process of refinement, consider an example of property marking. This example is mentioned by Pawson and Tilley (197, p89f), but I have added several nuances in order to concretize how mechanisms or context are continually refined (the refined part of the example is my own description). Property marking is chosen, because it is a very simple intervention with a single very specific outcome: Its purpose is to reduce the burglary rate in neighborhoods and in principle the intervention involves only two activities; marking certain goods in households and putting a sticker warning potential offenders on the front door or windows.

The empirical evidence for property marking has been mixed: Sometimes a reduced burglary rate is found, sometimes not (197). To come closer to an explanation of this mixed evidence, a theoretical basis is initially needed for understanding what mechanisms are affected by property marking that causes burglary rates to decline. Crime opportunity theory suggests that an offence is much more likely if an offender is motivated by *perceiving* that there are *suitable targets* (mechanism) and an *absence of capable guardians* (mechanism) (197). A property marking intervention might be hypothesized to reduce burglary rates by affecting how burglars perceive both targets and guardians: It could make participating households less attractive for burglars, because it signals that stolen goods from this household are more difficult for burglars to dispose of (deactivating mechanism: suitable targets). There could also be an increased potential for detection and prosecution of offenders by authorities if they are found in possession of stolen goods (activating mechanism: a capable guardian). These mechanism might be more likely to be (in)activated with property marking if the number of active offenders is small (context), since deterring those few offenders would have a drastic effect on burglary rates.

Empirically investigating these mechanisms and contexts might result in a need for refinement when they are implemented in reality or implemented in a different setting (this is my own made-up illustration). The “capable guardian” mechanism might only be activated if burglars find a higher risk of getting caught and/or prosecuted *persuasive enough* (refined mechanism). Activating this mechanism can result from any number of events and vary in effectiveness from burglar to burglar. It might be more likely to be activated in a context where domestic burglaries are *prioritized and receive publicity* (some burglar perceptions might be affected by newspaper articles demonstrating a case where property marking did lead to apprehension and convictions, some burglars could be deterred by noticing that extra funding to burglar task forces have been prioritized etc.). These events are created by the conditions set by context and are in most cases not part of the intervention, but instead unpredictable, emerging factors affecting the activation of this particular property marking mechanism. If activated, burglary rates might decline, but only if it is aligned with other mechanisms and only due to reduced burglaries in households with property marking (same outcome). Similarly, the “suitable targets” mechanism might only be deactivated if offenders are *convinced* that all activities in property marking are actually being conducted, i.e. that they have not found examples where households are just putting stickers on the windows and not actually marking any goods (refined mechanism). However, the mechanism could be activated again over time if burglars adapt, e.g. if there are available cash that cannot be marked but targeted instead (refined mechanism). Finally, the number of potential offenders might be small, but if offenders are not local opportunists, but foreign professionals, then they are much more likely to have effective distribution networks for disposing goods outside the jurisdictive of the police and they are less likely to get caught (refined context).

The example is meant to illustrate that what distinguishes realistic evaluation from experimental evaluation is that it is not the simple activity "property marking" that reduces burglary rates. Rather it is mechanisms triggered by the intervention that interacts with particular contexts or contextual events. These mechanisms can be activated or deactivated in many, many ways, some of which would be beyond the control of intervention designers. What characterizes the effects of interventions is therefore that they depend on factors that are emergent and unpredictable and that they could lead to non-linear outcomes (e.g. outcomes may diminish over time). In order to predict outcomes, increasingly refined CMO configurations must therefore be developed, a process that in principle can continue indefinitely, but at some point a pattern of regularity in mechanisms over different contexts will emerge leading to a much higher confidence in the prediction of outcomes (197).

But it is important to understand that since there can be many underlying mechanisms, whose refinement, relationship and dominance may be different from context to context, there will *never* be a deterministic relationship between intervention activities and desired outcomes, (202). Consequently, the regularity pattern are sometimes described as a "partial regularity" (202) meaning that there will always be exceptions to the rule in replicating the outcomes of previous experiences. So, when the Cochrane review on telehealthcare for COPD patients describe the search for why telehealthcare works in the following way: "*Typically, in COPD, there are a number of ingredients including some education, some assisted planning, emotional support, pragmatic advice, monitoring with equipment, etc. which, taken together, may or may not benefit the patient*" (26, p3), realists would claim that there is no point in searching for those "active ingredients" in interventions. Instead, if intended outcomes were not achieved, it is always possible to learn lessons that can be used to develop theories that can help future designs of interventions and evaluations. It is therefore paramount to investigate underlying reasons for activating/deactivating mechanisms and making those explanations publically available. This task is challenging in critical realism; it is illustrated that what was initially a simple intervention with a single outcome, can quickly become so nuanced that it becomes challenging to report. Imagine then the difficulties in seeking to explain the results from a cost-effectiveness evaluation of telehealthcare that is a much more complex intervention with several interacting activities and three outcomes (mortality, health-related quality of life and costs).

Critical realists have been criticized for not providing an operational methodology for the conduct of realistic evaluation (203,204) and evaluators have therefore struggled to implement realistic evaluation in practice (204). This implicates that it is generally difficult to empirically test applied realist theories, which risk bringing realistic evaluation into conflict with principles of scientific investigation, where theories should be falsifiable (205). Although some improvements have been made recently (199,206), realistic evaluation has been criticized for still being unclear in

its definition and operationalization of contexts, mechanisms and outcomes and especially in describing how they are distinguished (204).

6.3. CURRENT MEDICAL RESEARCH COUNCIL GUIDANCE FOR EXPERIMENTAL EVALUATION OF COMPLEX INTERVENTIONS

While adhering to the positivistic paradigm, The Medical Research Council (MRC) have published guidance of conducting evaluations of complex interventions that insists that experimental evaluation should be at the core of evaluation in order to maintain internal validity of the cause-and-effect conclusions drawn (207). It would otherwise be difficult to separate observed outcomes from what would have happened anyway (208). But at the same time, it is acknowledged that it is challenging to replicate findings from individual studies and theories and qualitative research is therefore necessary in order to provide lessons for future implementation (207).

In the MRC framework, the term “complex” is used to describe interventions that include several interacting components consisting of a large number of activities and a high frequency of interaction between these components (207). Furthermore, interventions are complex when several groups or organizational levels are involved and when potential relevant outcomes is large or the required behaviors of those targeted by the intervention is difficult (207). In this view, telehealthcare is complex, since it requires changed behaviors from both patients and healthcare professionals and involves not only a device with measurement peripherals, but also some training and a changed work flows and organization in and between healthcare delivery sectors. This complexity of telehealthcare solutions are acknowledged in the Cochrane review of telehealthcare to COPD patients (26) and by some health economists (209).

In order to develop and evaluate complex interventions, a systematic framework is recommended by MRC that consist of four more or less sequential phases (207). First, a *development phase*, where theory and existing evidence relevant for the intervention is explored that should allow investigators to formulate a logical model and hypotheses of intervention outcomes and to identify mediating or moderating variables. This could be based on traditional systematic reviews of the literature and other theories. Qualitative methods could assist in order to describe the inputs that the intervention requires, how the components fit together and to identify underlying processes that produces outcomes (207). Second, a *pilot and feasibility phase* is necessary to test the acceptability and applicability of the intervention activities and evaluation approach that have been created so far. This may take the forms of small scale models of the actual intended intervention and evaluation design or as designs that investigates the main uncertainties with implementation or

evaluation identified in the development phase. Examples of tests include estimating recruitment, retention, compliance, delivery and the relevance of outcomes. Qualitative methods could assist in addressing potential barriers or enablers affecting those aspects. It should also point to intervention components that are essential and which can be added, adapted or eliminated. Depending on the results of these studies further refinement of intervention components or evaluation are needed (207). Third, *full-scale evaluation* should be adopted in order to assess effects. Randomized designs must always be considered with large samples. Multilevel designs or some other variant of the basic RCT design are preferred to account for technical or ethical challenges with randomization (207). Process evaluations nested in the trial are recommended in order to explain differences in expected and actual outcomes as well as to identify facilitators or barriers to implementation into routine practice (208). A nested health economic evaluation should also be included to make the trial "more relevant" for decision-making and to assist in getting the evidence translated into practice (207). Finally, an *implementation phase* concludes the evaluation effort and includes scientific publication of results and descriptions of experiences that contain more accessible information and specific recommendations for deployment (207).

The main strength of the MRC framework is that it underlines that *evaluation is a process, not an approach.* This could in theory solve some of the challenges faced in trial-based economic evaluation of complex interventions by theoretically allowing refinement of intervention and identification of important factors that could potentially impact evaluation results *before* full-scale deployment of large and complicated initiatives and evaluations such as CRTs. It would allow *some* organizations, healthcare professionals and patients to become more experienced with the implementation of the intervention (accounting for some learning curve effects) and it could point investigators to potential confounding factors, wider organizational implications of implementation or how intervention activities are locally adapted.

Many of the MRC recommendations can be found in the TeleCare North evaluation and to a lesser extent in the KIH initiatives. The development of the intervention was largely decoupled from the evaluation in both TeleCare North and KIH. A theory of change and work flow analyses was applied in the KIH initiatives (115), but was largely neglected in TeleCare North. Instead, TeleCare North built on experiences gained from especially a pilot trial with nested economic evaluation in North Denmark Region (109,172) that demonstrated a potential for cost-effectiveness after 6 months measured as cost per QALY. In a sense, some would argue that this study functioned as both the underlying theory and the pilot study of the TeleCare North trial. This pilot study involved a non-commercial version of a tablet and peripherals and a hospital-to-home intervention and therefore required a refined technical solution and organization in order to make use of newer and more available technology and the cross-sectorial requirements. In seeking to update the

intervention, much attention was given to the feasibility of cross-sectorial implementation and too little attention to working with mechanisms and contexts thought to lead to effectiveness and cost savings. The pilot study, although probably cost-effective due to cost-savings, demonstrated statistically insignificant effectiveness (incremental QALY=0.027) and did not include cost to municipalities. A more conscious approach to improving and testing particularly the effectiveness outcomes of the trial and the impact on municipality cost could therefore have been warranted prior to large-scale evaluation. In addition, the pilot study was implemented in a different organizational setting than TeleCare North eventually did. Working with how this changed context could influence the outcomes of the trial might have helped in explaining the ultimate outcomes of the trial. But none of the initiatives put underlying theories to the test before large-scale implementation and evaluation. TeleCare North did move forward to a large-scale multilevel clinical trial with nested economic evaluation that was accompanied by a qualitative process evaluation of facilitators and barriers to implementation. This was in reality not done in KIH (even though it was called a large-scale initiative, it actually contained five pilot studies). The KIH initiatives contained some process evaluation as well but not a trial-based economic evaluation. TeleCare North was also the only initiative to be published scientifically. It is always easier to be smart in hindsight, but in the spirit of including what we know now in order to inform future evaluations, it could be argued that although many of the recommendations from the MRC framework was incorporated into the design and evaluation process of TeleCare North, we might still have moved too quickly to large-scale implementation and testing.

The MRC framework have been criticized for being too embedded in a clinical logic, for not engaging with complexity science and for having a tendency to oversimplify complex social reality (202,210,211). MRC presents a linear logic from a pre-specified program theory to large-scale implementation, which largely ignores the role of particularly context and that CMO-configurations needs to be refined before evaluation (210,211). Too little attention is therefore paid to its content in especially the development and feasibility phase (211). The largest role theory of context and mechanisms have in the MRC framework is in process evaluation that is suggested in the implementation phase in order to investigate reasons for discrepancies between trial outcomes and expectations (208). But process evaluations are not concerned with causality, but instead how interventions were implemented in practice and how outcomes were perceived by patients, healthcare professionals and other stakeholders (208). There is generally no feedback that can be used to quantify the effects that this implementation or these perceptions might have had on the outcomes of the evaluation. And then it becomes a little too late. So, the MRC framework is criticized for including the right topics or words, but without providing any real content of how to adapt an experimental evaluation setting to include context and mechanisms (202,210,211).

6.4. A REALIST SKETCH FOR HEALTH ECONOMIC EVALUATION OF COMPLEX TELEHEALTHCARE INTERVENTIONS

Some would rightfully state that it is impossible to have a standpoint in two different philosophies of science, meaning that experimental evaluation in general and trial-based economic evaluation by implication, cannot be truly realistic (198). However, albeit a difficult and daunting task, it does not exclude that lessons may be taken from one standpoint to inspire the development of the other. But how can a more theory-based evaluation approach be incorporated into trial-based economic evaluation that would maintain high internal validity while at the same time allow for greater external validity in order to keep answering what course of action is cost-effective and for whom, but also seek to focus more on why and under what circumstances telehealthcare is cost-effective?

There have been some developments into ways of incorporating more concrete realist thinking into experimental evaluation, e.g. realist randomized trials (200) and recently realist complex intervention science (211). But these more theory-based approaches to evaluation tend to be focused on explaining effectiveness (196). However, it should not be a controversial idea to seek starting a research agenda for realist health economic evaluation; the fact that health economists insist that pragmatic trials are better vehicles than explanatory trials for economic evaluations and that a lot of effort are put into sensitivity analyses (i.e. a right shift on an internal to external validity continuum) is a manifestation of an underlying logic that very few context-independent and deterministic relationships between interventions and outcomes can be found and that changed prerequisites and assumptions (i.e. changed theory) impact outcomes. This leads to the first principle.

> *Principle one: Trial-based economic evaluation of complex telehealthcare interventions should focus on external validity by developing and testing theory of what causes change mechanisms to be activated/deactivated in order to produce certain outcomes and incorporate how these are dependent of context.*

An adapted MRC framework for evaluation could be a starting point. Fletcher and colleagues have in July 2016 published an article suggesting several general techniques to include more “realism” into experimental evaluation studies by recommending that all MRC phases should be utilized and adapted (211).

In the *development phase*, some form of mixed-methods reviews can be used to synthesize the overall effects of interventions while also addressing how interventions work and interacts with context (211). Purposively sampled case studies (interviews, observations or surveys) are suggested as a way to explore

hypothesized theoretical mechanisms in different settings in order to understand how context interacts with mechanisms (211). Both activities should result in an initial program theory of change and explicitly describe how context would impact on the activation/deactivation of mechanisms (211).

Below is a suggestion of a program theory meant to demonstrate the potential value of realism for trial-based economic evaluation. Comparative effectiveness in telehealthcare might be achieved by "*adherence*", i.e. that increased effectiveness is mostly feasible if telehealthcare makes it easier for patients to adhere to the treatment and care opportunities relevant for their disease (e.g. monitoring, training, smoking cessation, dieting and pharmaceutical therapy). Another program theory could be "*patient capabilities*", i.e. that increased effectiveness is more achievable if telehealthcare allow new patient skills to be successfully developed in mastering COPD. Potential underlying mechanisms of change could be that these theories are only relevant if *patients* feel comfortable with the technology and are empowered by the opportunities to gain more access to counseling or support in managing their disease or in adhering to monitoring, training and medicine therapy. But it is also important to relate these program theories of effectiveness to program theories for costs. Under *usual care*, it could reasonably be hypothesized that increased severity of COPD leads to increased resource consumption and that there is some form of positive correlation between severity and resource consumption. However, a positive correlation does not exclude that all individual patients follow this trend. One program theory for lower costs could be called "*optimization*", which means that some patients – presumably COPD patients with intermediate severity and risk of events - who are treated by specialized healthcare professionals under usual care, might be safely switched to less specialized care *if* they *adhere* to monitoring, pharmacological therapy and training by help from telehealthcare. Potential underlying mechanisms can be found in the reactions of healthcare providers and personnel involved in telehealthcare and it could be that in order to achieve lower cost by optimization, *engagement* in learning new ways to monitor and caring for patients by healthcare professionals is necessary. Another program theory for cost could be called "*transference*", meaning that *if* increased patient capabilities could be achieved, then it would be possible to reduce the need for healthcare sector resources, because the patient's themselves or their relatives become actively involved in learning to monitor and care for the disease. This would require a mechanism of *trust and acceptance* for transference of responsibility by patients and relatives and possibly convincing safety evidence as judged by health professionals. Lastly, contextual factors hypothesized as important for achieving increased effectiveness and lower costs should be presented. These could be found in areas such as the technology applied, the targeted patients and in the organizational structures in which telehealthcare is implemented. In TeleCare North, we have documented that the severity of COPD might play a role and an under-met need for care in municipalities might affect costs, but generally little explicit knowledge exist. So, an overall program theory for cost-effectiveness of

telehealthcare might be that two intermediate outcomes must be increased first in order to increase health-related quality of life or reduce mortality by telehealthcare, adherence and capabilities. These are also necessary in order for cost-savings to occur. But both effectiveness and cost outcomes are depending on several underlying mechanisms to be (de)activated and the context in which telehealthcare is implemented.

A *feasibility/pilot phase* should be used to explore facilitators or barriers to implementation and the acceptability of participants and stakeholders in order to inform a revision of the intervention design and evaluation methods prior to large scale evaluation (211). However, there has been a tendency in feasibility studies to only explore if an intervention and evaluation is possible and acceptable or not; not what is possible and acceptable for whom and under what circumstances (211). More importantly in this phase; developed program theories should be put more to the test before large-scale evaluation in order to investigate if and how interventions or evaluation should be adapted before large-scale evaluation and to point to potential areas for learning effects and organizational implications (it is too late when evaluation has begun). It is therefore important to design interventions and develop evaluation simultaneously to allow for testing different concrete intervention activities or components (e.g. device and peripherals, training, monitoring, support) that could achieve intermediate outcomes and lead to the activation of mechanisms. This should be done across a range of purposively selected contexts that would allow for refining the underlying program theory (CMO configurations) (211). If it is appropriate, a multi-arm pilot study can be used to assist in identifying components that are essential in activating/deactivating mechanism but also to assess if some components could be adapted or eliminated (211).

Fletcher and colleagues have argued (211) that the *evaluation phase* should utilize developments in realist randomized trials (200), which are large-scale mixed-methods trials, that assess within-trial effects of interventions and how they are mediated by mechanisms and moderated by context as well as developing theories about factors that would enable or prohibit effectiveness in other contexts. Realist trials would usually apply complex designs, such as CRTs or other variants of RCTs and place emphasis on multi-arm or factorial designs in estimating the effectiveness of separate intervention components and in combination (200). Analytically, realist trials would also apply statistical mediation analysis in order to assess the activation/deactivation of mechanisms (200). In doing so, realist trialists would underline that it is not the concrete activities in interventions that are important but how they activate/deactivate mechanisms (200). Interaction with context is explored by heterogeneity analyses and by pre-planned coordination with other trials applying the same program theory between implementation contexts (200). Qualitative methods would be included as process evaluations meant to identify what the intervention did and point to a need for refining mechanisms or

refining how context affected them, but also point to relevant additional variables to be used in the follow-up measurement (200).

To my knowledge, there has only been one initiative that has used a well-founded theoretical approach in telehealthcare research. Perhaps based on the discussion in the aftermath of the results of telehealthcare in the WSD in the UK, initial work began in 2013-2014 on applying theory-based principles to telehealthcare on a new initiative testing the effectiveness and cost-effectiveness of telehealthcare in the UK (the "TECH model" applied in the later implemented Healthlines studies) (212). In developing the conceptual model, a systematic review to assess evidence of effects combined with a realist synthesis of qualitative investigations of patient experiences with telehealthcare a crude program theory was formulated (212). The model was then adapted to their empirical context by a qualitative study tapping into the potential role of telehealthcare valued by both patients and different health professionals and a survey to patients investigating the presence of and interaction of contextual factors on the emergent mechanisms (212). Based on these sources, a draft model was formulated to capture how, why and under what circumstances telehealthcare might bring about desired outcomes and this draft model was refined by a workshop consisting of representatives from both patients, healthcare personnel and providers, commissioners of services and academics which lead to the development of an initial program theory (212).

The results from this process identified three theoretical "components" for effectiveness and cost-effectiveness in telehealthcare (212). One was "self-management" which were hypothesized to be activated by several activities in interventions, e.g. by training behavior change techniques, allowing for self-monitoring of vital values, promoting self-efficacy, shared decision-making, motivational interviewing and personalized support from health professionals etc. (212). The second component was "treatment optimization" that could be activated in several ways, e.g. by risk stratification of patients, treatment intensification, using evidence-based protocols and regular review, promoting medicine adherence etc. (212). The third component was "care coordination" activated by shared records, effective communication between healthcare providers, regular monitoring of system performance etc. In addition, several important contextual factors were hypothesized to influence the effects of these components (e.g. the level of engagement of patients and primary care professionals, type and severity of disease and the wider health system in which interventions are embedded) (212).

The TECH program theory was then used to design two strategically coordinated telehealthcare pragmatic trials with nested 12 months trial-based economic evaluations and process evaluations (213). The chosen participants were patients with depression (609 patients, 307 received telehealthcare) and patients with cardiovascular risk (641 patients, 325 received telehealthcare) (213). The setup consisted of the same participating general practices, intervention- and research

staff and the same underlying theory of components, mechanisms and contexts. But the specific activities in each intervention, the data analyses and reporting were different from study to study (213). In July 2016, the results from the trial-based economic evaluations of both trials were published and provided mixed results (214,215). The cardiovascular trial concluded that it was likely that the Healthlines telehealthcare intervention for cardiovascular risk was cost-effective (net monetary benefit of £116 (95% CI −58 to 291) and a 77% probability of cost-effectiveness at threshold of £20,000 per QALY) (214). However, the depression trial concluded that the Healthlines telehealthcare intervention was unlikely to be cost-effective (net monetary benefit of -£143 (95% CI −£164 to -£122) and a 30% probability of cost-effectiveness at threshold of £20,000 per QALY). This mixed evidence leads to the second principle.

> *Principle two: Trial-based economic evaluation of complex telehealthcare interventions should be part of a continuous strategic process of knowledge creation that should ultimately answer questions of what was cost-effective for whom, why and under what circumstances. Since only partial regularities can be found, it is not an approach designed to be a one-off verdict of what was cost-effective or not.*

Traditional health economists would be disappointed by the results from Healthlines and might even claim that there is no need to adopt realistic evaluation methods in health economic evaluation, since evidence is still mixed. But is that fruitful? We have done this for 20 years without much progress (28). Not surprisingly, a critical realist would argue that this mixed evidence is exactly why we need realistic evaluation and that disbanding a theory-based approach is premature. Researchers and practitioners always learn something - both from successes and failures - that can be used in order to inform the theoretical logic underlying future studies. The journey to theory validation and refinement has only just begun. And it is also worth remembering that critical realists would argue that the world is so complex that unforeseen events always arise that affects outcomes in some way. But what realistic evaluation does, is that it focuses attention to working with contexts, mechanisms and outcomes and to making these experiences explicit and available for other researchers.

The published process evaluation points to some indications of the need for refinement in implementation in the Healthlines trials (216). In general, the overall theory is concluded as sound for developing interventions and for assessing effects and implementation occurred largely as planned. However, there were several contextual factors that could have affected effectiveness in both trials. In the depression trial, there were problems with engaging patients because they had busy lives and preferred a therapeutic approach to behavioral therapy and in the CVD

trial, patients were not really interested in making lifestyle changes, but joined the trial for altruistic reasons. Finally, primary care was not engaged either (216).

There are several considerations for future theory-based interventions and evaluation to be made based on the experiences with the TECH model and Healthlines studies that relates to context, mechanisms and outcomes. First, more work should be done in distinguishing between what is underlying mechanisms of change, context and what are outcomes, e.g. in the Healthlines studies it is not completely clear if self-management is a response to resources offered (a mechanism) or an intermediate outcome leading to more ultimate outcomes. It is also not clear if patient engagement is a mechanism or a context (it is described as context, but could also be a response to the intervention?). Second, more work on linking mechanisms and context to costs should be conducted and to describe how they are linked with effectiveness mechanisms. Third, it could also be time to ask ourselves again if the effectiveness part of cost-utility analyses (i.e. QALYs) is the most relevant outcome in its existing form in a hierarchy of outcomes for telehealthcare evaluation. Looking across recent trial-based economic evaluations of telehealthcare, a regularity has arisen, i.e. the incremental QALYs are still small and statistically insignificant (incremental QALYs in WSD (0.012, not significant) (125), Telescot (0.017, not significant) (127), TeleCare North (0.013, not significant) (129) and the two Healthlines studies (0.012, not significant (214); and 0.031, not significant) (215). This keeps being consistent with findings from Bergmo that published similar results from a review of QALY gains in telehealthcare research prior to the WSD publications (217). I do not recommend stopping the effectiveness part of economic evaluation, because we need to assess whether the intervention is at least as good or safe as comparative interventions. But as long as they are, and this seems to be case, the largest effects on outcomes might be found in minimizing costs, effectively reducing the economic evaluation to a cost-minimization analysis. It could also implicate that we need to be prepared to find additional effectiveness outcomes in trial-based economic evaluation, such as better access to treatment and care or more consistent care, even if this would reduce economic evaluations to cost-effectiveness analyses (although utility could also be access to treatment and care, but will require quantification in order to be applied for comparison with other technologies with the same objective). Bergmo have also previously suggested that benefits of telehealthcare instead of QALYs might be found in other outcomes such as access or waiting time (217). Fourth, although much effort was put into the development of the theoretical model for the Healthlines trials, it seems that much more attention could be made in testing theory and interventions before large-scale evaluation. That two GPs were used in “a run-in period” is mentioned in the trial protocol (213), but it seems that this was to test feasibility of implementation rather than testing and refining theory of mechanisms, context and outcomes. More systematic work on testing could be warranted before further evaluation in order to explicate or refine cost and effectiveness mechanisms

leading to outcomes and how context would affect their relationship as described above. This leads to the third principle.

> *Principle three: Trial-based economic evaluation of complex telehealthcare interventions should have an organization for testing CMO-configurations which would allow for climbing up the learning curve before large-scale evaluation.*

In general, mixed methods are recommended as part of getting more realism into experimental evaluation and cross-disciplinary research networks with researchers from multiple disciplines and practitioners is a requirement for good intervention science (200). But this needs an organization. I would propose that it could probably be found in the quality improvement (QI) tradition that is deeply embedded in practice. On the conceptual level, the QI tradition is especially useful for trial-based economic evaluation, because it is inherently realistic in its evaluation approach and terminology and because there is a particular focus on documenting learning. QI seeks to build knowledge of how to make improvements (218). This could be achieved by changing the characteristics of a product or service, its delivery process or the entire system to bring about predicted improvements. Central to QI is two things: First, predictions of improvements are based on an explicit theory and hypotheses of change. Second, sequential planned experimentations are used to test theories in practice. If changes did not lead to predicted improvements in the experiments, it is necessary to identify and describe the context that affected change in order to refine the theory (QI would identify those based on anomalies or variation in outcomes, so mixed-evidence is expected!). A framework to support this knowledge creation is the "Model for improvement", which applies both deductive and inductive learning phases in so-called Plan-Do-Study-Act (PDSA) cycles. In the "Plan" phase, a theory of why a concrete intervention would lead to improvement is developed and an experiment to test the intervention and theory is designed. In the "Do" phase, the test is conducted and any unexpected observations or problems are documented. The "Study" phase is concerned with analyzing data where a comparison with predictions are made and described. In the "Act" phase, refinements are made to the intervention and/or theory. By repeated use of experimentation by PDSA cycles, knowledge of what works, why and under what circumstances is increased and more confident predictions can be made. QI uses small scale and local experiments in order to develop and study interventions that should bring about improvements and in a MRC optic, the QI tradition is therefore most relevant in the intervention development and feasibility/pilot phase (218). The QI tradition could therefore be precisely what trial-based economic evaluation needs in order to bridge the gap between theory development and large-scale implementation and evaluation.

Many Danish regions has a huge focus on the model for improvement and have organizational structures that support the initiation, testing and dissemination of

quality improvement efforts across hospitals, e.g. DEFACTUM in Central Denmark Region (www.defactum.dk), the improvement organization in North Denmark Region(http://www.rn.dk/Om-Region-Nordjylland/Organisationsbeskrivelse/Forbedring)and Centre for Quality in the Region of Southern Denmark (www.centerforkvalitet.dk). But in Denmark, it is almost an embedded culture that the QI movement and comparative outcomes research are separate disciplines, where comparative outcomes researchers may have a tendency to be skeptical of the concept of "quality" in QI that is generally defined as fulfilling customer needs and see their interventions as quick fixes rather than sound interventions based on rigorous evaluation methods. QI practitioners on the other hand could advocate that comparative outcome researchers are not aware of real practice and oversimplify reality in laboratories. In order to utilize this organization, leadership is a prerequisite from both universities and regions in order to bring academics and practitioners together to try something different. On one hand, QI practitioners should be willing to accept that there is a hierarchy of evidence and that experimental trials are at the heart of this hierarchy. Equally important, it also needs a larger community of health economists that have had the frustrating experience of being challenged to explain the outcomes of trial-based economic evaluation. This leads to principle four.

> *Principle four: Trial-based economic evaluation of complex telehealthcare interventions should challenge existing general good practices of how to conduct and report trial-based economic evaluation.*

In order to be more able to develop, validate and refine CMO configurations, some changes to how trial-based economic evaluation should or should not be conducted and reported is necessary. One is tackling the fear of "data-dredging" that could lead to false positive results. Clinical trials are usually powered to detect changes in a single primary (usually efficacy or effectiveness) outcome which mean that it is hard to find statistically significant results in secondary analyses such as in secondary outcomes, in mediation analyses or in subgroups. This challenge should be taken seriously especially if studies are meant to inform clinical practice. However, it is already acknowledged that the purpose of trial-based economic evaluation is different because they are meant to inform funding decisions. Those decisions do not have to be dependent on statistical significant certainty and the only requirement is that any secondary analyses are pre-specified in a study protocol. But in complex telehealthcare interventions, it would only be feasible to pre-specify mechanisms and contexts that are included in the program theory. Central to realism is that unanticipated events would always occur that affects outcomes and those events are difficult to foresee a priori. Reporting how these unforeseen events affected mechanisms and cost-effectiveness is central to the refinement of theories that can be used in future design of interventions and evaluation, not only locally but also for researchers conducting mixed-methods systematic reviews. So, instead of automatically rejecting post hoc secondary

analyses, good practices for trial-based economic evaluation of complex telehealthcare interventions should actually recommend them and include them in reporting guidelines for trial-based economic evaluation (e.g. as extensions to the CHEERS-checklist).

Another challenge is the fear of comparing complex telehealthcare interventions that involves different concrete intervention activities and for changing concrete interventions to suit the local context in which it must operate when evaluation has begun. The strength of focusing on mechanisms is that it is not the concrete activities that directly affect outcomes, but underlying causal mechanisms. So, as long as the underlying logic or program theory is the same, comparing interventions with the same program theories should be possible and there should be more room for local adaptations in concrete intervention activities.

The above was meant as a sketch of some normative overall principles for trial-based economic evaluation of complex telehealthcare interventions that should be discussed and developed further. It is meant as a suggestion for how to move forward and is not a finished concrete concept to be tested in practice and how can it be when sketched by a single author new to the field without existing research or examples to build on.

Although, I have suggested possible ways to adhere to the suggested principles that can be built on very recently published evaluation research and described how existing organization in particularly the Danish regions can be strategically utilized, it is still a huge investment in time and effort to change existing trial-based economic evaluation of complex telehealthcare interventions. I would therefore like to finish off by giving voice to Bonell and colleagues who were one of the first research groups that suggested an integration of experimental and realistic evaluation in 2012 that addresses this issue in the following way:

> *"It might be argued that our plans are impossible because of the expense and time required for such painstaking assessment […]. we argue that this systematic approach would ultimately be more time- and cost-efficient than current, uncoordinated efforts. There are already huge numbers of evaluation studies conducted that focus on an unsystematic array of combinations of intervention components and settings, and which, for the reasons outlined above, often provide a poor basis for decision-making by policy-makers and practitioners"*(200, p2305)

CHAPTER 7. LITERATURE LIST

1. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;386(9995):743–800.

2. World Health Organization. The world health report 2002 - reducing risks, promoting healthy life. Vol. 16. 2003.

3. Centers for Disease Control and Prevention. Chronic Disease Overview [Internet]. Chronic Disease Prevention and Health Promotion. 2016. Available from: http://www.cdc.gov/chronicdisease/overview/

4. Lau, CJ; Lykke, M; Andreasen, AH; Bekker-Jeppesen, M; Buhelt, LP; Robinson, KM; Glümer C. Health Profile 2013 – Chronic Disease [Internet]. 2013. Available from: https://www.regionh.dk/fcfs/publikationer/Documents/Sundhedsprofil 2013 - Kronisk sygdom.pdf\nhttps://www.regionh.dk/fcfs/sundhedsfremme-og-forebyggelse/Sider/Sundhedsprofil.aspx

5. Danish Health Authority. Kronisk sygdom [Internet]. Overblik over kronisk sygdom. 2014. Available from: https://sundhedsstyrelsen.dk/da/sygdom-og-behandling/kronisk-sygdom

6. Halbert RJ, Natoli JL, Gano A, Badamgarav E, Buist AS, Mannino DM. Global burden of COPD: systematic review and meta-analysis. Eur Respir J [Internet]. 2006;28(3):523–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16611654

7. Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. Lancet [Internet]. 2007;370(9589):765–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17765526

8. WHO. Chronic obstructive pulmonary disease (COPD) [Internet]. Fact sheet 315. 2012. Available from: http://www.who.int/mediacentre/factsheets/fs315/en/index.html

9. European COPD Coalition. Prevalence of COPD in EU [Internet]. 2014. Available from: http://www.copdcoalition.eu/about-copd/prevalence

10. Danish Lung Association. COPD in Denmark [Internet]. 2012. Available from: www.lunge.dk

11. Hansen JG, Pedersen L, Overvad K, Omland Ø, Jensen HK, Sørensen HT. The Prevalence of chronic obstructive pulmonary disease among Danes

aged 45-84 years: population-based study. COPD [Internet]. 2008;5(6):347–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19353348

12. Flachs EM, Eriksen L, Koch MB. The disease burden in Denmark. 2015.

13. National Institute of Public Health. The Public Health Report, Denmark. 2007.

14. Wouters M. Economic analysis of the Confronting survey : an overview of results. Respir Med. 2003;97(Supplement C):S3-14.

15. OECD. Quarterly National Accounts. OECD Stats; 2016.

16. Danish Ministry of Finance. Economic Survey, December 2011 [Internet]. 2011. Available from: http://sim.dk/nyheder/nyhedsarkiv/2011/dec/oekonomisk-redegoerelse-december-2011.aspx

17. OECD. OECD Economic Surveys [Internet]. OECD Economic Surveys: Denmark 2012. Available from: http://www.oecd-ilibrary.org/oecd/content/book/eco_surveys-nzl-2005-en

18. Pedersen KM, Bech M, Karsten V. The Copenhagen Consensus Report. 2011.

19. OECD. Health Care Systems : Getting More Value for Money [Internet]. Vol. 2, OECD Economics Department Policy Notes. 2010. Available from: http://www.oecd.org/economy/productivityandlongtermgrowth/46508904.pdf

20. The Prime Ministers Office. A Denmark who stands together [Internet]. Government Platform 2011. Available from: http://www.stm.dk/publikationer/Et_Danmark_der_staar_sammen_11/Regeringsgrundlag_okt_2011.pdf

21. Pedersen KM, Bech M, Karsten V. The Danish Health Care System: An Analysis of Strengths, Weaknesses, Opportunities and Threats The Consensus Report. Copenhagen Consensus Center. 2008. 1-83 p.

22. Dinesen B, Haesum LK, Soerensen N, Nielsen C, Grann O, Hejlesen O, et al. Using preventive home monitoring to reduce hospital admission rates and reduce costs: A case study of telehealth among chronic obstructive pulmonary disease patients. J Telemed Telecare [Internet]. 2012;18(4):221–5. Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L364984101

23. Schou L, Østergaard B, Rydahl-hansen S, Rasmussen LS, Emme C,

Jakobsen AS, et al. Randomised trial of telemedicine-based treatment versus conventional hospitalisation in patients with severe COPD and exacerbation – effect on self-reported outcome. J Telemed Telecare. 2013;19(3):160–5.

24. Sorknaes AD, Bech M, Madsen H, Titlestad IL, Hounsgaard L, Hansen-Nord M, et al. The effect of real-time teleconsultations between hospital-based nurses and patients with severe COPD discharged after an exacerbation. J Telemed Telecare [Internet]. 2013;19(8):466–74. Available from: http://jtt.sagepub.com/lookup/doi/10.1177/1357633X13512067

25. Bøg MF, Christensen JB, Jensen MK, Christensen JB. Erfaringsopsamling i forbindelse med Udbredelse af Telemedicinsk Hjemmemonitorering. Odense; 2015.

26. McLean, S; Nurmatov, U; Liu, JLY; Pagliari, C; Car, J; Sheikh A. Telehealthcare for chronic obstructive pulmonary disease. Cochrane Database Syst Rev [Internet]. 2011;(7):1–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21735417

27. Polisena J. Meta-analysis on COPD. J Telemed Telecare [Internet]. 2012;18(4):242. Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L364984107

28. Wootton R. Twenty years of telemedicine in chronic disease management--an evidence synthesis. J Telemed Telecare [Internet]. 2012;18(4):211–20. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3366107&tool=pmcentrez&rendertype=abstract

29. Wainwright C, Wootton R. A review of telemedicine and asthma. Dis Manag Heal Outcomes [Internet]. 2003;11(9):557–63. Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L37159607

30. Polisena J, Coyle D, Coyle K, McGill S. Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations. Int J Technol Assess Heal Care. 2009;25:339–49.

31. Pare G, Jaana M, Sicotte C. Systematic Review of Home Telemonitoring for Chronic Diseases: The Evidence Base. J Am Med Informatics Assoc. 2007;14(3):269–77.

32. Tran K, Polisena J, Coyle D, Coyle K, Eh K, Cimon K, et al. Home telehealth for chronic disease management. 2008.

33. The Danish Agency for Digitilisation. The National Action Plan for Dissemination of Telemedicine [Internet]. 2012. Available from:

http://www.digst.dk/Servicemenu/English/Policy-and-Strategy/Strategy-for-Digital-Welfare/Telemedicine

34. National Institute for Health and Care Excellence. NICE technology appraisal [Internet]. 2016. Available from: https://www.nice.org.uk/About/What-we-do/Our-Programmes/NICE-guidance/NICE-technology-appraisal-guidance

35. Dillon A. Presentation by NICE on danish conference on prioritisation. Danish Regions; 2016.

36. Sørensen L, Nielsen J. Tanketom styring af danskernes sundhed : nedskæringer og teknologisk udvikling betyder voksende behov for medicinsk teknologivurdering : samtidig stopper Sundhedsstyrelsen sin indsats på området på ubestemt tid. Ingeniøren [Internet]. 1989; Available from: http://ing.dk/artikel/tanketom-styring-af-6249

37. Danish Health Authority. Yearly report on health technology assessment. 2008.

38. Alban A, Keiding H, Søgaard J. Retningslinjer for Samfundsøkonomiske analyser af lægemidler [Internet]. 1998. p. 1–8. Available from: https://sundhedsstyrelsen.dk/da/medicin/tilskud/generelle-tilskud/ansoegning/~/media/A798FDB668DD4BBEAC539AB2AE22025C.ashx\nhttps://sundhedsstyrelsen.dk/da/medicin/tilskud/generelle-tilskud/ansoegning/sundhedsoekonomiske-analyser

39. Danske Regioner. Koordineringsrådet for ibrugtagning af sygehusmedicin. 2016.

40. Danske Regioner. Rådet for anvendelse af dyr sygehusmedicin [Internet]. 2016. Available from: http://www.regioner.dk/radsdk

41. Danske Regioner. Model for vurdering af lægemidler. 2016;(16). Available from: http://www.regioner.dk/media/1287/model-for-vurdering-af-laegemidler.pdf (23-07-2016)

42. Danmarks lungeforening. Patienthistorier om KOL [Internet]. Patienter fortæller. 2015. Available from: https://www.lunge.dk/patienthistorier-om-kol

43. Reersted T. KOL og følelserne [Internet]. 2015. Available from: http://min.medicin.dk/Artikler/Artikel/69

44. Roisen, RR; Vestbo J. Global initiative for Chronic Obstructive Lung disease – Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease [Internet]. 2013. Available from: http://www.goldcopd.org/uploads/users/files/GOLD_Report_2013_Feb20.pdf

45. Pena VS, Miravitlles M, Gabriel R, Jimenez-Ruiz CA, Villasante C, Masa JF, et al. Geographic variations in prevalence and underdiagnosis of COPD: Results of the IBERPOC multicentre epidemiological study. Chest. 2000;118(4):981–9.

46. McCance K, Huether S, Brashers V, Rote N. Pathophysiology. The Biologic Basis for Disease in Adults and Children. 6th editio. Maryland Heights: Mosby Elsevier; 2010.

47. Kessler R, Partridge MR, Miravitlles M, Cazzola M, Vogelmeiere C, Leynaud D, et al. Symptom variability in patients with severe COPD: A pan-European crosssectional study. Eur Respir J. 2011;37(2):264–72.

48. Wedzicha J a, Donaldson GC. Exacerbations of chronic obstructive pulmonary disease. Respir Care. 2003;48(12):1204-1213-1215.

49. Barnes PJ, Celli BR. Systemic manifestations and comorbidities of COPD. Eur Respir J [Internet]. 2009;33(5):1165–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19407051

50. Smith MC, Wrobel JP. Epidemiology and clinical impact of major comorbidities in patients with COPD. Int J Chron Obs Pulmon Dis. 2014;9:871–88.

51. Rodriguez-Roisin R, Vestbo J. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease (Updated 2013). 2013.

52. Stenton C. The MRC breathlessness scale. Occup Med (Chic Ill). 2008;58:226–7.

53. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. Eur Respir J. 2009;34(3):648–54.

54. Christiansen T. Organization and financing of the Danish health care system. Health Policy (New York). 2002;59(2):107–18.

55. Olejaz M, Juul Nielsen A, Rujkjobing A, Okkels Birk H, Krasnik A, Hernández-Quevedo C. Denmark: Health system review. Health Syst Transit. 2012;14(2):155–76.

56. Pedersen KM. Sundhedsøkonomi. 1st ed. København: Munksgaard; 2013.

57. The Prime Ministers Office. Healthy throughout life [Internet]. 2002. Available from: http://forskningsbasen.deff.dk/Share.external?sp=S13c27ce0-ac5d-11da-bee9-02004c4f4f50&sp=Sau

58. Danish Health Authority. Forstærket indsats for patienter med kronisk

sygdom. 2010.

59. Danish Health Authority. Perspektiver for den forstærkede indsats for patienter med kronisk sygdom. 2014.

60. Blands J BL. COPD: Recommendations for early detection, monitoring, treatment and rehabilitation. 2007.

61. Lange P. Pharmacological treatment of COPD [Internet]. Treatment of COPD. 2016. Available from: https://www.sundhed.dk/borger/patienthaandbogen/lunger/sygdomme/kol/behandling-af-kol/

62. Nolte E, Knai C MM. Managing Chronic conditions – Experience in eight countries [Internet]. 2008. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.147.7245&rep=rep1&type=pdf

63. Danish Health Authority. National klinisk retningslinje for rehabilitering af patienter med KOL. 2014.

64. Danish Society of Respiratory Medicine. COPD Stratification [Internet]. Danish COPD guidelines. 2012. Available from: http://lungemedicin.dk/fagligt/101-dansk-kol-retningslinje-2012.html/#stratificering

65. The Danish Parliament. The Danish Health Act [Internet]. retsinformation.dk; 2014. Available from: https://www.retsinformation.dk/forms/r0710.aspx?id=152710#Kap1

66. Polisena J, Tran K, Cimon K, Hutton B, McGill S, Palmer K, et al. Home telehealth for chronic obstructive pulmonary disease: A systematic review and meta-analysis. J Telemed Telecare [Internet]. 2010;16(3):120–7. Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L358647046

67. Jakobsen AS, Laursen LC, Schou L, Emme C, Phanareth K V. Varying effect of telemedicine in the treatment of chronic obstructive pulmonary disease-a systematic review. Ugeskr Laeger [Internet]. 2012;174(14):936–42. Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L364927375

68. Franek J. Home telehealth for patients with chronic obstructive pulmonary disease (COPD): an evidence-based analysis. Ont Health Technol Assess Ser [Internet]. 2012 Jan;12(11):1–58. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3384362&tool=pmcentrez&rendertype=abstract

69. McLean S, Liu J, Pagliari C, Car J, Sheikh A. Telehealthcare for asthma. Cochrane Database Syst Rev [Internet]. 2009;(2). Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L355267549

70. Sicotte C, Pare G, Morin S, Potvin J, Moreault MP. Effects of home telemonitoring to support improved care for chronic obstructive pulmonary diseases. Telemed J E Health [Internet]. 2011;17(2):95–103. Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L362118949

71. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of ehealth on the quality and safety of health care: A systematic overview. PLoS Med [Internet]. 2011;8(1). Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-79951644557&partnerID=40&md5=89fe6299ef530cad141cc2a3915e45f7

72. Mistry H. Systematic review of studies of the cost-effectiveness of telemedicine and telecare. Changes in the economic evidence over twenty years. J Telemed Telecare [Internet]. 2012;18(1):1–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22101609

73. Sood S, Mbarika V, Jugoo S, Dookhy R, Doarn CR, Prakash N, et al. What is telemedicine? A collection of 104 peer-reviewed perspectives and theoretical underpinnings. Telemed J E Health [Internet]. 2007;13(5):573–90. Available from: http://www.liebertonline.com/doi/pdfplus/10.1089/tmj.2006.0073\nhttp://www.ncbi.nlm.nih.gov/pubmed/17999619

74. Goodwin N. The State of Telehealth and Telecare in the UK: Prospects for Integrated Care. J Integr Care [Internet]. 2010;18(6):3–10. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-78649913761&partnerID=tZOtx3y1

75. Doughty K, Monk A, Bayliss C, Brown S, Dewsbury L, Dunk B, et al. Telecare, telehealth and assistive technologies — do we know what we're talking about? J Assist Technol. 2007;1(2):6–10.

76. Darkins A, MA C. Telemedicine and Telehealth: Principles, Policies, Performances and Pitfalls. New York: Springer Publishing Company Inc.; 2000.

77. Oh H, Rizo C, Enkin M, Jadad A. What Is eHealth: A Systematic Review of Published Definitions. J Med Internet Res. 2005;7(1).

78. McLean S, Nurmatov U, Liu JLY, Pagliari C, Car J, Sheikh A. Telehealthcare for chronic obstructive pulmonary disease: Cochrane

Review and meta-analysis. Br J Gen Pract [Internet]. 2012;62(604):e739–49. Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L365948923

79. Miller E. Solving the disjuncture between research and practice, telehealth trends in the 21st Century. Health Policy (New York). 2007;82(2):133–41.

80. McLean S, Protti D, Sheikh A. Telehealthcare for long term conditions. BMJ. 2011;342(d120).

81. McLean S, Sheikh A. Does telehealthcare offer a patient-centred way forward for the community-based management of long-term respiratory disease? Prim Care Respir J [Internet]. 2009;18(3):125–6. Available from: http://dx.doi.org/10.3132/pcrj.2009.00006

82. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. Fourth edi. Oxford University Press; 2015.

83. INATHA. About INAHTA [Internet]. Available from: http://www.inahta.org/

84. Fox-Rushby J, Cairns J. Economic Evaluation. England: Open University Press, McGraw-Hill; 2005.

85. Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. Oxford University Press; 2006.

86. Claxton K, Sculpher M. Using value of information analysis to prioritise health research: some lessons from recent UK experience. Pharmacoeconomics [Internet]. 2006;24(11):1055–68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17067191

87. Gray A, Clarke P, Wolstenholme J, Wordsworth S. Applied Methods of Cost-effectiveness Analysis in Health Care. Gray A, Briggs A, editors. Oxford University Press; 2011.

88. Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR good research practices task force report. Value Heal. 2009;12(4):409–18.

89. Box GEP, Draper NR. Empirical Model-Building and Response Surfaces. Wiley; 1987.

90. Rothman K. Epidemiology. An Introduction. Second Edi. Oxford University Press; 2012.

91. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. BMJ.

2010;340(c332):1063–70.

92. Glick H, Doshi J, Sonnad S, Polsky D. Economic Evaluation in Clinical Trials. Oxford University Press; 2007.

93. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-Effectiveness Analysis Alongside Clinical Trials II—An ISPOR Good Research Practices Task Force Report. Value Heal [Internet]. 2015;18(2):161–72. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1098301515000169

94. O'Sullivan AK, Thompson D, Drummond MF. Collection of health-economic data alongside clinical trials: Is there a future for piggyback evaluations? Value Heal [Internet]. International Society for Pharmacoeconomics and Outcomes Research (ISPOR); 2005;8(1):67–79. Available from: http://dx.doi.org/10.1111/j.1524-4733.2005.03065.x

95. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. Bmj [Internet]. 2008;337(nov11 2):a2390–a2390. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.a2390

96. Fletcher R, Fletcher SW. Clinical Epidemiology: The Essentials. Fifth edit. Wolters Kluwer / Lippincott Williams & Wilkins; 2013.

97. Rovithis D. Do health economic evaluations using observational data provide reliable assessment of treatment effects? Health Econ Rev [Internet]. Health Economics Review; 2013;3(1):21. Available from: /abstract

98. Shiell A, Gerard K, Donaldson C. Cost of illness studies: An aid to decision-making? Health Policy (New York). 1987;8(3):317–23.

99. Jury T. Cash Flow Analysis and Forecasting: The Definitive Guide to Understanding and Using Published Cash Flow Data. Wiley Finance Series; 2012.

100. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [Internet]. Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration; 2011. Available from: http://handbook.cochrane.org/index.htm#chapter_15/15_incorporating_economics_evidence.htm

101. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health [Internet]. 2013;16(2):231–50. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/23538175

102. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: A report of the ISPOR-SMDM modeling good research practices task force-2. Value Heal [Internet]. Elsevier Inc.; 2012;15(6):804–11. Available from: http://dx.doi.org/10.1016/j.jval.2012.06.016

103. Berger ML, Dreyer N, Anderson F, Towse A, Sedrakyan A, Normand SL. Prospective observational studies to assess comparative effectiveness: The ISPOR good research practices task force report. Value Heal [Internet]. Elsevier Inc.; 2012;15(2):217–30. Available from: http://dx.doi.org/10.1016/j.jval.2011.12.010

104. Berger ML, Mamdani M, Atkins D, Johnson ML. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force . Value Health [Internet]. 2009;12(8):1044–52. Available from: http://dx.doi.org/10.1111/j.1524-4733.2009.00600.x

105. Briggs a, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. Health Econ. 1994;3(January):95–104.

106. Gold MR, Siegel JE, Russell LB, Weinstein M. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996.

107. North Denmark Region. About TeleCare North (in Danish) [Internet]. [cited 2016 Jan 27]. Available from: http://www.rn.dk/Sundhed/Til-sundhedsfaglige-og-samarbejdspartnere/TeleCare-Nord/Om-projektet

108. North Denmark Region. Report on the experiences gained in TeleCare North [Internet]. Available from: http://www.rn.dk/Sundhed/Til-sundhedsfaglige-og-samarbejdspartnere/TeleCare-Nord.aspx

109. Dinesen B. About the Telekat project [Internet]. [cited 2015 Aug 18]. Available from: http://www.telekat.eu/index.php?id=13

110. Region Nordjylland. Funding application for TeleCare North. 2012. p. 1–17.

111. Udsen F, Lilholt P, Hejlesen O, Ehlers L. Effectiveness and cost-effectiveness of telehealthcare for chronic obstructive pulmonary disease: study protocol for the Danish “TeleCare North” pragmatic cluster-randomized trial. Trials. 2014;15(178).

112. Region Nordjylland. Research design for TeleCare North. 2012. p. 1–8.

113. The Danish Agency for Digitilisation. Dissemination of telemedicine throughout Denmark [Internet]. Strategy for Digital Welfare 2013-2020. 2016. Available from: http://www.digst.dk/Digital-velfaerd/Initiativer-og-projekter/Projekter-i-Strategi-for-digital-velfaerd/Udbredelse-af-telemedicin-i-hele-landet_fokusomraade1

114. North Denmark Region. Business case for TeleCare North. 2012;1–10.

115. Lee A, Sandvei M, Hosbond T, Petersen J, Christiansen R. Final report on the experiences gained in KIH. 2015;(August):1–134.

116. Central Denmark Region. Clinically integrated home monitoring [Internet]. 2016. Available from: http://www.rm.dk/sundhed/faginfo/center-for-telemedicin/telemedicinske-projekter/hjemmemonitorering-k.i.h/

117. Udsen F, Hejlesen O, Ehlers L. A systematic review of the cost and cost-effectiveness of telehealthcare to patients suffering from chronic obstructive pulmonary disease. J Telemed Telecare. 2014;20(4):212–20.

118. Strauss A, Corbin J. Basics of qualitative research: Grounded theory procedures and techniques. Sage Publications; 1990.

119. Creswell J. Research Design: Qualitative, Quantitative and Mixed Methods Approaches. Fouth edit. Sage Publications; 2014.

120. Patton M. Qualitative Research & Evaluation Methods: Integrating Theory and Practice. Fourth edi. Sage Publications; 2015.

121. Bowen GA. Document analysis as a qualitative research method. Qualtative Res J [Internet]. 2009;9(no.2):27–40. Available from: http://dx.doi.org/10.3316/QRJ0902027\nhttp://\nhttp://dx.doi.org/10.3316/QRJ0902018

122. Miles M, Huberman A. Qualitative Data Analysis. Second edi. Thousand Oaks, Sage Pulications; 1994.

123. Rixon L, Hirani SP, Cartwright M, Beynon M, Doll H, Steventon A, et al. A RCT of telehealth for COPD patient's Quality of life: The Whole System Demonstrator Evaluation. Clin Respir J [Internet]. 2015;1–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26260325

124. Steventon a., Bardsley M, Billings J, Dixon J, Doll H, Hirani S, et al. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. Bmj [Internet]. 2012 Jun 21;344(jun21 3):e3874–e3874. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.e3874

125. Henderson C, Knapp M, Fernández J, Beecham J, Hirani S, Cartwright M, et al. Cost effectiveness of telehealth for patients with long term conditions

(Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. BMJ. 2013;346.

126. Cartwright M, Hirani SP, Rixon L, Beynon M, Doll H, Bower P, et al. Effect of telehealth on quality of life and psychological outcomes over 12 months (Whole Systems Demonstrator telehealth questionnaire study): Nested study of patient reported outcomes in a pragmatic, cluster randomised controlled trial. BMJ [Internet]. 2013;346(7897). Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84874641201&partnerID=40&md5=651e11d1ae013e7341ec115cce749c05

127. Stoddart A, van der Pol M, Pinnock H, Hanley J, McCloughan L, Todd A, et al. Telemonitoring for chronic obstructive pulmonary disease: a cost and cost-utility analysis of a randomised controlled trial. J Telemed Telecare [Internet]. 2015;21(2):108–18. Available from: http://jtt.sagepub.com/lookup/doi/10.1177/1357633X14566574

128. Pinnock H, Hanley J, McCloughan L, Todd A, Krishan A, Lewis S, et al. Effectiveness of telemonitoring integrated into existing clinical services on hospital admission for exacerbation of chronic obstructive pulmonary disease: researcher blind, multicentre, randomised controlled trial. Bmj. 2013;347(October):f6070.

129. Udsen F, Hejlesen O, Ehlers L. Cost-effectiveness of telehealthcare to patients with chronic obstructive pulmonary disease: Results from the Danish “TeleCare North” cluster-randomized trial. Forthcoming. 2015.

130. Udsen FW, Hejlesen OK, Ehlers L. Heterogeneity analysis of telehealthcare to patients with chronic obstructive pulmonary disease: The case of the Danish “TeleCare North” cluster-randomized trial. Forthcoming. 2016;

131. Lilholt PH, Udsen FW, Ehlers LH. Telehealthcare for patients suffering from COPD : Effects on health-related quality of life - Results from the Danish “ TeleCare North ” cluster-randomised trial. Forthcoming. 2016;1–16.

132. PA Consulting Group. Initial business case for the dissemination of telemedicine in Denmark. 2015.

133. The Danish Agency for Digitilisation. TeleCare Nord har vist vejen til landsdækkende telemedicin [Internet]. 2016. Available from: http://www.digst.dk/Servicemenu/Nyheder/Nyhedsarkiv/Digitaliseringsstyrelsen/2016/TeleCare-Nord-har-vist-vejen-til-landsdaekkende-telemedicin

134. The Danish Agency for Digitilisation. Landsdækkende udbredelse af telemedicin til borgere med KOL senest med udgangen af 2019 [Internet]. 2016. Available from: http://www.digst.dk/Digital-velfaerd/Udbredelse-af-

telemedicin-til-borgere-med-KOL

135. The Danish Agency for Digitilisation. Økonomien i telemedicin til borgere med KOL [Internet]. 2016. Available from: http://www.digst.dk/Digital-velfaerd/Udbredelse-af-telemedicin-til-borgere-med-KOL/Om-udbredelse-af-telemedicin-til-borgere-med-KOL/Oekonomien-i-telemedicin

136. The Danish Agency for Digitilisation. Statens business case-model [Internet]. Version 3.0. 2016. Available from: http://www.digst.dk/Styring/Business-case-model

137. The Danish Agency for Digitilisation. Fremtidens telemedicinske løsninger står på spring [Internet]. 2015. Available from: http://www.digst.dk/Servicemenu/Nyheder/Nyhedsarkiv/Digitaliseringsstyrelsen/2015/Fremtidens-telemedicinske-loesninger-staar-paa-spring.aspx

138. The Danish Agency for Digitilisation. Landsdækkende udbredelse af telemedicin til borgere med KOL [Internet]. 2016. Available from: http://www.digst.dk/Digital-velfaerd/Udbredelse-af-telemedicin-til-borgere-med-KOL/Om-udbredelse-af-telemedicin-til-borgere-med-KOL

139. Stigaard D. Regionsdirektør: Forfejlet kritik af business cases. Altinget.dk [Internet]. 2016 Apr 26; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/regionsdirektoer-forfejlet-kritik-af-business-cases

140. Kidholm K. Forsker: Når business cases bliver bestillingsarbejde. Altinget.dk [Internet]. 2016 Apr 26; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/forsker-naar-business-cases-bliver-bestillingsarbejde

141. Jensen GD, Pedersen K. KL om business cases: Fra misfostre til meningsfulde gevinster. Altinget.dk [Internet]. 2016 Apr 26; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/kl-om-business-cases-fra-misfostre-til-meningsfulde-gevinster

142. Zeberg RH. Styrelse: Business cases giver gode beslutningsgrundlag. Altinget.dk [Internet]. 2016; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/styrelse-business-cases-giver-gode-beslutningsgrundlag

143. Kidholm K. Forsker: Stil større krav til evalueringer af telemedicin. Altinget.dk [Internet]. 2016; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/forsker-stil-stoerre-krav-til-evalueringer-af-telemedicin

144. Pedersen KM. Professor: Uholdbare evalueringer af telemedicin. Altinget.dk [Internet]. 2015 Dec; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/professor-uholdbare-

evalueringer-af-telemedicin

145. Stigaard D. Regionsdirektør: Kampen om den telemedicinske sandhed. Altinget.dk [Internet]. 2016; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/regionsdirektoer-kampen-om-den-telemedicinske-sandhed

146. Pedersen KM. Professor: Hvornår skal man udrulle telemedicinske løsninger? Altinget.dk [Internet]. 2016; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/professor-hvornaar-skal-man-udrulle-telemedicinske-loesninger

147. Pedersen K. Professor: Telemedicinforsøg uden klare konklusioner. Altinget.dk [Internet]. 2015 Nov; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/professor-telemedicinforsoeg-uden-klare-konklusioner

148. Svenning W. Horsens: Derfor er telemedicin så svært. Altinget.dk [Internet]. 2016; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/horsens-derfor-er-telemedicin-saa-svaert

149. Lose S. Regionerne: Business cases er stadig den bedste løsning. Altinget.dk [Internet]. 2016 Apr 19; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/regionerne-business-cases-er-stadig-den-bedste-loesning

150. Pedersen KM. Professor: Telemedicinske business cases er ren feberfantasi. Altinget.dk [Internet]. 2016 Apr 19; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/professor-telemedicinske-business-cases-er-ren-feberfantasi

151. Christensen J. Forsker: Telemedicin viser stort potentiale under svære vilkår. Altinget.dk [Internet]. 2016; Available from: http://www.altinget.dk/artikel/forsker-telemedicin-viser-stort-potentiale-under-svaere-vilkaar

152. Ehlers LH. Professor: Telemedicin fungerer ikke med fingerknips. Altinget.dk [Internet]. 2016; Available from: http://www.altinget.dk/artikel/professor-telemedicin-fungerer-ikke-med-et-fingerknips

153. Udsen FW. Sundhedsøkonom: TeleCare Nord-evalueringen letter prioriteringerne. Altinget.dk [Internet]. 2016; Available from: http://www.altinget.dk/artikel/sundhedsoekonom-telecare-nord-evalueringen-letter-prioriteringerne

154. Kidholm K. Forsker: Telemedicin er sundhedsvæsnets vilde vest. Altinget.dk [Internet]. 2015 Dec 28; Available from:

http://www.altinget.dk/velfaerdsteknologi/artikel/forsker-telemedicin-er-sundhedsvaesenets-vilde-vest

155. Brandt A. Lungeforening: KOL-patienter viser vejen for telemedicin. Altinget.dk [Internet]. 2015 Dec 8; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/lungeforening-kol-patienter-viser-vejen-for-telemedicin

156. Andersen J. Regionerne: Ny uvurderlig viden om telemedicin. Altinget.dk [Internet]. 2015 Nov 24; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/regionerne-uvurderlig-viden-om-telemedicin

157. Gaardsted K. S: Telemedicin - sølvpapirshat eller patientrevolution. Altinget.dk [Internet]. 2015 Sep 15; Available from: http://www.altinget.dk/velfaerdsteknologi/artikel/s-telemedicin-soelvpapirshat-eller-patientrevolution

158. Lorenzen M. Sundhedsøkonomer: Besparelser via tele-medicin er grebet ud af luften [Internet]. Ingeniøren. Available from: http://ing.dk/artikel/sundhedsokonomer-besparelser-tele-medicin-er-grebet-ud-af-luften-181830

159. Pedersen KM, Thomsen C. Telemedicin: Tro, håb og evidens. Ugeskrift for laeger. 2016;

160. Bossen C, Jensen LG, Udsen FW. Boundary-Object Trimming: On the Invisibility of Medical Secretaries' Care of Records in Healthcare Infrastructures. Comput Support Coop Work J Collab Comput. 2013;23:75–110.

161. Bossen C, Jensen LG, Udsen FW. Evaluation of a comprehensive EHR based on the DeLoneand McLean model for IS success: Approach, results,and success factors. Int J Med Inform. 2013;82(10):940–53.

162. Gupta S. Intention-to-treat concept: A review. Perspect Clin Res [Internet]. 2011;2(3):109. Available from: http://www.picronline.org/text.asp?2011/2/3/109/83221

163. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. Bmj [Internet]. 2012;345(sep04 1):e5661–e5661. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.e5661

164. Rabe-Hesketh S, Skrondal A. Multilevel and Longitudinal Modeling Using Stata. Volume I: Continouos Responses. Third edit. Texas: Stata Press; 2012.

165. Hayes R, Moulton L. Cluster Randomised Trials. Boca Raton: Chapman &

Hall/CRC Press; 2009.

166. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated health economic evaluation reporting standards (CHEERS)-explanation and elaboration: A report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. Value Heal [Internet]. Elsevier; 2013;16(2):231–50. Available from: http://dx.doi.org/10.1016/j.jval.2013.02.002

167. Bachmann MO, Fairall L, Clark A, Mugford M. Methods for analyzing cost effectiveness data from cluster randomized trials. Cost Eff Resour Alloc [Internet]. 2007;5(Icc):12. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2020454&tool=pmcentrez&rendertype=abstract

168. Finansministeriet. Aftaler om den kommunale og regionale økonomi for 2016 [Internet]. 2015. 90 p. Available from: http://www.fm.dk/publikationer/2007/aftaler-om-den-kommunale-og-regionale-oekonomi-for-2008/download/~/media/Files/Publikationer/2008/Download/Aftale_om_den_kommunale_og_regionale_okonomi_2008_endelig_version.ashx

169. Messner W. Making the Compelling Business Case. Decision-Making Techniques for Successful Business Growth. Houndmills: Palgrave Macmillan; 2013.

170. Schmidt M. The Business Case Guide. Third edit. Solution Matrix Limited; 2016.

171. Greenhalgh T. Rapid responses to Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial [Internet]. 2013. Available from: http://www.bmj.com/content/344/bmj.e3874/rapid-responses

172. Haesum LKE, Soerensen N, Dinesen B, Nielsen C, Grann O, Hejlesen O, et al. Cost-utility analysis of a telerehabilitation program: A case study of COPD patients. Telemed e-Health [Internet]. 2012;18(9):688–92. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84869057985&partnerID=40&md5=87200bfb992018c29819b0f75c0d60cf

173. Sullivan SD, Mauskopf J a, Augustovski F, Jaime Caro J, Lee KM, Minchin M, et al. Budget impact analysis-principles of good practice: report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. Value Health [Internet]. Elsevier; 2014 [cited 2014 Mar 19];17(1):5–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24438712

174. Usher Institute of Population Health Sciences and Informatics. Telescot [Internet]. Available from: http://www.telescot.org/

175. Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. Lancet [Internet]. 2005;365(9454):176–86. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15639301

176. Sculpher M. Reflecting heterogeneity in patient benefits: the role of subgroup analysis with comparative effectiveness. Value Health [Internet]. International Society for Pharmacoeconomics and Outcomes Research (ISPOR); 2010;13 Suppl 1:S18-21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20618790

177. Sculpher M. Subgroups and heterogeneity in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):799–806.

178. Drummond M, O'Brien B. Clinical importance, statistical significance and the assessment of economic and quality-of-life outcomes. Health Econ. 1993;2(3):205–12.

179. Heyckendorff P, Emtekær LK. Exploring User Experience of a Telehealth System for the Danish TeleCare North Trial. In: Cornet R, Stoicu-Tivadar L, Hörbst A, Calderón C, Andersen S, Hercigonja-Szekeres M, editors. Digital Healthcare Empowering Europeans: Proceedings of MIE2015, 27-29 May 2015, Madrid, Spain (pp. IOS Press; 2015. p. 301–5.

180. Taylor R, Iglesias C. Assessing the Clinical and Cost-Effectiveness of Medical Devices and Drugs: Are They That Different? Value Heal. 2009;12(4):404–6.

181. Drummond M, Griffin A, Tarricone R. Economic evaluation for devices and drugs--same or different? Value Health [Internet]. 2009 Jun;12(4):402–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19138306

182. Sorenson C, Tarricone R, Siebert M, Drummond M. Applying health economics for policy decision making: Do devices differ from drugs? Europace. 2011;13(SUPPL. 2):54–8.

183. Chapman AM, Taylor CA, Girling AJ. Are the UK systems of innovation and evaluation of medical devices compatible? The role of NICE's Medical Technologies Evaluation Programme (MTEP). Appl Health Econ Health Policy. 2014;12(4):347–57.

184. Green W, Hutton J. Health technology assessments in England: An analysis of the NICE medical technologies evaluation programme. Eur J Heal Econ. 2014;15(5):449–52.

185. Kirisits A, Redekop WK. The economic evaluation of medical devices: Challenges ahead. Appl Health Econ Health Policy. 2013;11(1):15–26.

186. Craig J a., Carr L, Hutton J, Glanville J, Iglesias CP, Sims AJ. A Review of

the Economic Tools for Assessing New Medical Devices. Appl Health Econ Health Policy [Internet]. 2015;13(1):15–27. Available from: http://link.springer.com/10.1007/s40258-014-0123-8

187. Gelijns AC, Russo MJ, Hong KN, Brown LD, Ascheim DD, Moskowitz AJ. Dynamics of Device Innovation: Implications for Assessing Value. Int J Technol Assess Health Care. 2013;29(4):365–73.

188. Ramsey S, Luce B, Deyo R, Franklin G. The limited state of technology assessment for medical devices: facing the issues. Am J Manag Care. 1998;4:188–99.

189. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? BMJ [Internet]. 2001 Feb 10;322(7282):355–7. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1119583&tool=pmcentrez&rendertype=abstract

190. Puffer S, Torgerson DJ, Watson J. Cluster randomized controlled trials. J Eval Clin Pract [Internet]. 2005;11(5):479–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16164589

191. Gomes M, Ng ES-W, Grieve R, Nixon R, Carpenter J, Thompson SG. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. Med Decis Making [Internet]. 2012 [cited 2014 Oct 16];32(2):350–61. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3757919&tool=pmcentrez&rendertype=abstract

192. Häkkinen L, Hilmola O-P. ERP evaluation during the shakedown phase: lessons from an after-sales division. Inf Syst J. 2008;18(1):73–100.

193. Birch S. Making the problem fit the solution: evidence-based decision-making and “Dolly” economics. In: Donaldson C, Mugford M, Vale L, editors. Evidence-based Health Economics: From Effectiveness to Efficiency in Systematic Review. London: BMJ Books; 2002. p. 133–147.

194. Byford S, Sefton T. Economic evaluation of complex health and social care interventions. Inst Econ Rev. 2003;186(1):98–103.

195. Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. Bmj. 2008;336(7656):1281–3.

196. Anderson R, Hardwick R. Realism and resources: Towards more explanatory economic evaluation. Evaluation [Internet]. 2016;22(3):323–41. Available from: http://evi.sagepub.com/cgi/doi/10.1177/1356389016652742

197. Pawson R, Tilley N. Realistic Evaluation. Sage Publications; 1997.

198. Van Belle S, Wong G, Westhorp G, Pearson M, Emmel N, Manzano A, et al. Can “realist” randomised controlled trials be genuinely realist? Trials [Internet]. Trials; 2016;17(1):313. Available from: http://trialsjournal.biomedcentral.com/articles/10.1186/s13063-016-1407-0

199. Astbury B, Leeuw FL. Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. Am J Eval. 2010;31(3):363–81.

200. Bonell C, Fletcher A, Morton M, Lorenc T, Moore L. Realist randomised controlled trials: A new approach to evaluating complex public health interventions. Soc Sci Med [Internet]. Elsevier Ltd; 2012;75(12):2299–306. Available from: http://dx.doi.org/10.1016/j.socscimed.2012.08.032

201. Bate S, Mendel P, Robert G. Organising for quality: the improvement journeys of leading hospitals in Europe and the United States. Oxford: Radcliffe; 2008.

202. Blackwood B, O’Halloran P, Porter S. Review: On the problems of mixing RCTs with qualitative research: the case of the MRC framework and the evaluation of complex healthcare interventions. J Res Nurs. 2010;15(6):511–21.

203. Funnell S, Rogers P. Purposeful Program Theory - Effective Use of Theories of Change and Logic Models. Jossey-Bass; 2011.

204. Marchal B, van Belle BTC S, Josefien van Olmen B, Hoerée T, Kegels G. Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research review of realist evaluations in health systems research. Evaluation. 2016;18(2):192–212.

205. Popper K. The Logic of Scientific Discovery. London: Routledge; 2002.

206. Pawson R. The Science of Evaluation – A Realist Manifesto. London: Sage Publications; 2013.

207. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions : new guidance. BMJ. 2008;337.

208. Moore G, Audrey S, Barker M, Bond L. Process evaluation of complex interventions. DecipherUkNet [Internet]. 2014;19-45-75. Available from: http://decipher.uk.net/wp-content/uploads/2014/11/MRC-PHSRN-Process-evaluation-guidance.pdf

209. Husereau D, Jacobs P, Manns B, Hoomans T, Marshall D, R T. Economic Evaluation of Complex Health System interventions: A Discussion Paper. Edmonton; 2014.

210. De Silva MJ, Breuer E, Lee L, Asher L, Chowdhary N, Lund C, et al.

Theory of Change: a theory-driven approach to enhance the Medical Research Councils' framework for complex interventions. Trials [Internet]. 2014;15(1):267. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24996765

211. Fletcher A, Jamal F, Moore G, Evans RE, Murphy S, Bonell C. Realist complex intervention science: Applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. Evaluation [Internet]. 2016;22(3):286–303. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27478401\nhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4946011

212. Salisbury C, Thomas C, O'Cathain A, Rogers A, Pope C, Yardley L, et al. TElehealth in CHronic disease: mixed-methods study to develop the TECH conceptual model for intervention design and evaluation. BMJ Open [Internet]. 2015;5(2):e006448–e006448. Available from: http://bmjopen.bmj.com/cgi/doi/10.1136/bmjopen-2014-006448

213. Thomas CL, Man M-S, O'Cathain A, Hollinghurst S, Large S, Edwards L, et al. Effectiveness and cost-effectiveness of a telehealth intervention to support the management of long-term conditions: study protocol for two linked randomized controlled trials. Trials [Internet]. 2014;15(1):36. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3906859&tool=pmcentrez&rendertype=abstract

214. Dixon P, Hollinghurst S, Edwards L, Thomas C, Gaunt D, Foster A, et al. Cost-effectiveness of telehealth for patients with raised cardiovascular disease risk: evidence from the Healthlines randomised controlled trial. BMJ Open [Internet]. 2016;6(9):e012355. Available from: http://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2016-012355

215. Dixon P, Hollinghurst S, Edwards L, Thoms C, Foster A, Davies B, et al. Cost-effectiveness of telehealth for patients with depression: evidence from the Healthlines randomised controlled trial. BJPsych Open [Internet]. 2016;2:262–269. Available from: http://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2016-012355

216. O'Cathain A, Drabble SJ, Foster A, Horspool K, Edwards L, Thomas C, et al. Being human: A qualitative interview study exploring why a telehealth intervention for management of chronic conditions had a modest effect. J Med Internet Res. 2016;18(6).

217. Bergmo TS. Using QALYs in telehealth evaluations: a systematic review of methodology and transparency. BMC Health Serv Res [Internet]. 2014;14(1):332. Available from: http://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-14-

332

218. Moen R, Nolan T, Provost L. Quality Improvement Through Planned Experimentation. Third edit. McGraw-Hill; 2012.

SUMMARY

This thesis contains the results from the Danish TeleCare North trial meant to assess the effectiveness and cost-effectiveness of a telehealthcare solution implemented in North Denmark Region from 2013-2015. The TeleCare North trial demonstrated no difference in health-related quality of life and the telehealthcare solution was not cost-effective for all included COPD patients. But there was a potential to target the solution to patients with severe COPD. The results also indicate that implementation could have a strong impact on cost-effectiveness, more so than health- or socio-demographic factors.

The results from the TeleCare North trial were used directly in a national decision to implement the telehealthcare solution to patients with severe COPD in Denmark and lead to considerable debate nationally. This debate could be viewed as an actual account of the usefulness of health economic evaluation for decision making meant to inform adaptation of the health economic evaluation approach.

Based on developments in realist evaluation and experiences with conducting the evaluation of TeleCare North, four principles for health economic evaluation of complex telehealthcare interventions is outlined in order to facilitate future health economic designs of telehealthcare that should ultimately answer if telehealthcare is cost-effective, for whom, why and under what circumstances.

ISSN (online): 2246-1256
ISBN (online): 978-87-7112-847-5

AALBORG UNIVERSITETSFORLAG