PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND

Dissertations in Social Sciences and Business Studies

UNIVERSITY OF EASTERN FINLAND

ERJA MUSTONEN

Telephone-based health coaching for chronic disease patients: evaluation of short- and long-term effectiveness of health benefits and costs

Telephone-based health coaching for chronic disease patients: evaluation of short- and long-term effectiveness of health benefits and costs

Erja Mustonen

Telephone-based health coaching for chronic disease patients: evaluation of short- and long-term effectiveness of health benefits and costs

Publications of the University of Eastern Finland
Dissertations in Social Sciences and Business Studies
No 240

University of Eastern Finland
Kuopio
2021

Grano Oy
Jyväskylä, 2021
Editor in-Chief: Markus Mättö
Editor: Anna Karttunen
Sales: University of Eastern Finland Library
ISBN: 978-952-61-3688-2 (print)
ISBN: 978-952-61-3689-9 (PDF)
ISSNL: 1798-5749
ISSN: 1798-5749
ISSN: 1798-5757 (PDF)

Author's address:	Faculty of Social Sciences and Business Studies University of Eastern Finland KUOPIO FINLAND
Doctoral programme:	Welfare, Health and Management Doctoral programme
Supervisors:	Professor Johanna Lammintakanen, Ph.D Department of Health and Social Management University of Eastern Finland KUOPIO FINLAND
	Professor, Miika Linna, D.Sc (Tech) Department of Industrial Engineering and Management the HEMA Institute Aalto University ESPOO FINLAND
Rewievers:	Professor Minna Kaila, MD Faculty of Medicine University of Helsinki HELSINKI FINLAND
	Professor Arto Ohinmaa, Ph.D Health Policy and Management, School of Public Health University of Alberta EDMONTON CANADA
Opponent:	Professor Minna Kaila, MD Faculty of Medicine University of Helsinki HELSINKI FINLAND

Mustonen, Erja
Telephone-based health coaching for chronic disease patients: evaluation of short- and long-term effectiveness of health benefits and costs
Kuopio: Itä-Suomen yliopisto, 2021
Publications of the University of Eastern Finland
Dissertations in Social Sciences and Business Studies; 240
ISBN: 978-952-61-3688-2 (print)
ISSNL: 1798-5749
ISSN: 1798-5749
ISBN: 978-952-61-3689-9 (PDF)
ISSN: 1798-5757 (PDF)

ABSTRACT

The aim of this study was to evaluate the effectiveness of a 1-year telephone-based health coaching intervention among high-risk chronic disease patients using a multimethod, multidisciplinary longitudinal approach. The study was conducted using a randomized controlled trial design; 1534 type 2 diabetes, coronary artery disease and cardiac heart failure patients were randomized into an intervention group (usual care and monthly telephone health coaching; N=1034) and a control group (usual care; N=501).

Effectiveness was evaluated based on four dimensions — clinical outcomes and cost-effectiveness in the short term (1-year follow-up) and social and health care costs and mortality and morbidity in the long term (8-year follow-up). The data were collected from patient health records and research nurses' measures, from patients with a 15D questionnaire on the health-related quality of life and national health and social care registries. The factors associated with effectiveness were also studied by interviewing health coaches (N=7; additional results in the summary). The evaluation process and results were reviewed and discussed from the perspective of rational decision making. Analyses were conducted using modified intention-to-treat (included available results), intention-to-treat (all allocated patients) and per protocol (patients who participated in the study) methods. In the sub-studies, statistical and health economic analyses were used, and interview material was analysed using inductive content analysis.

In the short term, significant improvements in diastolic blood pressure due to the health coaching were found, and health coaching increased health-related quality of life with acceptable costs. In the long term, severe chronic disease complications occurred less frequently, and the total social and health care costs were lower in the intervention group from 2.5 years onwards. Statistically significant differences were found in the per protocol analysis. Based on the health coaches' interviews, the learning of coaching skills took 1–3 years, and continuous support, mentoring and quality assurance were essential in developing the coaching skills. The coaches also observed that it took time for patients to integrate behaviour changes into their daily lives. Therefore, the evaluation of health coaching interventions should extend to at least 3 years using a multidisciplinary, multidimensional approach.

From the rational decision-making viewpoint, understanding the nature of the intervention is essential for decision makers to set realistic targets and to evaluate them in a timely fashion. The overall results suggest that health coaching has a positive effect on health, quality of life and social and health care costs, particularly for those patients able and willing to participate in the intervention. Therefore, health coaching could be the part of self-care support in the chronic care delivery system.

Keywords: health coaching, self-care, coronary artery disease, diabetes mellitus, type 2, heart failure, cost effectiveness, effectiveness, quality of life, health care costs, decision making, randomized controlled trial

Mustonen, Erja
Terveysvalmennuksen vaikuttavuus kroonisesti sairailla potilailla: lyhyen ja pitkän aikavälin terveyshyötyjen ja kustannusten arviointi
Kuopio: Itä-Suomen yliopisto, 2021
Publications of the University of Eastern Finland
Dissertations in Social Sciences and Business Studies; 240
ISBN: 978-952-61-3688-2 (nid.)
ISSNL: 1798-5749
ISSN: 1798-5749
ISBN: 978-952-61-3689-9 (PDF)
ISSN: 1798-5757 (PDF)

TIIVISTELMÄ

Tutkimuksen tarkoituksena oli arvioida yhden vuoden kestävän, puhelimitse toteutetun terveysvalmennuksen vaikuttavuutta kroonisesti sairaiden, korkean riskin potilaiden kohdalla käyttämällä monimenetelmällistä ja monitieteellistä pitkittäistutkimusasetelmaa. Tutkimuksessa käytettiin satunnaistettua kontrolloitua tutkimusasetelmaa, jossa 1534 sisäänottokriteerit täyttävää tyypin 2 diabetes-, sepelvaltimotauti- ja sydämenvajaatoimintapotilasta arvottiin joko interventioryhmään (N=1034) tai kontrolliryhmään (N=501). Interventioryhmän potilaat saivat normaalien sosiaali- ja terveyspalvelujen lisäksi puhelimitse terveysvalmennusta noin kerran kuukaudessa ja kontrolliryhmän potilaat käyttivät tavanomaisia sosiaali- ja terveydenhuollon palveluja.

Vaikuttavuutta arvioitiin neljästä näkökulmasta: kliinisiä muuttujia ja kustannus-vaikuttavuutta lyhyellä aikavälillä (yhden vuoden seuranta) ja sosiaali- ja terveydenhuollon kustannuksia sekä kuolleisuutta ja sairastavuutta pitkällä aikavälillä (kahdeksan vuoden seuranta). Lyhyen ajan tutkimusaineisto kerättiin potilastietojärjestelmistä ja tutkimushoitajien mittauksista (kliiniset tiedot) ja kustannus-vaikuttavuusaineisto kerättiin potilailta 15D kyselylomakkeella (terveyteen liittyvä elämänlaatu) ja kustannukset kansallisista rekistereistä. Pitkän ajan seuranta-aineisto kerättiin kansallisista rekistereistä. Lisäksi tutkittiin terveysvalmentajien näkemyksiä terveysvalmennuksen vaikuttavuuteen liittyvistä tekijöistä haastattelemalla tutkimushankkeessa toimineita terveysvalmentajia (N=7). Vaikuttavuuden arvioinnin arviointiprosessia ja tutkimuksen tuloksia tarkasteltiin rationaalisen päätöksenteon näkökulmasta. Tutkimuksessa käytettiin modifioidun hoitoaikeen, hoitoaikeen ja protokollan mukaisia aineistoanalyysejä. Määrälliset aineistot analysoitiin käyttämällä tilastollisia ja terveystaloustieteellisiä menetelmiä ja haastatteluaineisto analysoitiin induktiivisella sisällönanalyysillä.

Terveysvalmennukseen osallistuneiden potilaiden diastolinen verenpaine laski tilastollisesti merkitsevästi ja terveysvalmennus lisäsi elämänlaatua kohtuullisin kustannuksin lyhyellä aikavälillä. Pitkällä aikavälillä hoitoaikeen mukaisissa analyyseissä kuolemia ja vakavia kroonisten sairauksien komplikaatioita esiintyi vähemmän ja sosiaali- ja terveydenhuollon kokonaiskustannukset olivat matalammat interventioryhmässä. Kustannukset kääntyivät kahden ja puolen vuoden seurannan jälkeen interventioryhmän hyväksi. Tilastollisesti merkitseviä eroja havaittiin protokollan mukaisissa analyyseissä. Terveysvalmentajien valmennustaitojen omaksuminen kesti yhdestä kolmeen vuoteen, ja olennaista osaamisen kehittymisessä oli jatkuva tuki ja terveysvalmennuksen laadunseuranta. Terveysvalmentajien mukaan myös potilaiden käyttäytymisen muutosten integroiminen jokapäiväiseen elämään vie aikaa. Siten terveysvalmennuksen vaikuttavuuden arvioinnin seuranta-aikaa on tarpeen laajentaa vähintään kolmevuotiseksi käyttäen monitieteellistä ja monimenetelmällistä arviointia.

Rationaalisen päätöksenteon näkökulmasta intervention luonteen tunteminen on ensisijaista, jotta päätöksentekijät osaavat asettaa realistiset tavoitteet ja arvioida niitä oikea-aikaisesti. Päätulosten mukaan terveysvalmennusinterventiolla on myönteisiä vaikutuksia kroonisesti sairaiden potilaiden terveyteen, elämänlaatuun ja sosiaali- ja terveydenhuollon kustannuksiin varsinkin kohdennettuna niihin potilaisiin, jotka ovat kykeneviä ja halukkaita terveysvalmennukseen. Siten terveysvalmennus voisi olla yksi omahoidon tukimuoto kroonisten sairauksien hoidon palveluvalikoimassa.

Asiasanat: terveysvalmennus, itsehoito, sepelvaltimotauti, diabetes, sydämen vajaatoiminta, vaikuttavuus, kustannukset, elämänlaatu, rationaalisuus, päätöksenteko, satunnaistettu vertailukoe

Acknowledgements

This project began in 2006 when TERVA, development and research program was established. The journey has been long and sometimes rocky and could not have succeeded without many exceptional people and their valuable contributions. I wish to express my sincere gratitude to these people for supporting me during the completion of my doctoral thesis.

I wish to express my gratitude to my principal supervisor, Professor Johanna Lammintakanen, PhD for her support and target-oriented, clear and detailed guidance during this long process. She asked me very fruitful questions that expanded my thinking significantly. I am also most grateful to my second supervisor, Professor Miika Linna, whose expertise, enthusiasm, encouragement and support were invaluable for this thesis.

I am deeply grateful to the pre-examiners, Professor Minna Kaila from the University of Helsinki and Professor Arto Ohinmaa from the University of Alberta, for their valuable comments, which helped me to improve this summary significantly.

Warmest thanks go to my scientific writing group, many members of which have contributed significantly to the TERVA program. I am privileged to have had the opportunity to collaborate with all of you. To Martti Talja, professor, MD and head of TERVA who hired me to work on a TERVA project, your endless support and encouragement has been instrumental over the years. I warmly thank Pilvikki Absetz, PhD, for all the support, encouragement and collaboration during this process. Your influence on this study is inestimable. You helped all of us to understand the basic principles relating to strength-based behavioural change, and it has been my privilege to work with you on many projects. I warmly thank Kristiina Patja, MD for her support and sharing her innovative thinking and expertise in medical sciences and health promotion. I warmly thank my long-time co-worker Risto Kuronen, MD for the partnership and kind support over the years. We have shared some unforgettable moments in our development projects, with both successes and failures. I also want to thank Iiris Hörhammer, PhD for her support, careful reading, valuable comments and contributions to this work.

I am also most grateful to all the health coaches who bravely threw themselves into the new challenges in the TERVA project. You provided very valuable information from your experiences. Thank you Päivi Suikki, Soile Borgman, Mervi Jaakkola, Ritva Kaikkonen, Marjo Kanervio, Tuula Kinnunen and Kirsi Reijonen.

I am thankful for my many co-workers. I am privileged to have worked with you during these years and to have received your support in different phases. Much appreciation goes to a great TERVA project team from Pfizer and Sitra: Jari Turpeinen, Mari Luodes, Timo Ovaska, Virpi Pyykkö and, Ilmo Parvinen. I also thank Anssi Auvinen and Kari Tokola from University of Tampere, members of TERVA scientific group, and co-workers from MediNeuvo, Diabetes Association, Heart Association and Prowellness – we built TERVA together. Thanks also to the many great people from POTKU (Patient in the Driver's seat) and Better Everyday Life (Parempi Arki) projects: 'the boys from South Ostrobothnia': Jaakko Pihlajamäki, Matti Rekiaro, Raimo Rintala, project managers and all project workers – we worked together over seven years and shared many great moments in our development work. I am thankful to Anneli Hujala for your great cooperation in our projects and your support in my studies. I am also thankful to my closest co-workers in the last years in Päijät-Häme: Mari Sisso, Sari Hokkanen, Anna Päätalo, Heli Lehtovaara and Risto Raivio – we shared many memorable years together. I am thankful to my dear friend Outi Kähkönen for all the encouragement and fruitful telephone discussions during this process.

I am especially grateful to my mother and late father who always encouraged and supported me in my studies. I warmly thank my siblings Hilkka, Hannu, Eeva and Helena and their families, with whom I have shared good times and bad. My late brothers Erkki and Heikki cannot share this day with me, but I know they are proud of me in heaven.

I also thank my daughter Emilia and my son Sami. You were quite young at the beginning of this process. During these years I tried to guide and advise you to the best of my ability, and now you both provide me with fresh thoughts and support me in concrete ways, such as reading my texts and helping to solve IT problems. I am so proud of you and thankful to have you in my life. Always believe in yourselves and your dreams. Warmest thanks also go to Pete. You brought joy, and many unforgettable adventures into my life and always supported me in my work, even on lazy days.

Finally, I thank the Päijät-Häme Joint Authority for Health and Wellbeing, Pfizer, the Finnish Innovation Fund (SITRA) and the Onni and Hilja Tuovinen Foundation for their financial support.

In Lahti, November 2020

Erja Mustonen

Table of contents

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AEA	American Evaluation Association
AMI	Myocardial infarction
APR	Ambulatory and primary care related patient group
BMI	Body mass index
CA	Cost analysis
CABG	Coronary artery bypass grafting
CAD	Coronary artery disease
CBA	Cost–benefit analysis
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptability curve
CHF	Congestive heart failure
CI	Confidence interval
CUA	Cost–utility analysis
CVD	Cardiovascular disease
DCCT	The Diabetes Control and Complications Trial Research Group
CONSORT	Consolidated Standards of Reporting Trials
DRG	Diagnosis-related group
EGHSPA	Expert groups on health system performance assessment
EHR	Electronic health records
EU	European Union
GOAL	Good ageing in Lahti region - program
HbA1c	Glycated haemoglobin
HDL	High density lipoprotein
HILMO	National Discharge Registry
ICER	Incremental cost-effectiveness ratio
ITT	Intention-to-treat
LDL	Low-density lipoprotein
LTC	Long-term care
mITT	Modified intention-to-treat
NCSP	Nordic Classification of Surgical Procedures
NYHA	New York Heart Association - classification
OEDC	Organization for Economic Cooperation and Development

PAS	Patient Administration system
PP	Per protocol
PTCA	Percutaneous transluminal coronary angioplasty
PVD	Peripheral vascular disease
QoL	Quality of life
S-kol	Serum total cholesterol
QALY	Quality adjusted life years
RCT	Randomized controlled trial
Sos. HILMO	Care Register for Social Welfare
TERVA	Telephone-based health coaching development and research program
T2D	Type 2 diabetes
UAP	Unstable angina pectoris
UK	United Kingdom
UKPDS	UK Prospective Diabetes Study
US	United States
WTP	Willingness to pay

1 Introduction

1.1 The burden of chronic disease in health care

The burden of chronic disease is a major challenge in health care. It has been estimated that chronic diseases are responsible for 70%–80% of total healthcare costs in European Union (EU) countries (Rieken et al. 2013). Therefore, the key question for political–administrative decision makers is how to manage chronic diseases more economically, particularly the prevalence of health behaviour-related diseases such as type 2 diabetes (T2D) and coronary artery disease (CAD) increase. It is estimated that in 2030, 50% of the population aged 65–69 will have at least two chronic conditions and 69% of deaths will be caused by chronic or non-communicable diseases (Scheller-Kreinsen et al. 2009; Barnett et al. 2012). T2D is one of the most prominent risk factors for cardiovascular disease (CVD), which in 2014 was the leading cause of death in the EU (37.1%) and in Finland (37.5%) (The Diabetes Control and Complications Trial Research Group (DCCT) 1993; Stratton et al. 2000; Eurostat 2014). The increased prevalence of macro- and microvascular complications of T2D (such as diabetic neuropathy, nephropathy and retinopathy) have resulted in increased disability and social and healthcare costs (Tamayo et al. 2014). It has been estimated that approximately 500,000 people in Finland live with T2D, resulting in approximately 1.5 billion Euros in total health care costs in 2011. The complications of T2D increase health care costs; costs without complications were approximately €3,036 per patient per year, and costs with complications were approximately €7,069 per patient per year (Jarvala et al. 2010).

1.2 Self-care support as a part of chronic disease management

Self-management support interventions have been recognized as an essential part of chronic disease management (Panagioti et al. 2014), but the evidence of the effectiveness of self-care interventions is heterogeneous (see Trappenburg et al. 2013). One global problem in chronic disease management is adherence to treatment; approximately 50% of chronic disease patients comply with recommendations (World Health Organization 2003); 50% of patients take medicine as recommended; and 30% of patients adhere to healthy diets (Haynes et al. 2002; Pitkälä et al. 2005). Conventionally, self-management support interventions focus on the disease itself, emphasizing coordinated care, following evidence-based clinical guidelines and encouraging patient compliance to treatments; however, they focus less on patients' individual needs and behaviour (Ellrodt et al. 1997; Mattke et al. 2007). Recently, the trend of self-management support interventions is to move from compliance toward concordance and shared decision making by using coaching methods (Routasalo et al. 2009). Health coaching is a patient-centred and goal-oriented approach to self-care support based on shared decision making and collaborative goalsetting facilitated by motivational interviewing. It emphasizes and supports patients' autonomy instead of compliance (Hayes et al. 2008; Palmer et al. 2013; Olsen 2014; Härter et al. 2016.) The evidence on the effectiveness of health coaching is based mainly on short-term follow-up studies with mixed outcomes; concluded small effects or no effects and the effectiveness in long-term is defective (Dennis et al. 2013; Kivelä et al. 2014; Hale & Giese 2017; Tiede et al. 2017).

1.3 Context of the study

Currently, municipalities are responsible for organizing social welfare and health care in Finland. The basic services can be provided by a municipality alone or jointly by municipal authorities with other municipalities. Social welfare and health care services may also be purchased from other municipalities, organizations or private service providers. Specialized medical care (secondary care) is organized by hospital districts. (Ministry of Social Affairs and Health 2019.) The aim of the ongoing health and social service reform is to transfer responsibility for the organisation of social and health care service from municipalities to the autonomous regions larger than municipalities, called counties. The reform includes both structural reform and development of social and health care services. (Finnish Government 2020.)

The study was based on the Finnish health coaching development and research project (TERVA) conducted in the Päijät-Häme Social and Health District and that had a population of 212,000 in 2006–2009. In this region, the population aged over 65 has increased more rapidly than in other parts of Finland, and the cost of chronic diseases, such as diabetes, CAD and heart failure (HF) is high, particularly in secondary health care. The main aim of the health coaching program was to manage the burden of chronic disease by supporting patient self-care and health behaviour. The self-care support was expected to lead to improved clinical outcomes and quality of life (QoL) and to contribute to a more efficient use of social and health care services. (TERVA protocol 2007.)

Similar health coaching programs were previously carried out in the US, the UK (Birmingham Own Health) and Italy by Pfizer Health Solutions and have also been tested in Finland at the behest of the Finnish Innovation Fund (SITRA). To date, according to the researcher's knowledge, scientific research of those previous projects has only been conducted in UK (Steventon et al. 2013). The coaching model (engaging, informing, involving, empowering) is based on behaviour change and health coaching techniques delivered by specially trained health coaches. The local decision makers in Päijät-Häme gave consent to introduce the TERVA health coaching program, except in the biggest city in the region (Lahti). Thus, the final study included 12 municipalities with a population of 112,000. The TERVA project was carried out and funded by a private–public partnership with four main partners—the Finnish Innovation Fund (SITRA), Pfizer, the Päijät-Häme Joint Authority for Health and Wellbeing and Business Finland. The health coaching program was running down the end of October in 2009. The study cohort in this dissertation is based on the

TERVA trial (trial registration NCT00552903, registration date 1 November 2007, updated 3 February 2009.

1.4 The aim of the study

The aim of the study was to evaluate the effectiveness of a 1-year telephone health coaching intervention on health benefits and social and health care costs in the short and long terms among chronically ill patients. The specific aims of this study were:

1. To evaluate the effectiveness of health coaching on clinical outcomes in the short term (1 year) (Sub-study I).
2. To evaluate the cost-effectiveness of health coaching in the short term (1 year) (Sub-study II).
3. To evaluate the impact of health coaching on social and health care costs in the long term (8 years) (Sub-study III).
4. To evaluate the effectiveness of health coaching on clinical endpoints in the long term (8 years) (Sub-study IV).
5. To describe health coaches' perceptions of the factors associated with the effectiveness of health coaching (additional results in the summary).

The study design is shown in Figure 1.

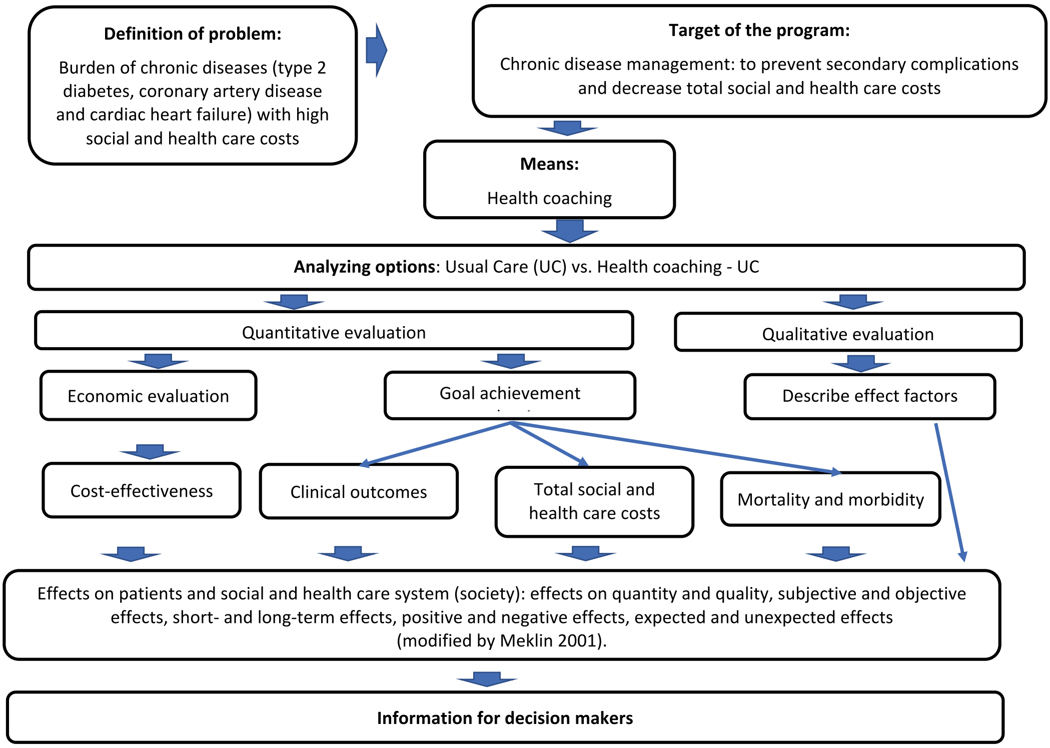


Figure 1. Study design for evaluating the effectiveness of telephone-based health for chronic disease patients.

This study provides new information on the short and long-term effectiveness of telephone health coaching among high-risk chronic disease (T2D, CAD and CHF) patients. Previous studies have mainly focused on short term effectiveness of health coaching (e.g. Dennis et al. 2007; Kivelä et al. 2014: Dejonghe et al. 2017).

1.5 The relationship of the study to health management sciences

This study belongs to the field of health management science, which focuses on the organization, management and decision making of health care organizations. Research on this topic is often multidisciplinary and is closely linked to administrative sciences, health sciences and medicine. (Vuori

2005, 21–25; Niiranen & Lammintakanen 2011, 113; Salminen 2011, 310.) Evaluation research is a common approach in several disciplines, including health management science. The general purpose of evaluating effectiveness is to provide information for decision makers, service producers, citizens and society in general (Sinkkonen & Kinnunen 1999; Rajavaara 2006, 9; Drummond 2008.) In health economics, the goal of evaluation is to measure alternative health options in terms of cost and to help decision makers set priorities (Sintonen & Pekurinen 2006, 10–11). Traditionally, public sector management theories emphasize effective and rational decision making (Gabor 1976).

However, the health care system is complex and includes many interest groups (e.g. professionals, policy makers, managers and patients). Additionally, the decision-making environment might be turbulent due to funding and political reasons. Therefore, it is challenging for decision makers to set clear targets for development work and evaluation because there are many needs and interests (Shapira 1997, 3; Sanderson & Gruen 2006, 1, 10.) Multidisciplinary research may produce new information, as important factors might be 'hiding' between the traditional sciences (Mikkeli & Pakkasvirta 2007, 6–8). An analogy between administrative sciences and economics has been presented by Herbert A. Simon; administrative sciences conform to the same theory (economic behaviour) and functional (profit seeking) classification as economics (Salminen 2011, 310).

In this study, rational decision making provides the general framework for the research problem, process and methodology. The evaluation of the effectiveness of a complex self-care intervention is considered according to phases of rational decision making and is discussed in the context of rational decision making.

2 The principles of rational decision making as the basis for evaluating effectiveness

In general, the complexity of organizations often restricts rational decision making. Complexity refers to many interest groups and numerous different functions that must be managed simultaneously, and joint effect is difficult to predict thus the management of complex organization require also multidimensional knowledge for decision making (Harisalo 2008, 27–28).

The definitions of decision-making take account of this complexity. Decision making is described as a coherent and rational process in which alternative interests and perspectives are considered and the most optimal alternative is selected. This process has been seen as a way to increase rationality in policy making (Weiss 1972, 2; Shapira 1997; Adair 2006, 1; Vuori 2006, 39.) Rationality is associated with decision making, and the concept of being rational is understood as taking a reasonable, logic and systematic approach in relation to, for example, cost reduction, resource allocation and business development. Rationality has also been defined as the compatibility between choice and value (Salminen 1993, 59; Axelsson & Engström 2001; Oliveira 2007).

The theory of rational decision making has been developed in economics and assumes that 'economic man' is rational. It represents an ideal and standard model of evaluation approaches. (Sinkkonen & Kinnunen 1994, 71; Elster 1996; Sanderson & Gruen 2006, 8; Hunsson 2007.) Rational decision making emphasizes the detailed definition of a problem, target setting, relevant and reliable information and a systematic process to arrive at logical decisions. Alternative options need to be compared, and the consequences of each decision need to be understood (Gabor 1976; Simon 1979; Russ et al. 1996; Vedung 2003, 9; Oliveira 2007; Jeanes 2019.) According to Uzonwanne (2016), rational decision making is the most promising, effective and functional process for leaders, managers and individuals, especially when stakeholders, investments and high stakes are involved. Being the opposite of intuitive decision making, in rational decision making individuals use facts and a step-by-step procedure to arrive at a conclusion (Russ et al. 1996).

The first step in making a well-reasoned decision is to identify the problem and set realistic targets (see Figure 1). Target setting is important, as targets are instruments that enable measuring and evaluating choices that have been made (Øvretveit 1998, 183; Harisalo 2008, 19–21, 147; Uzonwanne 2016). Usually, defined values are permanent and guide peoples' or groups' behaviour (Virtanen 2007, 47). Harisalo (2008) has described the difficulty of target setting; short-term targets might conflict with long-term ones or targets might be too demanding, insignificant, symbolic or idealistic; unambiguous or ambiguous; official or unofficial; suitable or unsuitable for the circumstances (Harisalo 2008, 19–21, 147). However, money is a common measure in commercial and non-commercial organizations (Simon 1979, 106).

The theory of rational decision making emphasizes comprehensive and reliable information in decision making, but it also emphasizes the values 'behind' the decisions (Gabor 1976; Vedung 2003, 9; Harisalo 2008, 146–147, 149). The value of information depends on how important it is for decision making (Feldman & March 1981; Rossi 2004, 127–218). Information also determines the consequences of different options and what options are favoured. In economics, quantitative and numeric information is greatly emphasized, but information can also be qualitative (Harisalo 2008, 154.) Gabor (1976) discusses the facts and values in decision making; if the facts on which a decision is based are not verifiable, then the decision is not rational. It can be said that what cannot be verified cannot be considered rational. According to Gabor (1976, 278) "any decision-making process includes both fact and values". However, an organization that only relies on quantitative information might be rather conservative because it does not have the courage to make decisions based on qualitative information. Therefore, the main question is who sets the evaluation assignment and how information is given to decision makers. (Virtanen 2007, 39.)

A number of theories, methods and models (e.g. iconic, graphic, symbolic and mathematic) have been developed to facilitate decision making and choosing between two or more alternatives. In decision making, it is critical to choose the most appropriate model or method for evaluation. Descriptive decision theory explains and predicts how people make decisions, whereas normative decision theory concerns how people ought to choose when making decisions. Examples of normative theories are mathematical game theory, decision making under risk or ignorance, utilitarianism, classical mathematical probability, frequency, Bayesian decision theory, epistemology and social choice theory. (Sanderson & Gruen 2006, 18; Peterson 2008; Binmore 2009.) In this study, comparisons between study groups (usual care vs usual care and health coaching) were made using statistical and economic evaluation methods (see Table 2). The target of statistical analysis is to evaluate how the results of a sample can be generalized to the overall population and to make reliable inferences. These are addressed in statistical decision theories that attempt to deal with uncertainty in the data.

The theory of rational decision making also emphasizes that individuals should know all alternatives and consequences, but in real life this is often impossible (Simon 1979, 105). For this reason, Simon (1979) is interested in limitations of efficient decision making and presented the concept of bounded rationality. In real life, choices that are merely acceptable are made because of incomplete information, complex problems, uncertainty, our own limited processing capacity, the time available, the conflicting goals of decision makers, the lack of agreed criteria, foolishness and error (Simon 1979, 118–121; Jeanes 2019.) The cost of gathering relevant information and experiences might also restrict rationality; therefore, economic inputs from organizations and understanding the nature of investments are necessary. However, the necessary information is not always available. A lack of time also restricts rationality, as there is not always enough time to consider and assimilate information about different options (Simon 1979, 78; Tomer 1992; Boos & Jacquemart 2000; Sinervo 2011, 74.) Additionally, decision makers might also worry about criticism and therefore might not be ready to take risks in decision making. This is why decisions are bureaucratic; it is difficult to criticize conservative decisions. (Nutt 2003; Kahneman 2012, 237.)

3 Evaluation of effectiveness in health care

3.1 Evaluation research

‘Evaluation is an elastic word that stretches to cover judgements of many kinds’ (Weiss 1972, 1). Evaluation research employs theories and methods commonly used in traditional scientific disciplines, such as sociology, political science, administrative science, economics, educational science and legal science, with different research methods, such as ethnography, survey research, randomized experiments and cost–benefit analysis (CBA) (Weiss 1972, 4; Sinkkonen & Kinnunen 1994, 7, 21; Øvretveit 1998,1; Berk & Rossi 1999, 3.) The difference between evaluation research and basic research is the use of practical focus (specific for decision making), gathering data for the purpose of judging value and the element of comparing alternatives (Weiss 1972, 6; Øvretveit 1998, 13). Different definitions of evaluation and evaluation research exist; some emphasize effectiveness (Fink 1993) and some emphasize goal achievement (St Leger et al. 1992), depending on the viewpoint of the evaluator (Öretveit 1998, 12, 276). Some include decision making as part of their definitions, such as Weiss (1972), Øvretveit (1998) and Vedung (2003): ‘The purpose of evaluation research is to measure the effects of a program against the it set out to accomplish as a means of contributing to subsequent decision making about the program and improving future programming’ (Weiss 1972; 4). ‘Evaluation is attributing value to an intervention by gathering reliable and valid information about it in systematic way, and by making comparisons, for the purposes, of making more informed decisions or understanding causal mechanism or general principles’ (Øvretveit 1998; 9). Vedung (2003) also includes evaluation as part of practical decision making and defines the concept as the assessment of decisions, administration, outputs or public sector results (Vedung 2003, 3).

3.2 Effectiveness as a concept

In the literature, the definition of the concept of effectiveness depends on the scientific discipline and paradigm the author represents (Axelsson & Engström 2001; Meklin 2001, 107; Konu et al. 2009; Silvennoinen-Nuora 2010, 80). The common element in the definitions is the observed change in a measure compared to the desired target or baseline (Berk & Rossi 1999, 5; Meklin 2001, 107; Koskinen-Ollonqvist, Pelto-Huikko & Rouvinen-Wilenius 2005, 6–7). In administrative science and health management sciences, effectiveness is described as the impact of an action, policy, program or service on the desired outputs or on reaching the goals (Berk & Rossi 1999, 5; Lumijärvi & Jylhäsaari 1999; Vuori 2005a, 62, 66; Productivity Commission 2013). In health care, effectiveness describes the benefits of interventions measured by improvements in health outcomes in a typical population, such as in a general hospital or treatment centre setting (Smith 2005, 14). In health economics, effectiveness is defined as a net change in the outcome of a health state or QoL in normal conditions (Mandelblatt et al. 1997; McGuire 2001, 8–14; Sintonen & Pekurinen 2006, 53–55). Effectiveness indicators are often described as the outputs or end results, for example QoL, life years gained, morbidity or mortality (Axelsson & Engström 2001; Konu et al. 2009).

Effectiveness may consist of various effects and may include several factors. For example, it can be described as below according to Vedung (1997) and Meklin (2001).

1. Effects on clients and society: Actions or interventions focused on clients might affect taxpayers.
2. Effects on quantity and quality: There are enough services, but quality is poor, and vice versa.
3. Subjective and objective effects: Effects depend on the perspective of whoever is observing them (e.g. clients or professionals).
4. Long-term and short-term effects: Costs generate the first, and benefits occur later.
5. Effects appear in another sector or another municipality: For example, the long-term benefits of exercise are seen in the social and health care sectors.

6. Positive and negative effects: For example, in budgeting positive effects are emphasized, but negative effects are ignored.
7. Expected and unexpected effects: Some unexpected and negative effects can appear.

The concept of effectiveness is widely used, but it has been defined differently in different studies (Konu et al. 2009; Simonen et al. 2011; Pohjola 2012, 9; Simonen 2012, 9; Klemola 2015, 41). The concept seems to be difficult to concretize, and it is often understood in a simplistic way. Effectiveness is commonly associated with the treatment outcome, such as the effects on operation, goal orientation and costs. In political–administrative decision making, it is primarily used in defining goal states and in justification issues (Simonen et al. 2011; Simonen 2012, 9). Clear definition and expression 'compared with what', is essential in measuring effectiveness (Berk & Rossi 1999, 5; Konu et al. 2009).

In this study, effectiveness is compared between two research groups using a randomized controlled trial (RCT) design and is understood as the benefit of intervention with multidimensional effects—effects on patients (clinical outcomes and clinical endpoints), effects on society (social and health care costs), quality and quantity effects (cost-effectiveness and health-related quality of life [HRQoL]), short- and long-term effects (1-year and 8-year follow-ups) and unexpected effects in the qualitative part of the study (see Meklin 2001; Smith 2005; Productivity Commission 2013). Effectiveness measures are presented in Chapter 5.5 and Table 2 in page 49.

3.3 Approaches of evaluation of effectiveness in health care

In health care, evaluation typically focuses on treatments, services, policies or organizational interventions at different levels (individual, population, large population, system level). Various aspects can be measured, such as attitudes, values, knowledge, behaviour, budgetary allocations, agency service patterns and productivity (Weiss 1972, 39; Øvretveit 1998, 1, 17; Koskinen-Ollonqvist et al. 2005.) The evaluation of effectiveness in health care can also be considered either from a service system, a single-process (e.g. Silvennoinen-Nuora 2010) or an intervention perspective. Essential from a management viewpoint is to evaluate costs and productivity, accessibility, patient-centeredness, cost-effectiveness (costs related to accessibility, patient-centeredness, effectiveness, quality and equity), effectiveness, quality and safety. In measuring quality and effectiveness, the timeline is key; results can be acceptable and quality can be high in the short term, but measurable effectiveness can only be verified in the long term (Lumijärvi 1999; OECD 2013, EGHSPA 2016, 24; Hämäläinen et al. 2016, 10–11). This is a common problem in evaluation; the development of outcomes is often slow, and one cannot wait long enough for the true effect. Additionally, single effects are easier to evaluate but do not produce enough information for a service system perspective (Räsänen et al. 2006; Kettunen et al. 2017, 9.)

However, evaluating effectiveness is essential in health care; the fragmented service system with limited resources requires much effort to enhance efficiency and productivity. The adoption of effectiveness evaluation has been slow in health care, as the service system is complex and performance measurement varies across patient segments/disease categories. The definition of effectiveness and the evaluation of effectiveness is ambiguous in the broad context, where environment, circumstances and peoples' behaviour affect the final outcomes. (EGHSPA 2016, 24; Häkkinen & Peltola 2016, 66, 82–83.)

3.3.1 Experimental design with economic and non-economic approaches

Øvretveit (1998) presents four perspectives of evaluation in health care: the experimental, the economic, the developmental and the managerial (Øvretveit 1998, 33). This study is based on experimental evaluation with an RCT design and investigates whether the intervention results in any improvements in health or resource use. Effectiveness was evaluated from an economic (cost-effectiveness) perspective and a non-economic (goal achievement and costs) perspective, according to the classification of Sintonen and Pekurinen (2006) (Figure 2).

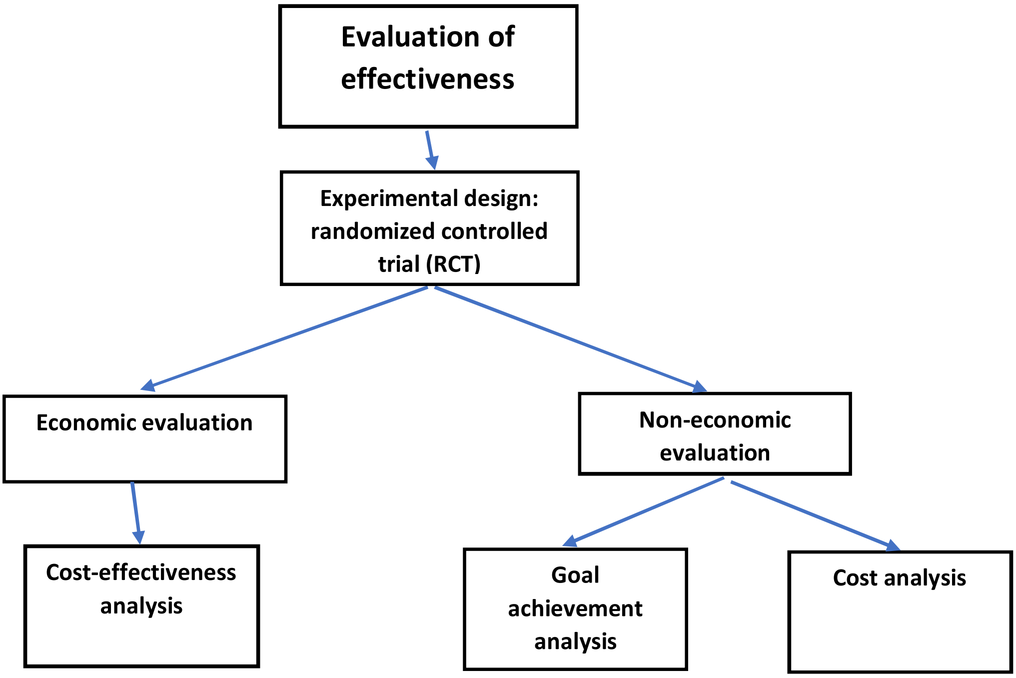


Figure 2. Evaluation frame of this study (modified from Øvretveit 1998 and Sintonen & Pekurinen 2006)

Experimental approach: Randomized controlled trial

In evaluating the effectiveness of health care interventions, an RCT is the 'gold standard'. An RCT is conventionally used in controlled circumstances to evaluate the safety or efficacy of new drugs but might be impractical or irrelevant when assessing complex interventions or changes in health service delivery. (Schulz et al. 2010; Clarke et al. 2019.) In quantitative research, randomization should always be considered to prevent selection bias if possible (Craig 2008). Intention-to-treat (ITT) analysis is used to avoid basic complications of RCTs, such as non-compliance and missing outcomes. ITT analysis includes all those patients who were originally allocated to trial ('once randomized, always analysed') and ensures comparability between groups. Therefore, the estimated treatment effect is generally conservative because dropouts and compliant subjects are mixed in the final analysis. (Gupta 2011; Brody 2016; Ranganathan et al. 2016; McCoy et al. 2017.) Per protocol (PP) analysis includes those patients who received the intervention strictly according to the study protocol. It provides an estimate of the true efficacy of intervention but might exaggerate the effect of treatment. (Brody 2016; Ranganathan et al. 2016; McCoy et al. 2017.) The importance of ITT analysis is often highlighted, but both analyses can introduce bias into the conclusions (safety and efficacy). Consolidated Standards of Reporting Trials (CONSORT) guidelines recommend that both types of analyses should be conducted for all planned outcomes, as this allows readers to interpret the effect of intervention. (Schultz et al. 2010; Brody 2016; Ranganathan et al. 2016.) Modified intention-to-treat (mITT) analysis is 'between' ITT and PP analyses; it includes fewer subjects than ITT analysis but more than PP analysis. mITT excludes some specific subjects from the ITT analysis, such as subjects who died before receiving treatment, subjects who were enrolled but who were later found not to meet the inclusion or exclusion criteria and subjects not taking all scheduled drugs or participating in the intervention. (Brody 2016.) According Abraha and Montedori (2010), approximately 50% of clinical trials employ mITT analysis using various types of descriptions of exclusion.

Economic evaluation

Economic evaluation has two main features. First, it deals with both the inputs and the outputs, which can be described as the costs and the consequences. Second, it is concerned with choices. Therefore, economic evaluation is the comparative analysis of alternative courses of action in terms of both their costs and consequences. Economic evaluation is often used along with experimental evaluation (Øvretveit 1998, 109; Sintonen & Pekurinen 2006, 250; Drummond et al. 2015, 3–4.) Economic evaluation methods include cost-benefit analysis (CBA), cost-effectiveness analysis (CEA) and cost–utility analysis (CUA). In CBA analysis, the costs and consequences of an intervention are expressed in monetary terms. In CUA analysis, costs are compared with quality-adjusted life years (QALYs), which quantify the strength of people's preferences for health states as defined by the HRQoL measure. (Hawthorne et al. 2001; Sintonen & Pekurinen 2006, 249–254; Drummond et al. 2015, 5–10.)

This study employs CEA and CUA to estimate the costs and health gains of alternative interventions. Effectiveness is evaluated using, for example, disease-specific instruments or life-years gained. The cost effectiveness of health care intervention is determined by calculating the incremental cost-effectiveness ratio (ICER), the difference in cost between two possible interventions divided by the difference in their effect. A common way to compute the ICER is to calculate the cost-effectiveness acceptability curve (CEAC), which indicates the probability of the cost-effectiveness of the intervention at different levels of willingness to pay (WTP) for the additional health outcome. (Van Hout 1994; Hawthorne et al. 2001; Sintonen & Pekurinen 2006, 249–254; Drummond et al. 2015, 5–10.) Instruments for economic evaluation can be divided into disease-specific and generic instruments. Disease-specific instruments measure the experience of a particular illness or condition or its treatment, and measures are likely to be sensitive to change. Generic measures are designed for general purpose usage and are not linked to particular diseases or treatments (Kind 2001). According to Drummond (2001), 'A disease-specific scale may have the maximum responsiveness to change, whereas the 'utility' or preference-based measure may have the potential to influence public policy and resource allocation decisions, as it enables quality-adjusted life-years (QALY) to be calculated' (Drummond 2001, 347). In comparing a disease-specific instrument to a generic instrument, for example in diabetes patients, generic instruments have been seen as more informative than disease-specific ones (Parkerson et al. 1993; Anderson et al. 1997). Although QALYs are considered one of the most important measures of effectiveness in health care, studies relating to HRQoL are limited (Räsänen et al. 2006).

Examples of the most used generic instruments to measure QoL are the EQ-5D, HUI, SF-36 and 15D questionnaires, which are available in a number of languages (Drummond 2001). These instruments measure different dimensions of QoL, such as physical and social functions, pain, emotions and sense functions (Hawthorne et al. 2001). In this study, 15D questionnaire was used to measure HRQoL. The 15D instrument measures HRQoL among adults over 16 and consists of 15 items—mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort and symptoms, depression, distress, vitality and sexual function. It is mostly used to measure single interventions, such as to evaluate drugs, surgical procedures, rehabilitation and interventions in internal medicine but is also used in national surveys, for example, in Finland and Denmark (Sintonen 2001.) In an economic analysis comparing the sensitivity of EQ-5D, SF-6D and 15D in patients with T2D, especially those suffering from CAD and diabetic retinopathy, 15D is recommended (Kontodimopoulos et al. 2012). People can feel the difference in QoL with a change of 0.02–0.03 in the 15D score. 15D can be used as a profile measure or a single index number on a scale of 0–1 (0=dead, 1=completely healthy). (Sintonen 2001.) In social care, the evaluation of effectiveness is not as common (Pohjola et al. 2012). However, a generic QoL instrument, the Adult Social Care Outcomes Toolkit (ASCOT), has been developed to measure the ability to function in everyday life (van Leeuwen et al. 2015; van Loon et al. 2018).

Non-economic evaluation

Non-economic evaluation is used in this study according to Figure 2. This evaluation is divided into two categories by Sintonen and Pekurinen (2006) — goal-achievement analysis and cost analysis (CA)/cost-minimization analysis (CMA). The target of goal-achievement analysis is to find the most effective action or combination of actions despite limitations and costs. Effectiveness can be evaluated in relation to the needs of individuals or society. (Kind 2001; Smith 2005; Productivity Commission 2013.) The basis of the goal-achievement approach consists of 'what should be' criteria (usually norms, laws and other official standards) compared to 'what is' and how the desired goals have been achieved. It is important to define from whose viewpoint goals have been set (see Simon 1979). The weakness of the approach is that the evaluation might be narrow and restricted only to the desired results (Sinkkonen & Kinnunen 1994, 82–85). Evaluation studies usually aim at justifying the effectiveness of interventions, but it is difficult to get explicit 'yes' or 'no' answers, and results are often open to interpretation (Ettelt et al. 2015).

CA compares the cost of interventions, and CMA compares the costs of two similar interventions to determine which is less expensive (Hawthorne et al. 2001; Sintonen & Pekurinen 2006, 249-254; Drummond et al. 2015, 5–10). CA or cost-description analysis focuses only on costs and comparing alternative interventions, projects or treatments. It is important to specify the perspective of the analysis and what costs are included. For example, Drummond et al. (2015) suggests four categories of cost be considered—the resource use of the health sector, the resource use of the patient, the resource use of other sectors and productivity change (Sintonen & Pekurinen 2006, 250; Drummond et al. 2015, 219.)

3.3.2 Special features of evaluating complex interventions

Intervention can basically be defined as 'an action which results in change' (Øvretveit 1998, 7). In health care, non-pharmacological complex interventions are widely used, and their evaluation is not so linear compared to simple interventions (Campbell et al. 2000). According to the Medical Research Council, the characteristics of complex interventions are as follows:

1. The number of interactions between components within the experimental and control interventions
2. The number and difficulty of behaviours required by those delivering or receiving the intervention
3. The number of groups or organizational levels targeted by the intervention
4. The number and variability of outcomes
5. The degree of flexibility or tailoring of the intervention permitted (Craig et al. 2008)

Examples of complex interventions are service delivery and organization, interventions directed at health care professionals' behaviour, community interventions, group interventions and interventions directed at individual patients. Evaluating the effectiveness of complex interventions is challenging and has specific features. (Campbell et al. 2000.) Craig et al. (2008) emphasize two key questions in evaluating complex interventions: 1) Is the intervention effective in everyday practice? 2) How does the intervention work, that is, what are the active ingredients and how do they exert an effect? Realist evaluation asks what works, for whom and under what circumstances (Bonell et al. 2012; Fletcher et al. 2016). The use of integrated approaches (quantitative and qualitative methods) is particularly useful in evaluating interventions that are difficult to explore or capture using quantitative methods alone, such as interventions that involve social or behavioural processes and try to change patient or professional behaviour (Haynes 1998; Campbell et al. 2000; Oakley et al. 2006.)

Campbell et al. (2007) emphasize the importance of reporting the context in which the intervention was carried out, for example, the socio-economic background, the health care service system and the characteristics of the population. Further, a detailed process evaluation is important because the findings can explain why an intervention works or does not work or has unexpected consequences. Therefore, process evaluation should be integral to RCTs. Further, a clear description of the intervention is essential (Oakley et al. 2006; Craig et al. 2008; Moore et al. 2015.)

Health coaching interventions have characteristics of complex interventions (see Craig et al. 2008). For example, they involve social or behavioural processes and change patient or professional behaviour (Campbell et al. 2000; Oakley et al. 2006). In this study, qualitative methods were also used to explain and elucidate the effects of intervention.

4 Health coaching: concept and effectiveness

The literature search was conducted with information specialist in January and February in 2019 using systematic search strategy. Terms “health coach*” AND “effect*” was used in searches according to article title, abstract and keyword from the following databases: Scopus (356), PubMed (327) and Cinahl (206). Altogether 171 articles remained after removing duplicates. In Scopus, the first publication of health coaching was 1990 and from 1990 to 2012 less than 20 publications were published each year and after that the number of publications has been steadily grown until 2018 where 79 studies had been published. All 171 articles found were reviewed according to titles and abstracts. Then articles were divided three categories: definition of health coaching, effectiveness of clinical and behavioral changes and effectiveness of health care utilization and costs. The most essential articles with relation to effectiveness of health coaching (25) have been described in Appendix 1.

According to literature review self-management support interventions are promising approaches to manage chronic disease (Trappenburg 2013; Kivelä et al. 2014; Panagioti et al. 2014). In the last few decades, the use of behavioural change theories and models has increased in chronic care management (Butterworth et al. 2007). However, the difficulty of adhering to treatment still exists (World Health Organization 2003). It is important to distinguish three concepts that have been widely used in relation to self-management; these are compliance, adherence and concordance. In the 1970s, the concept of compliance was occurred in medicine. It is considered to indicate an authoritarian relationship between professional and patient, suggesting that patients passively follow orders. Such an authoritarian relationship might weaken patients’ self-efficacy and thus their capability to care for themselves. (Bell et al. 2007.) The concept of adherence has most commonly been defined as ‘the extent to which patients follow the instructions they are given for prescribed treatments’ (Bissonnette 2008, 634), but it also emphasizes communication, cooperation and partnership in decision making between health care professionals and patients (Bell et al.

2007; Gardner 2008). The concept of concordance can be used in relation to a coaching relationship that emphasizes the importance of communication and interaction, for example patients as equal partners with health carers. Patients are seen as experts on their own lives (Routasalo et al. 2009; Gardner 2014.) Bell et al. (2007) suggest that concordance is synonymous with patient-centred care. The transition from professional-based compliance thinking to patient-centred concordance thinking depends on a change of mindset on the part of professionals, decision makers and patients (Routasalo et al. 2009).

Conventional chronic disease management programs aim to improve patients' self-management skills in increasing treatment adherence, such as keeping appointments with health care professionals and taking prescribed medicines. They focus more on the disease itself (diagnoses, complications or symptom management), emphasizing coordinated and comprehensive care pathways and algorithms built upon evidence-based clinical guidelines, and focus less on the patient's individual needs or behaviours. (Ellrodt et al. 1997; Wagner et al. 2001; Mattke et al. 2007.) Butterworth et al. (2007) discuss the differences between traditional health education and ideal health coaching. Traditional education is task-oriented, provides advice and shares information based on structured assessment and a treatment adherence plan. The aim is to manage disease and its complications. Ideal health coaching is client-oriented and empathetic. It supports self-efficacy and takes a whole-person approach in which behaviours are prioritized for maximum impact on overall health. As late as the 1990s, the nursing literature referred to nursing coaching as a practice framework that complements patient teaching and supportive therapy.

Olsen (2014, 24) defines health coaching as 'a goal-oriented, client-centred partnership that is health-focused and occurs through a process of client enlightenment and empowerment'. It is based on a partnership between the coach and the individual, shared decision making (a decision is reached together with professional and patient) and collaborative goal setting facilitated by motivational interviewing (Palmer et al. 2013; Olsen 2014; Hale & Giese 2017). Wolever et al. (2013) emphasize that the patient-centred process is based upon behaviour change theory and is delivered by health professionals with diverse backgrounds. Usually, health coaching is conducted by certified health coaches or specially trained health care professionals (Olsen 2014). The role of the health coaches' involves active listening, understanding, facilitating, applauding, supporting, motivating, providing feedback and helping patients to weigh options, make choices and identify and overcome challenges in the process of changing for the better (Lindner et al. 2003; Hayes 2008).

Generally, the evidence on the effectiveness of health coaching has mainly been evaluated in the short term. The follow-up time of most studies has been 12 months or less, and only a few studies have evaluated effectiveness in the long term. According to reviews by Dennis et al. (2013) and Kivelä et al. (2014), health coaching is effective in terms of physiological, behavioural, psychological and social outcomes. However, the systematic review findings of Dejonghe et al. (2017) were mixed, with follow-up times of 24–98 weeks for rehabilitation and prevention. Three of seven studies for each setting found statistically significant effectiveness. The findings were also mixed in single effectiveness studies, particularly in terms of clinical outcomes (Vale et al. 2003; Wolever et al. 2010; Karhula et al. 2015; Sherifali et al. 2015; Wayne et al. 2015; Willard-Grace et al. 2015; Sharma et al. 2016; Tiede et al. 2017; Chapman et al. 2018; Tuluce & Kutluturkan 2018; Panagioti et al. 2018). From a health care utilization and economic viewpoint, particularly cost-effectiveness, the effectiveness of health coaching has been found to be limited (Hale & Giese 2017). The evidence on effectiveness mainly suggests that health coaching does not reduce health care utilization or result in cost savings in the short term (Wennberg et al. 2010; Hutchison & Breckon 2011; Lin et al. 2012; Steventon et al. 2013; Benzo et al. 2015; Billot et al. 2015; Jonk et al. 2015; Härter et al. 2016; Wagner et al. 2016; Hale & Giese 2017; Scuffham et al. 2018). Studies with long-term follow-up are rare. In one example, Byrnes et al. (2018) achieved significant reduction in overall mortality and lower total health insurance costs in a 6.35-year follow-up due to health coaching.

5 Materials and methods

5.1 Target group and patient identification

The present study included 12 municipalities with a total population of 112,000. Altogether 5500 participants (4.9%) were identified from electronic patient records in secondary care according to the laboratory inclusion criteria based on The Finnish Current Care Guidelines. A research nurse verified that the patients met the inclusion and exclusion criteria by checking their medical records and 2594 patients (2.3%) met the inclusion criteria (Table 1). Patients with more than one disease were allocated to their most prominent disease group using the following hierarchy: 1) CHF; 2) CAD; 3) T2D. An information letter and a consent form were sent to all eligible patients in four batches during 12 months in 2007–2008, and 1535 patients (59.2%) gave consent and enrolled in the study.

Table 1. Inclusion and exclusion criteria for the TERVA program.

Inclusion criteria:	**Exclusion criteria:**
• Resident in the region of Päijät-Häme aged 45 years or older • One of the following diagnoses: • Heart failure of class II or III according to the New York Heart Association (NYHA) classification and a history of hospital admission for heart failure within the last 2 years • History of myocardial infarction or cardiac revascularisation procedure and one of the following (treated or untreated): blood pressure above 140/85, total serum cholesterol concentration >4.5 mmol/L, serum low-density lipoprotein concentration >2.5 mmol/L • Type 2 diabetic on medication and serum HbA1c >7% without clinically evident cardiovascular diseases (e.g. MI, stroke, peripheral vascular disease)	• Inability to cooperate or participate • Pregnancy • Life expectancy less than 1 year • Patients with major elective surgery planned within 6 months • Patients have had major surgery within the last 2 months

5.2 Randomization

A total of 2594 eligible patients were randomized using a Zelen design with a 2:1 ratio for the intervention or control group; the ratio imbalance was considered in the statistical power calculations. In a Zelen design, eligible patients are randomized to either an intervention group or a control group before consent to avoid disappointment bias and subjective bias in the recruitment process (Homer 2002). To ensure balanced distribution within disease groups and different municipalities between the study groups, stratified randomization with permuted blocks was used. The algorithm was based on computer-generated random numbers. The study group was informed to the patients by health coaches after the initial measurements.

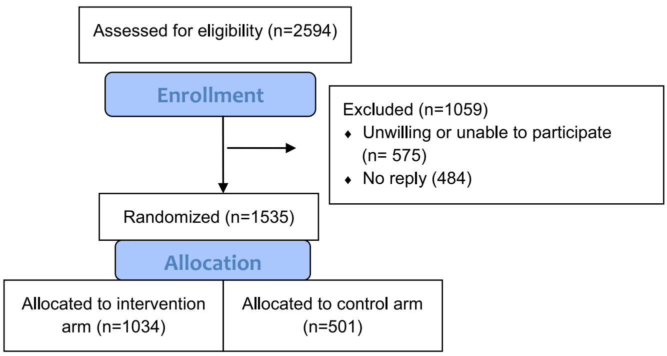


Figure 3. Randomization of the telephone health coaching study.

5.3 Intervention

All the health coaches worked in health coaching centre which was set up in the city of Lahti. Seven experienced (at least five-year work-history) registered nurses and public health nurses from secondary health care (hospitals), occupational health care and primary health care (health centres) were hired as full-time health coaches. They received a 4-week training program in motivational interviewing techniques (trained by a psychologist specializing in strength-based behaviour), lifestyle change coaching, the content of a telephone health coaching program and call centre technology.

Self-regulation theory (integrated behaviour changes components and behaviour change techniques) supported by evidence (that is, self-monitoring, goal setting, action planning and feedback) was used (see Michie et al. 2009). The intervention included eight key recommendations for patients developed by Pfizer Health Solutions that were modified for the Finnish health care system, as follows.

Patients:

1) know how and when to call for help
2) learn about the condition and set goals
3) take medicines correctly
4) get recommended tests and services
5) act to keep the condition well
6) make lifestyle changes and reduce risk
7) build on strengths and overcome obstacles
8) follow up with specialists and appointments

Self-management booklets prepared in collaboration with the Finnish Heart Association and the Finnish Diabetes Association were sent to patients to support progress towards the key recommendations. Health coaches utilized technology and a traffic light system to track patients' progress in relation to the key recommendations. Health coaches were not fully integrated into the care teams in primary care but had access to and the possibility to document patient health and coaching status in the primary and secondary care electronic health records (EHRs).

The patients in the intervention group were called approximately 10–11 times over 12 months. The calls were of four types—brief engagement calls, broader needs assessment calls, monthly coaching calls and evaluation calls. It was possible to have a brief follow-up call between the coaching calls if needed. Quality control of the length, frequency and content of calls was performed, and the coaches were tutored individually and in groups throughout the intervention by a psychologist. The first quality assurance measures were taken after 2 months; calls were typically long, up to 60 minutes, and were based on a coach-driven information provision model with very little concrete goal setting and action planning. To improve the quality of coaching, an explicit structure according to a self-regulation model was developed with the coaches, and the maximum number of topics during one call was limited to three. During the individual tutoring sessions, recorded coaching calls were reviewed to identify and strengthen key coaching skills, such as active listening, posing open questions, reflection and summarizing the patient talk (Rollnick, Miller & Butler 2007). The length of calls decreased by approximately 30 minutes (50%) as a result of the quality assurance actions together with coaches.

5.4 The usual care

The treatment of T2D and CAD in Finland is based on The Finnish Current Care Guidelines, which are independent, evidence-based clinical practice guidelines (Type 2 Diabetes; Cardio Vascular Disease: Current Care Guidelines Abstract, 2020). Primary care is responsible for chronic care provided by general practitioners together with nurses, some of which specialize in single diseases such as asthma, diabetes, CVD or mental health problems. T2D patients typically have between two and six planned annual visits to a doctor or nurse depending on how well the disease is being controlled. The wards in primary health care provide basic care for patients with less severe cases and those who are unable to cope at home. The treatment of CAD patients is mainly provided in secondary care, in addition to one or two primary care visits per year (Cardio Vascular Disease: Current Care Guidelines Abstract, 2020). In general, patients with these conditions who have complications are treated for acute needs in secondary care hospitals and transferred back to primary care. Those patients who need long-term care (LTC) receive home care, care at service home facilities or nursing homes and in-patient care at the primary care level.

5.5 Data collection and analysis

This study consists of four sub-studies (I–IV) and additional interview material. Multiple methods of data and material collection and analysis were used. A summary of the aims of the sub-studies, the effectiveness measures, the data collection and the analysis of each phase is presented in Table 2.

Table 2. The aims, evaluation measures, data sources and analysis of Sub-studies I–IV and additional interview material in health coaching study.

The aim of the study	Evaluation of effectiveness	Measures		Data/material source	Analysis
To evaluate the effectiveness of health coaching on clinical outcomes in the short term (one year) (Sub-study I)	Goal-achievement analysis	Clinical outcomes (1 year): Blood pressure, waist circumference, Glycated haemoglobin (HbA1C) (Type 2 diabetes patient), Serum total cholesterol (S-Kol), Serum Low-density lipoprotein (SLdl), New York Heart Association - classification (NYHA class) (Heart Failure patient)	10-percentage point increase the proportion of patients reaching the target level in the intervention group	Electronic health records (EHR) laboratory measures, and clinical measurements by research nurse	Modified intention to treat (mITT) Statistical analysis: Chi square tests, regression analysis
To evaluate cost-effectiveness of health coaching in the short term (one year) (Sub-study II)	Cost-effectiveness analysis	Cost-effectiveness (1 year): Quality of life (15D) and total social and health care costs	Incremental cost-effectiveness ratio (ICER) = Difference in costs (C=€) divided by the effect (E=QoL) (ICER = (C1–C2)/(E1–E2), (all, Type 2 diabetes, coronary artery disease)	15D questionnaires, National registries: National Discharge Registry (HILMO) and Care Registers for Social Welfare (SosHILMO), National price list for unit costs of health care services in Finland	Modified intention to treat (mITT), Differences in the mean costs and outcomes and corresponding cost-effectiveness ratio (ICER), cost-effectiveness acceptability curve (CEAC)

The aim of the study	Evaluation of effectiveness	Measures		Data/material source	Analysis
To evaluate the impact of health coaching on social and health care costs in the long term (eight year) (Sub-study III)	Cost analysis	Total social and health care costs (8 yrs): primary care visits, primary care wards, secondary care outpatient, secondary care inpatient, home care, service homes, nursing home	Incremental costs €/patient (all, Type 2 diabetes, coronary artery disease)	National registries: National Discharge Registry (HILMO) and Care Registers for Social Welfare (SosHILMO), National price list for unit costs of health care services in Finland	Intention to treat (ITT) and per protocol (PP) analysis, differences in mean costs (bootstrapped 95% Confidence Interval), cumulative cost curves for each research group
To evaluate the effectiveness of health coaching on clinical endpoints in the long term (eight year) (Sub-study IV).	Goal-achievement analysis	Morbidity and mortality (8 yrs): Primary outcome: A composite variable of death from cardiovascular causes or stroke or myocardial infarction (AMI) or unstable angina pectoris (UAP). Secondary outcomes: 1) A composite variable for cardiovascular mortality or morbidity 2) and events of death from any cause, AMI, stroke, renal insufficiency, peripheral vascular disease 3) and hospitalization due to congestive heart failure (CHF).	Hazard Rate (HR)= Hazard in the intervention group/ Hazard in the control group	National registries: National Discharge Registry (HILMO) and Care Registers for Social Welfare (SosHILMO)	Intention to treat (ITT) and per protocol (PP) analysis. Statistical methods: Chi square and t-tests, Cox proportional hazard regression, Kaplan-Meier estimator curve
To describe health coaches' perceptions the factors that associated with the effectiveness of health coaching (additional results in summary).	Qualitative analysis	Semi-structured interview	Interview themes: adaptation coaching skills, effect factors of health coaching, interaction with health care providers and implementation	Interview material from health coaches in TERVA-program	Inductive content analysis

Clinical outcomes (I)

The clinical outcomes after 1 year of health coaching were evaluated in Sub-study I, using the principles of goal-achievement analysis. The outcomes were systolic and diastolic blood pressure, serum total and LDL cholesterol concentration, waist circumference, glycated haemoglobin (HbA1c) for T2D patients and NYHA class for CHF patients. The target level was defined according the Finnish Current Care Guidelines, and the target effect was a 10-percentage point increase in the proportion of patients reaching the target level in the intervention arm. Data was collected using two methods; research nurses measured patients' weight, blood pressure and circumference, and laboratory test results were collected directly from EHRs in secondary health care.

A mITT analysis was performed; the analysis included data from all those patients with data given upon entry and at the end of the 1-year follow-up. Next, data of 1250 patients were calculated to provide adequate statistical power to find a 10-percentage point difference between the intervention arms. Differences between research groups were evaluated using statistical methods. A significance level value of ≤ 0.05 was used (see Tähtinen, Laakkonen & Broberg 2011, 92; Holopainen & Pulkkinen 2012, 165).

Cost-effectiveness (II)

The aim of Sub-study II was to evaluate the cost-effectiveness of 1 year of health coaching.

HRQoL was measured using the 15D generic self-administered instrument (Sintonen 2001). 15D questionnaires were sent to the patients in the intervention and control groups at the beginning of the intervention and 1 year after. The mITT method was used.

Cost and utilization data were collected from the national registries maintained by the National Institute for Health and Welfare. The hospital benchmarking database, the National Discharge Registry, includes secondary care data (the use of hospital outpatient care, all types of outpatient visits and hospital admissions) related to diagnoses (diagnosis-related grouping, DRG) and Care Registers for Social Welfare that includes all types of long- and short-term institutionalized care, housing and residential services and home care services. Primary care data were collected from the primary health care EHRs from 2007 until 2011, after which the EHRs were integrated into national registries (AvoHilmo) that provided data for 2012–2016. The use of a unique social security code enabled full linkage to the national registries providing comprehensive data about each individual's use of social and health care.

The DRG cost weights for hospitalizations and outpatient visits were based on individual-level cost-accounting data from several hospitals. The unit cost estimates for social care encounters and bed days were derived from the national price list for unit costs of health care services in Finland (Kapiainen et al. 2011). Extracting the patient-level data from the patient administration systems (with diagnosis and activity information) made it possible to group each individual encounter type using the Ambulatory and Primary Care Related Patient Groups (APR) grouper, a grouping system equivalent to the DRG used in hospital care (Honkasalo et al. 2014). The APR groups were supplemented with cost weights indicating the relative consumption of resources. Cost weights were based on large samples of time measurements in primary care contacts and procedures to compile a relative value scale. All costs were deflated using the price index for public health care provided by Statistics Finland.

Data was bootstrapped by generating 1000 replicates. Bootstrapping is a data-based simulation method assessing statistical precision. The observed sample is chosen randomly from an unknown probability distribution. The differences in mean costs and outcomes and the ICER were analysed. It was completed by calculating the CEAC derived from the bootstrap replicates. CEAC indicates the probability of the cost effectiveness of the intervention at different levels of WTP for the additional health outcome (Van Hout et al. 1994).

Social and health care costs (III)

The aim of Sub-study III was to evaluate the effectiveness of telephone health coaching on social and health care costs. (In this summary, the term 'social care cost' is used instead of 'long-term care', which was used in Sub-study III). Cost analysis included the costs of different service types—primary care visits, primary care wards, secondary care outpatient, secondary care inpatient, home care, nursing home and service home in the 8-year follow-up. Cost and utilization data based on national registries was collected between 2007 and 2015, as well as on Sub-study II. These registries enabled using both ITT and PP analysis in the cost analysis. ITT analysis included those patients who were originally allocated to the intervention and control groups. In PP analysis, patients who did not participate in any activities related to the study after giving their consent, for example, those who did not return study questionnaires or participate in the clinical measurements in Sub-study I, were excluded.

Differences in mean costs between research arms were calculated. In assessing the statistical significance of differences, non-parametric bootstrapping was used, calculating 95% confidence intervals (CIs) for incremental total costs at 8 years of follow-up. The cumulation of cost over time was assessed by drawing cumulative cost curves for each research arm. To assess the consistency of the intervention effect in different patient groups, the total health care costs were calculated for T2D and CAD sub-groups and for different service types.

Mortality and morbidity (IV)

The aim of Sub-study IV was to evaluate the effectiveness of telephone health coaching on mortality and morbidity in the long term by comparing study groups. The primary outcome was the first occurrence of a composite cardiovascular variable, that is, death from cardiovascular causes, non-fatal stroke, non-fatal myocardial infarction (AMI) or unstable angina pectoris (UAP). Secondary outcomes were death from cardiovascular causes or stroke or AMI; death from any cause or stroke or AMI; and death from any cause or stroke or AMI or UAP, coronary artery bypass grafting (CABG), percutaneous transluminal coronary angioplasty (PTCA) or HF, or peripheral vascular disease (PVD). The other outcomes were death (all causes), AMI (fatal or non-fatal), stroke (fatal or non-fatal), renal insufficiency, PVD and hospitalization due to CHF. (see the Look AHEAD -research group 2013).

Data was collected between 2007 and 2015 from the Finnish national registries based on the unique identification code (see Sub-studies II and III). Data was linked the patient cohorts to the registers and retrieved comprehensive data on all diagnoses, diagnostic and treatment procedures, service contacts in social and health care and mortality for each individual based on the International Classification of Diseases 10th revision (ICD-10) codes; the Finnish version of the Nordic Classification of Surgical Procedures (NCSP) codes for diagnostic and treatment procedures; and the respective NordDRG patient grouping classifications. The registries included the hospital benchmarking database, the national discharge registry (HILMO), the Hospital Discharge Register and the Cause of Death Register by Statistics Finland.

Baseline characteristics in the intervention and control groups were tested using chi square and t-tests. Cox proportional hazard regression was used to compare the risk (hazard rate, HR) of primary and secondary endpoints between the intervention and control groups. The Kaplan–Meier estimator curve was used to report the proportion of patients who had an event in the primary endpoint, and Cox regression was used to report HRs and the 95% CI for each endpoint. All statistical analyses were conducted using Stata version 15.0.

Health coaches' perceptions of the telephone-based health coaching interventions

Seven health coaches were interviewed in semi-structured interviews to better understand the health coaching intervention and to obtain their perceptions on the effect factors of the intervention (see Lewin et al. 2009). Interview themes arose from the health coaching process and the preliminary results of the study (see Tuomi & Sarajärvi, 2009, 75). Interview themes were adoption coaching skills, the effect factors of health coaching, the perceptions of interactions between health coaches and usual care professionals and the implementation. Interview themes were sent to the coaches in advance because the TERVA project ended almost 10 years prior to the interviews. Interviews were carried out during autumn 2017 and spring 2018 and lasted 30–90 minutes. All interviews were recorded and transcribed verbatim by an external transcriber. Interview material included 103 pages using 12-pt font and 1.0 line spacing. Interview material was analysed using the principles of inductive content analysis—simplification, grouping and abstraction. Interview material was read through several times. The meaningful expressions describing the perceptions of health coaches on factors of health coaching intervention were identified from the material. Original expressions having the same meaning were classified into sub-categories according to the themes, and main categories were formulated (Tuomi & Sarajärvi, 2009, 96–97, 117–118).

6 Results

A flow chart of the study (the number of patients analysed, data sources and analysis methods in Sub-studies I–IV) is presented in Figure 4.

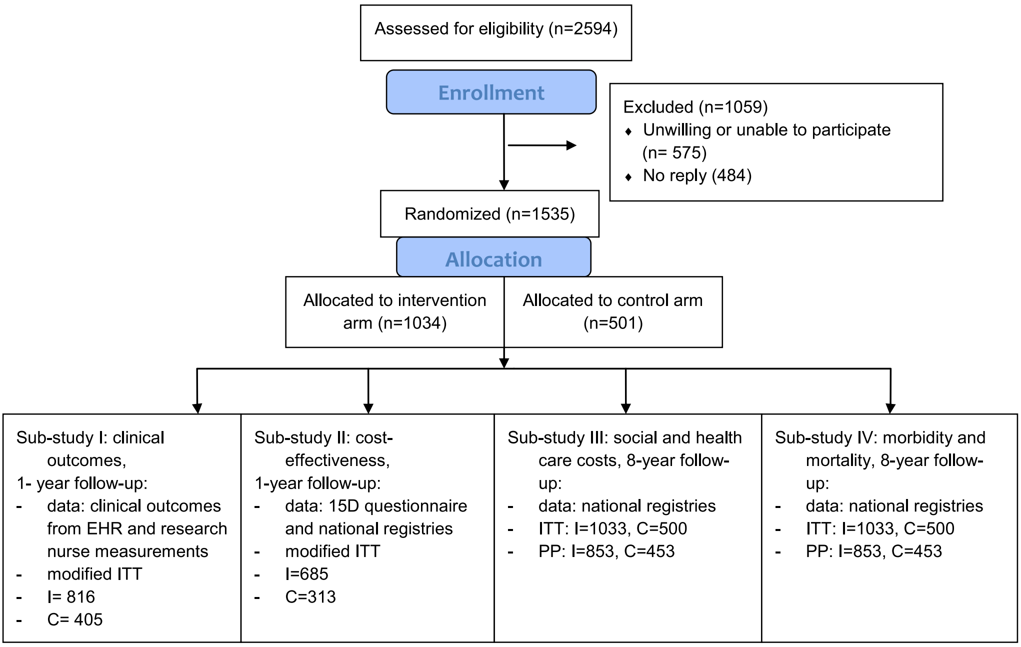


Figure 4. Flow chart of the health coaching study (I=intervention group, C= control group, EHR=electronic health record, ITT=intention to treat, PP=per protocol).

The baseline of the study is presented in Table 3, and there were no significant differences between the intervention and control groups. In Sub-study I, the baseline was also tested by disease group, and there were no differences in age, sex, self-reported duration of disease and age at diagnosis, blood pressure (systolic, diastolic), total cholesterol, high density lipoprotein (HDL), LDL, body mass index (BMI), waist circumference, daily smokers, lipid lowering medication, HbA1c, oral antidiabetic drug and insulin, oral antidiabetic drug, insulin and NYHA class between the intervention and control groups (I). Socio-economic status was not asked about in the questionnaires.

Table 3. Characteristics of the patients at baseline in telephone health coaching study (ITT=intention to treat, PP=per protocol).

	ITT intervention (N=1033)	**ITT control (N=500)**	**PP intervention (N=853)**	**PP control N=453**	**Non-co-operators intervention (N=180)**	**Non-co-operators control (N=47)**
Age, mean years	65.0	65.4	64.9	65.1	65.1	67.9
Age over 75(%)	14.1	16.2	12.7	15.0	21.1	27.6
Male sex (%)	60.6	57.8	61.8	58.5	56.1	61.7
Type 2 diabetes, N (%)	769 (74.3)	355 (71.0)	637 (74.7)	324 (71.5)	132 (73.3)	35 (74.4)
Coronary artery disease, N (%)	172 (16.6)	97 (19.4)	145 (17.0)	90 (19.8)	27 (15.0)	7 (14.9)
Congestive heart failure, N (%)	92 (8.8)	43 (8.6)	44 (8.8)	39 (8.6)	21 (11.6)	5 (10.6)
Multimorbidity (%)	48.8	48.0	48.1	48.7	52.8	40.4
Number of chronic conditions (mean)	1.76	1.67	1.72	1.66	2.0	1.78
Blood pressure (systolic), mean	141.9 (N=812)	143.5 (N=404)	-	-	-	-
Blood pressure (diastolic), mean	83.5 (N=812)	84.7 (N=404)	-	-	-	-
Body mass index, mean	25.5 (n = 404)	24.5 (n = 812)	-	-	-	-
Quality of life (15D), mean	0.861 (n = 470)	0.843 (n = 962)	-	-	-	-
Glycated haemoglobin (%)	7.7 (n = 224)	7.5 (n = 415)	-	-	-	-
Serum total cholesterol (mmol)	4.4 (n = 93)	4.3 (n = 250)	-	-	-	-
Serum high-density lipoprotein (mmol)	1.26 (n = 93)	1.26 (n = 245)	-	-	-	-
Serum low-density lipoprotein (mmol)	2.35 (n = 91)	2.23 (n = 245)	-	-	-	-

6.1 Short-term effectiveness

6.1.1 Clinical outcomes

In Sub-study I, the analysis included 1221 patients (80%) having data on primary endpoints both at entry and at the end of follow-up. Laboratory measures of lipids were available in EHRs only for a fifth of the patients, and HbA1c measures were available only for 54% of the T2D patients. The follow-up period began from the beginning of health coaching in the intervention group and from the date the study group was informed for the control group. The follow-up time ended 1 year after.

A significant difference was found in diastolic blood pressure. A decrease to 85 mmHg or lower was found in 48% of the intervention group and in 37% of the control group (difference of 10.8%, 95% CI 1.5%–19.7%). There were no significant differences in any other clinical outcomes. However, the target levels of systolic blood pressure and waist circumference were reached more frequently in the intervention group.

6.1.2 Cost-effectiveness

In Sub-study II, 998 patients completed the 15D questionnaire at the beginning and at the end of the 1-year follow-up. The cost data was obtained for all patients using the National Discharge Registries. The number of patients by sub-group was (I=intervention/C=control); CHF group: 56/27; CAD group: 124/68 and T2D group: 505/218.

The cost effectiveness was the greatest for patients with T2D; the ICER was €20,000 per QALY. In the CAD group, the ICER was €40,278 per QALY, and in the CHF group, costs increased with no marked effect on QoL. The overall ICER was €48 000 per QALY. An improvement of 0.008 in QoL was achieved in the T2D group, with a small increase in the cost of care (€160 per patient). In the CAD group, the improvement in QoL was higher (0.018), with an increase in the cost of care (€725 per patient). (Table 4).

Table 4. Incremental costs, quality of life and cost-effectiveness ratios in the disease sub-groups and in the whole study group in telephone health coaching study (II).

	Cost (€), mean (95% CI)			**QoL (15D), mean (95% CI)**			**ICER (€/QALY)**
	Interven-tion	**Control**	**Incre-mental cost**	**Interven-tion**	**Control**	**Incre-mental effect**	
Type 2 diabetes	1948 (1673–2222)	1788 (1204–2371)	160 (–406–726)	0.008 (0.003–0.014)	0.000 (–0.009–0.009)	0.008 (–0.002–0.0018)	20 000
Coronary artery disease	2510 (1806–3214)	1785 (984–2585)	725 (–389–1839)	0.019 (0.007–0.030)	0.001 (–0.014–0.016)	0.018 (–0.001–0.037)	40 278
Conges-tive heart failure	4469 (1955–6983)	2214 (–105–4533)	2255 (–1669–6180)	0.013 (–0.007–0.032)	0.015 (–0.015–0.046)	–0.003 (–0.037–0.032	–
All	2256 (1940–2571)	1824 (1345–2302)	432 (–135–999)	0.011 (0.006–0.015)	0.002 (–0.006–0.009)	0.009 (0.000–0.018)	48 000

According to the cost-effectiveness plane (Figure 5), the intervention was more effective compared to care as usual but was also more costly. Regarding the bootstrapped ICERs, 89% of the data points fell into the northeast quadrant, indicating increased QoL at an incremental cost, and 9% fell into the southeast quadrant, indicating increased QoL at a decreased cost. Only 2% of the data points fell into the northwest quadrant, and less than 1% fell into the southwest quadrant.

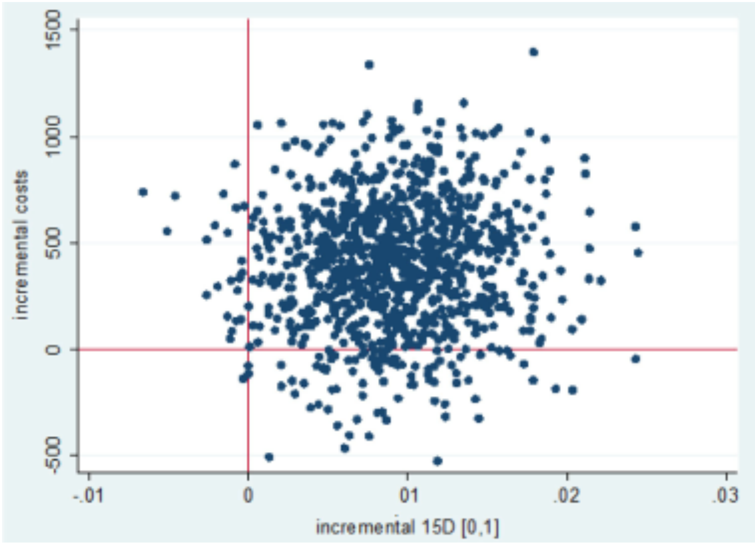


Figure 5. Distribution of bootstrapped incremental costs and health-related quality of life of telephone health coaching in 1-year follow-up (II).

CEACs are presented for all participants and for the T2D and CAD sub-groups in Figure 6. If the decision makers were willing to pay €50,000/QALY, the probability of cost-effectiveness is 55% for all patients and 75% for T2D patients.

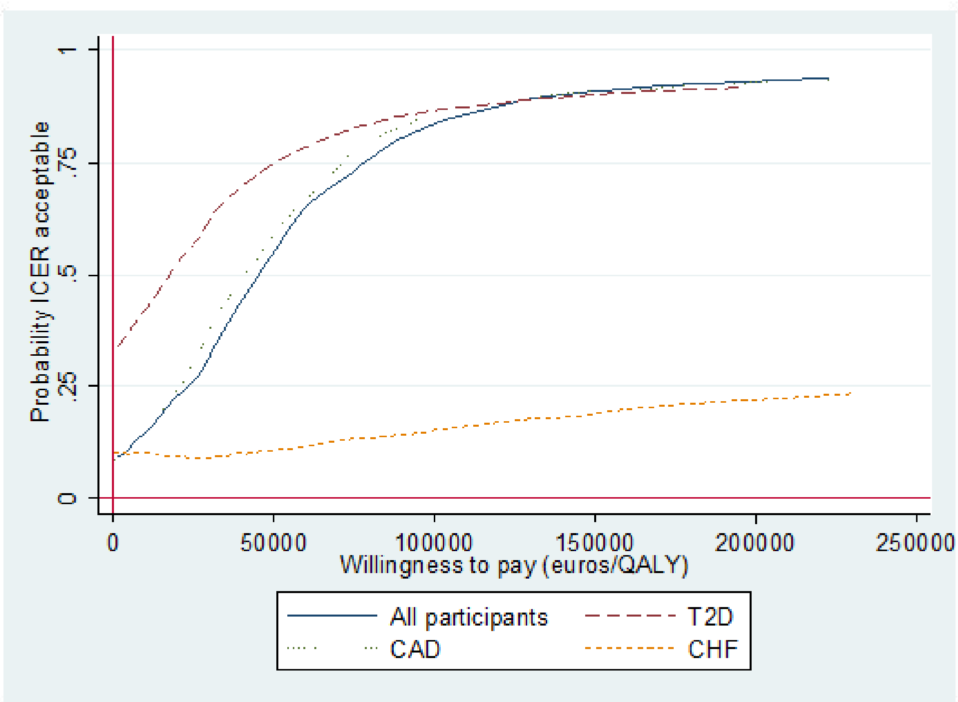


Figure 6. Cost-effectiveness acceptability curve f (CEAC) or all patients and for disease sub-groups (T2D=type 2 diabetes, CAD= coronary artery disease, CHF=cardiac heart failure (II).

6.2 Long-term effectiveness

6.2.1 Social and health care costs

In Sub-study III, the ITT analysis of the 8-year registry-based follow-up cost data included 1033 patients in the intervention group and 500 patients in the control group. One patient in each group was missing from the Finnish national registries, probably due to emigration. The PP analysis included 853 patients in the intervention group and 453 in the control group.

During the first 2 years, the cumulative costs were higher in the intervention group. After that, the costs in intervention group were lower to the end of follow-up. In the ITT analysis, the total costs were 3% (€1248) lower in the intervention group (€39,667 per patient in the intervention group and €40,916 per patient in the control group). The difference was not statistically significant (95% CI from €-6374 to €2217; p=0.2). In the PP analysis, the total cost was €35,863 in the intervention group and €41,816 in the control group; the cost saving due to the intervention was 14% (€-5953). The difference was statistically significant (p=0.02), indicating a 98% probability that the intervention is cost-saving compared to care as usual (95% bootstrapped CI from €-9842 € to €-1132) (Figure 5). Cumulative costs per patient per year and number of patients at risk per year in the ITT and PP analyses are presented in Appendix 2, Tables 1 and 2.

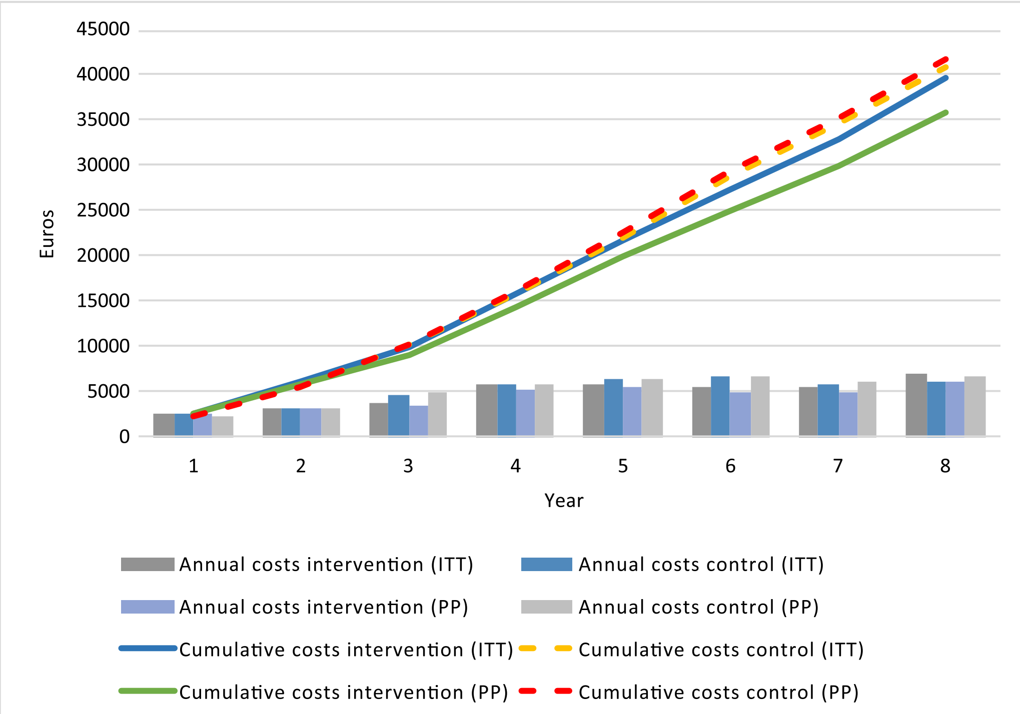


Figure 7. Cumulative and annual costs per patient in intention to treat (ITT) and per protocol (PP) analyses of telephone health coaching in 8-year follow-up (III).

In the sub-group ITT analysis, the average cost was €-3126 (7%) lower for patients with T2D. The difference was not statistically significant (p=0.18) between the study groups, whereas in the PP analysis the average cost was €-7287 (17%) lower per patient. This difference was statistically significant (95 % CI from €-12,528 to €-1760; p=0.02) with a 98% probability that the intervention was cost-saving compared to care as usual. The results were mixed among the CAD group (including HF patients); ITT analysis showed a cost increase of €3543 (10%) per patient due to the intervention, and PP analysis showed a cost saving of €-3101 (8%). The results were not statistically significant for either of the groups.

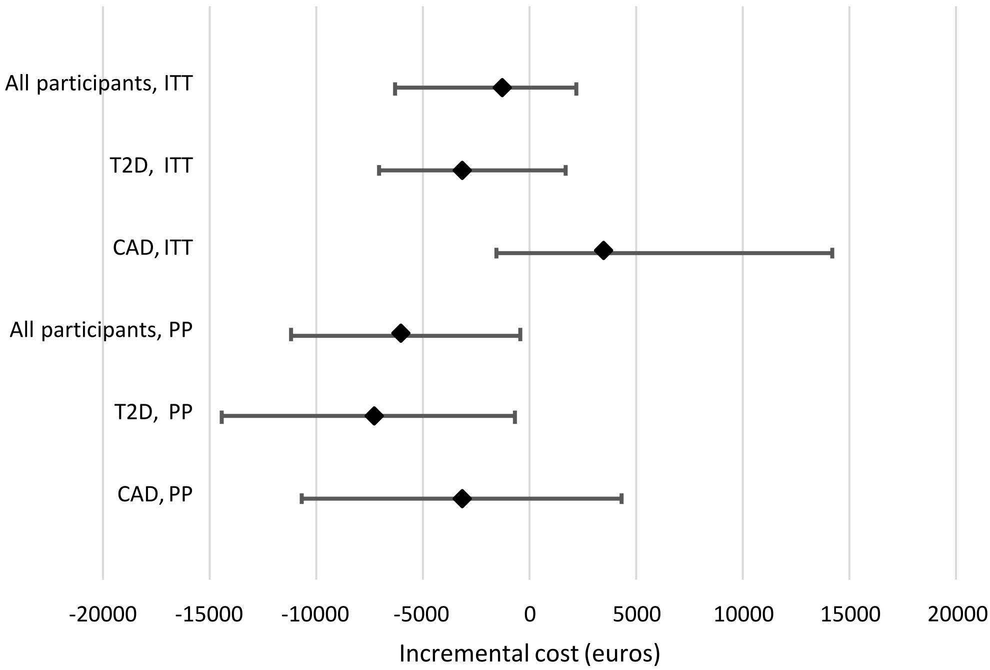


Figure 8. Mean difference in 8-year cumulative cost and bootstrapped confidence intervals for intention-to-treat (ITT) and per-protocol (PP) analyses among all participants and type 2 diabetes (T2D) and coronary artery disease (CAD) sub-groups (III).

Overall, lower costs were accrued for primary care visits, secondary care inpatient care and nursing homes in the ITT and PP analyses. However, the costs were higher in the intervention group for secondary care outpatients in both analyses. The ITT and PP analyses showed mixed results for cost differences in primary care wards, home care and service homes.

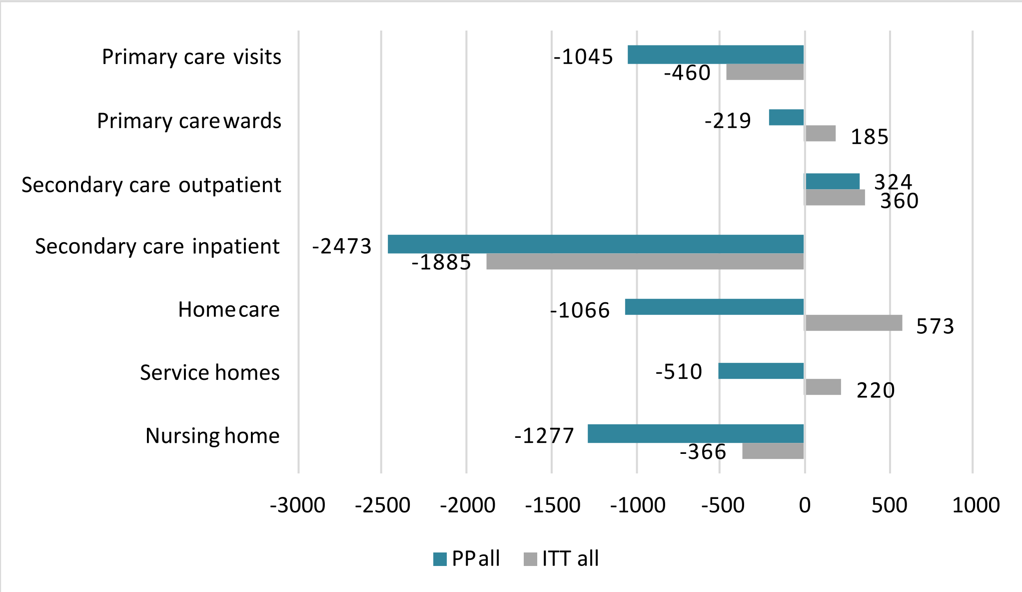


Figure 9. The social and health care costs (€) divided by service types per patient in intention to treat (ITT) and per protocol (PP) analyses in the 8-year follow-up including all patient groups in telephone health coaching study (III).

In the sub-group ITT and PP analyses, patients with T2D had lower costs in the intervention group, except for secondary care outpatients and mixed findings regarding home care (Figure 10).

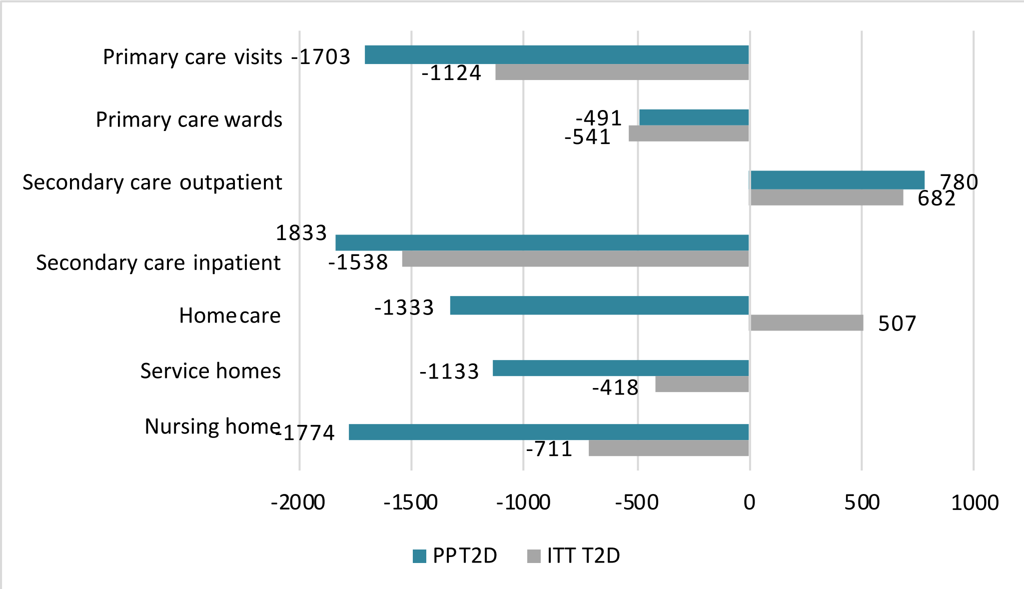


Figure 10. The social and health care costs (€) divided by service types per patient in intention to treat (ITT) and per protocol (PP) analyses in 8-year follow-up among patients with type 2 diabetes (T2D) in telephone health coaching study (III).

Among the CAD patients, the lower costs were mostly for secondary inpatient care in both the ITT and PP analyses. However, the ITT analysis for CAD clearly showed higher costs for social care and primary care wards in the intervention group compared to the PP analysis (Figure 11).

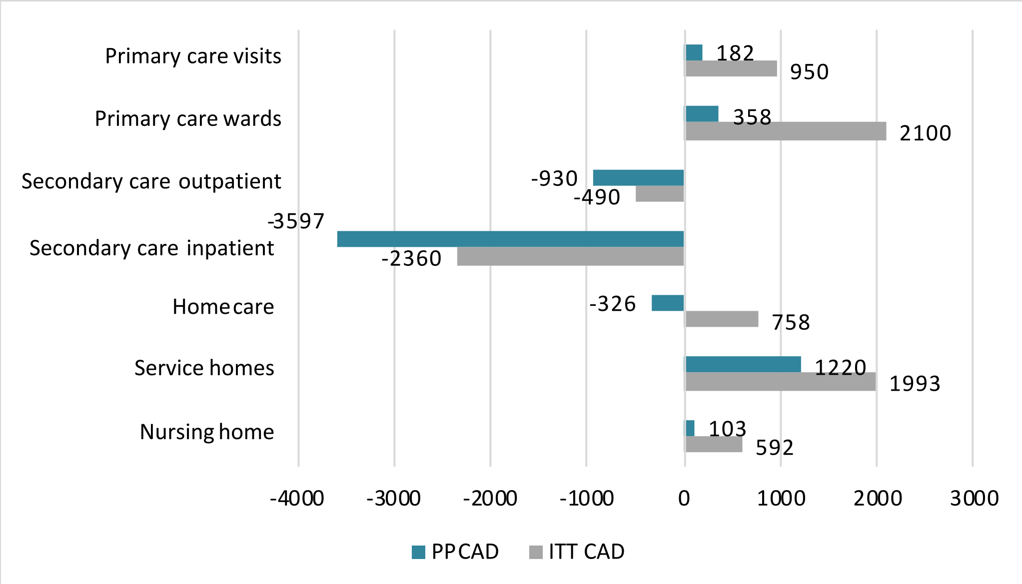


Figure 11. The social and health care costs (€) divided by service types per patient in intention to treat (ITT) and per protocol (PP) analyses in 8-year follow-up among patients with coronary artery disease (CAD) in telephone health coaching study (III)

6.2.2 Morbidity and mortality

In Sub-study IV, the ITT analysis included 1033 patients in the intervention group and 500 patients in the control group. The PP analysis included 853 patients in the intervention group and 453 in the control group.

All the tested event rates were lower in all outcomes in the intervention group, but differences were not statistically significant in the ITT analysis (Figure 12). The composite primary outcome event rate per 100 person years was 3.45 in the intervention group and 3.88 in the control group, and the HR in the intervention group was 0.87 (95% CI, 0.71–1.07; p=0.19). The ITT sub-group (T2D, CAD) analysis revealed no statistically significant effects.

Outcomes	Patients with event no.	Intervention (n=1033) no. of events (rate/100 person years)	Control (n=500) no. of events (rate/100 person years)	HR	p-value
Primary outcome					
Death from CVD causes or non-fatal stroke or non-fatal AMI or UAP	394	254 (3,45)	140 (3,88)	0,87 (0,71-1,07)	0,19
Secondary outcomes					
Death from CVD causes or non-fatal stroke or non-fatal AMI	374	241 (3,27)	133 (3,68)	0,87 (0,70-1,07)	0,20
Death from any cause or non-fatal stroke or non-fatal AMI	542	355 (4,82)	187 (5,18)	0,90 (0,76-1,08)	0,28
Death from any cause or non-fatal stroke or non-fatal AMI or UAP or CABG or PTCA or HF or PVD	684	452 (6,14)	232 (6,43)	0,93 (0,79-1,08)	0,34
Other outcomes					
Death, all causes	410	269 (3,65)	141 (3,91)	0,92 (0,75-1,12)	0,40
AMI, fatal or non-fatal	158	96 (1,30)	62 (1,72)	0,74 (0,54-1,01)	0,06
Stroke, fatal or non-fatal	102	61 (0,83)	41 (1,14)	0,71 (0,48-1,05)	0,09
Renal insufficiency	76	44 (0,60)	32 (0,89)	0,65 (0,41-1,04)	0,07
Peripheral vasc disease	93	61 (0,83)	32 (0,89)	0,92 (0,61-1,41)	0,70
Hospitalization due to CHF	258	168 (2,28)	90 (2,49)	0,89 (0,69-1,15)	0,30

Hazard ratio

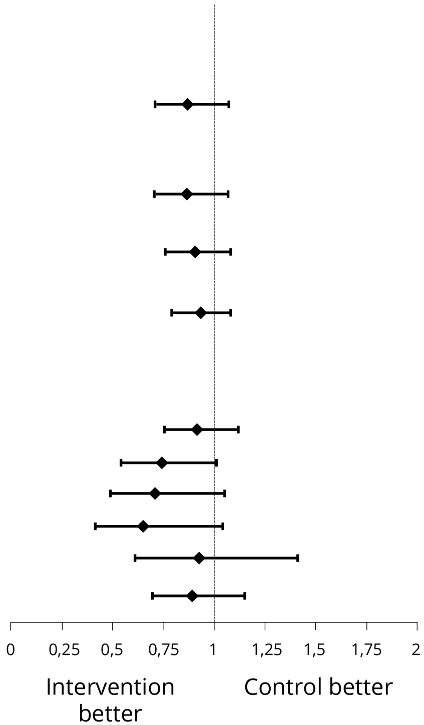
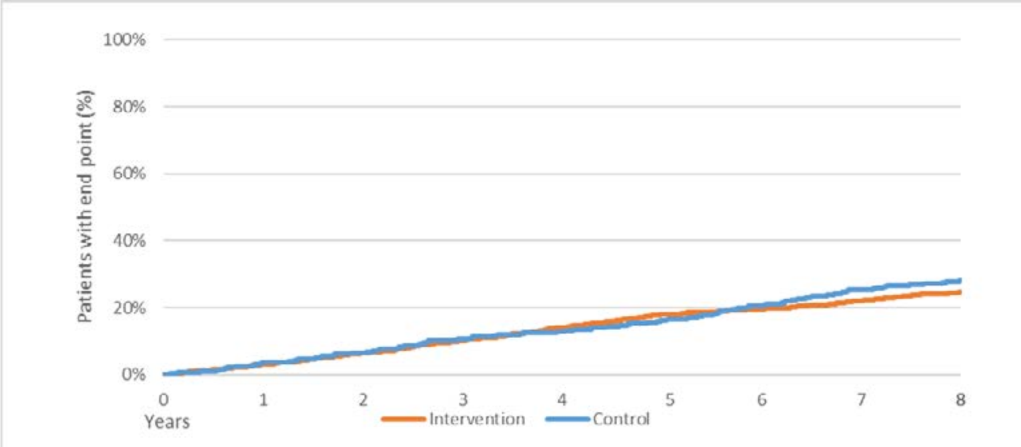


Figure 12. Cox proportional hazard ratio (HR), patient with events and p-values: comparing the primary, secondary and other outcomes between the intervention and control groups in intention to treat (ITT) analysis (IV). (Abbreviations: CVD=cardio vascular disease, AMI= myocardial infarction, UAP=unstable angina pectoris, CABG= coronary artery bypass grafting, PTCA= Percutaneous transluminal coronary angioplasty, PVD= Peripheral vascular disease, CHF=cardiac heart failure)



No at risk/year	1	2	3	4	5	6	7	8
Intervention	1004	970	941	897	854	821	793	764
Control	486	477	455	444	431	409	382	359

Figure 13. Kaplan–Meier estimates of the cumulative proportion of patients with a primary endpoint event in intention to treat (ITT) analysis in telephone health coaching study. The number of patients at risk in the control and intervention groups is presented in the table below the figure (IV).

The PP analysis showed statistically significant benefits for those who received the intervention as per protocol; renal insufficiency occurred more often in the control group with 0.86 events per 100 person years compared to the intervention group with 0.49 events per 100 person years (HR in the intervention group 0.56; 95% CI, 0.34–0.94; p=0.02). In the subgroup analysis, statistically significant differences were also found among CAD patients in two outcomes—death from any cause or stroke or AMI or UAP or CABG or PTCA or CHF or PVD (HR 0.73, CI 0.54–0.99, p=0.04) and renal insufficiency (HR 0.35, CI 0.13–0.97, p=0.04). Tables showing the results of sub-group (all patients, T2D and CAD) analyses are presented in Appendix 3.

6.3 Health coaches' perceptions of effect factors of telephone-based health coaching

6.3.1 Adoption of a 'health coach identity'

The health coaches' interviews were conducted to better understand the health coaching intervention (see Oakley et al. 2006). All seven coaches had previously worked as nurses or public health nurses in occupational health care, primary health care or secondary health care for at least 5 years before the health coaching program.

The key factors in adopting coaching skills were the training program (4 weeks), 'empowerment skills' (motivational interviewing and strength-based coaching), quality assurance (e.g. listening to coaching calls) and work supervision that was conducted in a way that enabled shared learning. All coaches mentioned the training was very well planned and intensive, and they felt privileged to participate.

Quality assurance was introduced 6 months after the start of the intervention and, according to the coaches, helped them become aware of the need to improve their empowerment skills. This was accomplished using motivational interviewing methods and led to a change process among coaches. Quality assurance and supervision were conducted by a researcher with a doctorate degree in psychology and it was seen as the most important factor for the adoption of coaching skills. Quality assurance included listening to one's own recorded coaching calls together with the supervisor (psychology), the idea of which all the coaches said was frightening in advance but which they soon realized was the only way to develop their coaching skills. Patient empowerment was the most essential difference compared to previous work as a nurse or a public health nurse.

> "….so, listening to calls and empowerment communication were some kind of shock - your working was evaluated and listened and analysed…"
> "…I think, that at the beginning of the project nobody knew how long the change from nurse to the health coach takes – the nature of work changed so radical compared previous work…"

Generally, the coaches felt that the systematic approach to evaluation, analysis and development of their work was new compared to their working history. However, the change process was slow; the coaches said it took 1–3 years and needed to be supported by continuous supervision and quality assurance.

The coaches worked in an open-plan office and were exposed to others in a way they were not used to, as they had previously worked in private office spaces alone with their patients. However, this shared space was referred to as one of the most important factors enabling shared learning. All health coaches emphasized that shared experiences and learning together and learning from each other were essential factors in the change process. Once they were familiar with each other's ways of working they were able to easily consult each other. Support from the team and the managers was also emphasized.

> "…so, we 'picked up' colleges' good working methods and expressions and thus we empowered each other's"

Additionally, training and supervising continued during the whole project, thus enabling continuing coaching development. Coaches also emphasized a multidisciplinary approach in training; medicine, nursing science and behavioural science were integrated.

6.3.2 The effect factors of health coaching

The effect factors of the health coaching were classified into three categories: 1) factors associated with the health coaching process itself, 2) communication and 3) the patient–health coach relationship. Factors associated with the health coaching process itself were regular contact, 'real health care service' and accessibility to the service. Factors associated with communication were active listening, motivational interviewing, empowerment and target-oriented and structured phone calls. Factors associated with the patient–health coach relationship were patient centeredness, individuality, comprehensiveness and confidentiality.

As the patients were highly committed to receiving coaching (nearly 90% of patients completed the coaching (I)), the coaches came to know the patients well over the 1-year period. This enabled the development of a confidential relationship, and the coaches felt the patients were more open and communicative than during face-to-face appointments. Active listening was challenging at first, but it was seen as a very important factor for the patients' own empowerment process. From the coaches' viewpoint, the concept of 'patient-centred' was understood more deeply than before. The other strengths were that phone calls were target-oriented and structured; phone calls did not 'sprawled', and structure of phone calls was tool for coaches. It was also very important that the coaching service was a part of the health care system; the coaches had access to and the right to update the patients' EHRs regarding coaching and health status. They also prepared patients for visits with health care providers and reflected with patients after the visits. Telephone-based health coaching was easy to access for patients, as they did not need to travel to the service. This was a real benefit, especially in more remote parts of the Päijät-Häme region where the distance to travel to social and health care services is significant.

> "...when you realize that you should have been quiet in this moment and give possibility to patient think and consider issues and then express his/her own words....being quiet in telephone is really, really hard and challenging..."
> "...coaching was regular, appointed and once a month was really good frequency to support selfcare..."

6.3.3 Interaction with health care providers

Health coaching was delivered as a centralized call-centre service. Health coaches visited health centres and met health care providers, including doctors, nurses and other professionals, to clarify their role in supporting patients in their self-management. The coaches felt the major obstacle to collaboration was health care providers' attitudes towards health coaches and the coaching service. They assumed the health care providers saw coaches as competitors and felt somehow threatened by this new group of professionals.

Health coaching was understood as an ‘overlapping’ service in usual care (e.g. ‘health coaches provide same service as diabetes nurses in the usual care’), and benefits for patients were not seen. However, health coaches felt this kind of service was ‘about 10 years ahead of time’. Despite visits and active contact with other health care professionals, genuine collaboration never developed, and this was felt to be the most unpleasant aspect of the health coaching process.

6.4 Summary of the results

A summary of the study results is presented in Table 5. The table also includes an interpretation of the results for decision making. The interpretation of effectiveness depends on the viewpoint of the discipline, especially in long-term follow-up. However, all measured outcomes showed non-negative benefits in the intervention group. The short-term results suggest that intervention is equal or preferred compared to usual care. In long-term intervention is indifferent or preferred. The overall results (I-IV) througout the follow-up suggest that intervention is preferred in targeted patient groups. Health coaches’ learning process and high-quality health coaching take time and need support to be realized in practice.

Table 5. Summary of the study results of telephone health coaching.

Target of intervention	Evaluation of effectiveness	Participants	Results	Interpretation	Decision making: 'goodness of intervention'
To achieve national clinical standards of secondary prevention	Goalachievement analysis	Intervention n= 816 Control n= 405 modified intentiont to treat (mITT)	Significant difference was found in one outcome (diastolic blood pressure) and target levels of systolic blood pressure and waist circumference were reached more frequently in the intervention group.	All defined clinical targets were not achieved.	Intervention is equal: value-based choice
To achieve increased quality of life with lower cost	Costeffectiveness analysis	Intervention n=685 Control n= 313 modified intention to treat (mITT)	Incremental cost-effectiveness ratio (ICER; €/QALY): All patients €48, 000, coronary artery disease (CAD) €40, 278, type 2 diabetes (T2D) €20, 000 Cost-effectiveness acceptability curve (CEAC): willingness to pay (WTP) with €50, 000 / Quality adjusted life years (QALY), the probability of cost-effectiveness is 55% for all patient and 75% for type 2 patients (T2D) patients.	Intervention increased quality of life with acceptable costs	Intervention is preferred (health economic viewpoint)
To achieve cost savings in chronic disease management	Cost analysis	Intervention n=1033 Control n= 500 intention to treat (ITT)	Total costs were 3% lower in intervention group, difference was not statistically signifigant (P=0.2)	Intervention saved costs, but the results are not generalizable in population level	Intervention is indifferent (statistical viewpoint) Intervention is preferred (health economic viewpoint)

Target of intervention	Evaluation of effectiveness	Participants	Results	Interpretation	Decision making: 'goodness of intervention'
		Intervention n=853 Control n=453 per protocol (PP)	Total costs were 14% lower in intervention group, difference was statistically signifigant (P = 0.02)	Intervention is cost saving	Intervention is preferred
To prevent mortality and secondary complications of chronic disease	Goal-achievement analysis	Intervention n=1033 Control n= 500 intention to treat (ITT)	Hazard rate was < 1 in all outcomes, difference was not statistically significant in any endpoint	Intervention prevents secondary complications, but results are not generalizable in population level	Intervention is indifferent
		Intervention n=853 Control n=453 per protocol (PP)	Hazard rate was < 1 in all outcomes, differenece was statistically significant in renal insufficient P=0.02, (all patients), mortality P=0.04 and renal insufficience P=0.04 (coronary artery disease (CAD) patients)	Intervention prevents mortality and secondary complications	Intervention is preferred
To provide highquality health coaching intervention	Descibe the effect factors from health coaches' perspective (aditional results in summary)	N=7 (health coaches)	1. Adaptation coaching skills took 1-3 years and the key factors were trainig program, work supervision, quality assurance and shared learning 2. Health coaching effect factors were process itself, communication and health coach - patient relationship 3. Interaction with health care providers was not successful	Professionals need time, support and quality assurance to adopt coaching skills that effect factors enables. Early phase 'deep' information, interaction and cooperation between usual care professional is essential.	Understanding the nature of intervention is essential for decision makers and evaluation of complex intervention should extend to three years.

7 Discussion

7.1 Main results regarding the short-term effectiveness of the health coaching intervention

In the short term, the effectiveness of health coaching was evaluated in terms of clinical outcomes (I) and cost-effectiveness (II). Clinical changes after the 1-year coaching were quite modest; only one clinical endpoint (diastolic blood pressure) showed a statistically significant difference due to the intervention. However, positive changes were also seen in systolic blood pressure and waist circumference. In general, the effectiveness of health coaching has mainly been evaluated in the short term with follow-up of 12 months or less. The results were similar with mixed findings, especially in clinical outcomes (e.g. Dennis et al. 2013; Kivelä et al. 2014; Dejonghe et al. 2017). In this study, the clinical targets might be too strict to achieve in the short term because health behaviour changes may have a delayed impact (I).

Based on the health coaches' interviews, it might have been unrealistic to expect significant impacts in the short term. First, the learning of coaching skills began 6 months after the beginning of the program, and individual adoption of coaching skills took at 1–3 years. It is important to concretize coaching concepts at the practical level, such as what 'empowerment' or 'motivational interviewing' really mean in interactions and communication between professionals and patients. Butterworth et al. (2007, 299) emphasized the importance of motivational interviewing, stating 'To date motivational interviewing-based health coaching is the only technique to have been fully described and consistently demonstrated as causally and independently associated with positive behavioural outcomes.' Thus, the importance of continuous quality assurance of health coaching is emphasized. Health coaches also highlighted this aspect that quality assurance implemented by psychology. According to the review by Kivelä et al. (2014), the most promising results in health coaching were achieved in studies in which the coaches were trained by psychologist. Second, according to the health coaches' interviews, the patients' behavioural changes must be gradually integrated into their daily lives. This takes at least 6 months, and therefore the effects of intervention tend to be delayed (see Absetz et al. 2009; Kivelä et al. 2014).

In the CEA, the intervention was more effective compared to care as usual but was also more costly (cost-effectiveness plane in the north-east). However, health coaching programs may well be acceptable, as an overall ICER of €48,000 was found. The cost-effectiveness was highest for patients with T2D (ICER €20,000) (II). Generally, the evidence of the cost-effectiveness of health coaching has been limited (Hale & Giese 2017), and changes in cost and QoL may not be found in short-term follow-up (Drummond 2001). According to Hale and Giese (2017), one of the three cost-effectiveness studies in their review affirmed the cost-effectiveness of health coaching; Jonk et al. (2015) found a significant reduction in outpatient and total expenditures in the short term (6 months), whereas Morello et al. (2016) and Wagner et al. (2016) found no cost reduction in a 1-year follow-up (Jonk et al 2015; Morello et al. 2016; Wagner et al. 2016). Panagioti et al. (2018) found similar findings for multimorbid elderly people in a 20-month follow-up; lower levels of emergency care use but increased use of planned services with higher costs and increased QALY were observed (cost-effectiveness plane in the north-east).

Health coaching particularly improved HRQoL in CAD patients (0.018 change in 15D score), but there was less improvement for patients with T2D (0.008 change in 15D score). Possible explanations for the difference between the patient groups can be found in their medical history and in the delivery of services in the health care system. For patients with T2D, the diagnosis was given an average of 10 years before the study, whereas most of the CAD patients were recruited for the health coaching program a few months after a PTCA. In usual care, the self-care support for CAD patients is typically not arranged so systematically, whereas T2D patients receive regular treatment and self-care support from specially trained diabetes nurses. Further, health coaches' emphasized factors that might increase QoL, such as regular contact (once a month), establishing a patient–health coach relationship, active listening, patient-centeredness, individuality, comprehensiveness, confidentiality, empowerment and using motivational interviewing techniques. Thom et al. (2016) identified similar factors based on health coaches' viewpoints, but patients have also emphasized the same elements. The motivation for lifestyle changes and self-management is high after an acute cardiovascular attack (see Evans 2009). Therefore, the health coaching support for CAD patients might be focused in the optimal phase of the disease process to improve HRQoL. Standard follow-up visits after acute cardiovascular attacks are arranged more frequently according to the Finnish clinical guidelines for primary and secondary health care, resulting in higher costs than for T2D patients (II.)

7.2 Main results regarding the long-term effectiveness of the health coaching intervention

In the long term, effectiveness was evaluated in terms of social and health care costs (III), morbidity and mortality (IV). The total social and health care costs were lower in the intervention group, as were morbidity and mortality (severe clinical events). The differences were not statistically significant for all the patients in the ITT analysis, whereas statistically significant differences were found in the PP analysis for patients who received health coaching. In the sub-group ITT and PP analyses, among T2D patients the costs were lower in secondary inpatient care, primary care visits and primary care wards. Among CAD patients, the lower costs were found mostly in secondary inpatient care in both the ITT and PP analyses, whereas higher costs were clearly found in the ITT analysis for social care and primary care wards compared to the PP analysis.

Long-term effectiveness was evaluated by linking the study cohort data to national registries. This allowed a long-term follow-up of all participants. Generally, the number of effectiveness studies of health coaching with long-term follow-up is small; according to the author's knowledge, the costs of social care have not been previously evaluated. To date, Byrnes et al. (2018) have had the longest follow-up time (6.35 years), finding a significant reduction in overall mortality and lower total health insurance costs in a prospective parallel-group case-control study with matched controls.

In this study, social and health care costs were higher in the intervention group during the first 2 years, but by the end of the 8-year follow-up costs were reduced in the intervention group. A possible explanation for the higher costs at the beginning is that one target of the health coaching intervention was to empower patients to manage their own disease and actively use health care services according to the guidelines, that is to ‘get recommended tests and services’ and to ‘follow up with specialists and appointments’. The intervention and study protocol did not include additional investigations or clinical visits (see Drummond 2001). Health coaches prepared patients for visits with health care providers, after which they reflected with the patients, thus forming a ‘bridge between clinician and patient’ (Bennet et al. 2010, Thom et al. 2016). Health coaching helps patients use health services more effectively and might result in additional visits and costs in the first few years. Evidence of short-term effectiveness mainly verifies that health coaching does not reduce health care utilization or result in cost savings. Wagner et al. (2016) found similar results in a 1-year follow-up; the average costs were higher in the intervention group in the first year compared with the costs for the control group. In a 2-year follow-up, Härter et al. (2016) found that hospital admissions were higher in intervention group patients with multiple diseases, whereas a significant reduction was found in hospital admissions for HF patients. Wennberg et al. (2010) found that a targeted telephone care-management program was successful in reducing the medical costs and hospitalization rate of patients in a 12-month follow-up.

Severe clinical events were more common in the control group than in the intervention group, but the findings were not statistically significant in the ITT analysis. Health coaching might decrease the risk of cardiovascular events for those patients who are able and willing to follow through with the intervention. According to the author’s knowledge, this is the first RCT study evaluating the effectiveness of health coaching on morbidity and mortality among T2D and CAD patients in an 8-year follow-up. Long-term evaluation has been recommended in several studies (e.g. Kivelä et al. 2014; Karhula et al. 2015; Härter et al. 2016; Hale and Giese 2017). Byrnes et al. (2018) found a significant reduction in overall mortality among CVD patients after a 6-month coaching program using a matched-control RCT with a 6.3-year follow-up.

The effectiveness of intensive, target-driven multifactorial interventions in the long term has been previously reported. For example, it has been shown that complications of T2D can be prevented or delayed by good glycaemic control and the management of behavioural lifestyle risk factors, such as obesity, smoking and an unhealthy diet (Orchard et al. 1990; UKPDS group 1998; Stratton et al. 2000; Gaede et al. 2008). According to a review by Angermayr and co-workers (2010), the longest follow-ups have not extended beyond 5 years, and often clinical improvements (e.g. BMI, blood glucose, blood lipids or blood pressure) were found but were not sustained during the post-intervention follow-up and had no effect on mortality. The AHEAD group (2013) and Ueki et al. (2017) found similar results, that is, significant improvements in risk factors during the intensive intervention phase but no effects on mortality or morbidity in the long term. These target-driven interventions often emphasize compliance rather than support patient autonomy and competence and may not be able to produce sustainable changes (Ueki et al. 2017), whereas the target of coaching interventions is to support self-regulation skills, such as self-monitoring, goal setting and action planning, in order to make small but sustainable changes.

7.3 Decision making perspective

Basically, the purpose of evaluating effectiveness is to provide information for decision making at the political/civil servant level (see Rajavaara 2006, 9; Drummond 2008). Rational decision making emphasizes target setting because targets are instruments for measuring and evaluating choices that have been made. In this study, the definition of problem was clear in a general level; increased chronic disease with increased complications and costs. However, a more detailed definition of the problem would have revealed the extent and significance of the problem, for example, showing the prevalence of chronic disease and the total costs of chronic care divided by the sectors. However, providing real-time, reliable and comprehensive information is challenging in real life, especially at the local level, due to defective documentation and fragmented information technology systems.

Health coaching intervention was tested as a means to affect a defined problem of chronic disease. A health coaching program was established to support patients who were already 'inside' the health care system. According to health coaches' interviews, interaction between health coaches and health care professionals did not work, mainly due to the attitudes of the latter group. The health coaches were felt to be a threat, and the benefits for patients were not recognized. Therefore, the context in which the intervention is carried out is important, as is the process evaluation, because the findings can explain why intervention works or does not work or has unexpected consequences (see Oakley et al. 2006; Craig et al. 2008). According to Grol and Wensing (2004) and Elissen et al. (2013), the other barriers to implementing self-management support are insufficient adoption of the empowerment paradigm, lack of awareness and knowledge, poor attitudes and lack of motivation to change. This study presents similar findings. Therefore, a better understanding and a stronger commitment to an intervention among decision makers and health care professionals in the early phase might lead to better interaction between professionals and thus to better chronic care in the population.

In health economics, the cost-effectiveness plane is used to demonstrate the ratio of QoL and costs. From the decision-making perspective, decision is clear if intervention is the more costly and decrease QoL or intervention is cheaper and increase QoL. The other options require a value judgement from decision makers. In this study, the cost-effectiveness plane showed that intervention was more costly and increased QoL in the short term. Cost reduction was not found until after 2 years of follow-up, and we were not able to investigate QoL in the long term. In real life, it is rarely possible to extend the follow-up period; consequently, the costs, time and availability of necessary information might restrict rational decision making.

In this TERVA project, the decision to run down health coaching service was based on short-term clinical outcomes. Results of the effectiveness of health coaching were expected too early by politicians and civil servants, and the nature of the complex intervention was ignored. The other results (Sub-studies II, III, IV) were not available at the time of decision making. Therefore, understanding the nature of the investment is essential in decision making, and it is recommended that follow-up time be expanded to at least 3 years.

Rational decision making emphasizes quantitative, numeric-based information. Knowledge can also be qualitative, but it is difficult to describe and give exact values. In this study, additional qualitative research, that is, health coaches' interviews after the study, provided a more detailed understanding of the effects of the health coaching intervention. It is recommended that qualitative evaluation also be included in short-term evaluation (see Oakley et al. 2006; Paronen 2015, 12, 147).

The evidence also suggested that possibilities to comply to principals of rational decision making are troubled and exaggerated (Harisalo 2008, 150–151; Paronen 2015, 12). In this study, from the decision-making viewpoint, the behavioural-based complex health coaching intervention was evaluated traditionally only in terms of short-term clinical outcomes (numeric). Effectiveness was narrowly evaluated from the perspective of one discipline. Therefore, the decision to eliminate the health coaching service was based on that early stage and on incomplete and narrow information. It is recommended (e.g. Harisalo 2008, 153) that in the early stage target setting and evaluating should be planned from different perspectives using multidisciplinary, scientific knowledge to acquire a more complete picture for decision making. However, time (especially the cycle of political decision-making), uncertainty, conflicts and ambiguity often restrict rationality in real life.

7.4 Validity of the study

In this study, an RCT design was used in a real-life setting, thus providing the strongest evidence of the potential effects of an intervention. Conventionally, RCTs are conducted in controlled circumstances where uncontrolled factors and 'non-resource inputs', such as staff attitudes, the social environment, patient history and personal resilience, do not exist. In this study, the intervention was complex; that is, it included patients' and professionals' behavioural change processes. Further, the health coaching program was seen by some as a threat and as competition to usual care, and we do not know how it affected the patients' usual care (whether patients were supported and motivated more than usual). By using a qualitative method alongside the RCT, these explanatory factors were revealed. The literature also recognizes that non-experimental methods are needed in evaluating complex interventions (e.g. Koskinen-Ollonqvist et al. 2005, 5; Virtanen 2007, 115; Craig et al. 2008; Mackenzie et al. 2010).

In this study, the sample size was originally calculated to evaluate the short-term effectiveness of health coaching, and it may have been too small to observe the long-term effects and statistically significant differences. For example, the combined CAD group included only 264 patients in the intervention group and 142 patients in the control group in the long-term evaluation. It is quite common in clinical trials that the sample size calculation is performed for the primary clinical endpoint, and, for example, measurement of QoL is often a second endpoint (Drummond 2001). Patients were selected based on clinical inclusion criteria, but the exclusion criteria were not controlled strictly enough. Therefore, there were patients in both group who did had have the capability to participate in the health coaching intervention; they gave consent but did not participate in any other activities related to the study, such as returning questionnaires or participating in the clinical measurements.

Data collection was accomplished using multiple methods—clinical measurements, laboratory tests (EHR), questionnaires, registries and interviews. In Study I, 80% of clinical measurements were done by a nurse (waist circumference, BMI and blood pressure) both at the beginning and at the end of follow-up. However, laboratory measures of lipids (cholesterol) were available in EHRs only for a fifth of the patients, and HbA1c measures were available only for 54%. The explanation for the low availability of laboratory tests results is that additional laboratory tests were not done for the study; the results were collected using routine test frequency according to Finnish Current Care Guidelines. In Sub-study II, HRQoL was measured using the validated 15D instrument (see section 4.6). Evaluating the effectiveness of self-care interventions using 15D, Leal et al. (2017) found similar changes in QALYs in the short term for T2D patients in a cluster RCT. The response rate in Sub-study II was 65%, which is considered good for questionnaire surveys.

Long-term follow-up Sub-studies III and IV were based on linkages to Finnish national registries. Only the information for two patients was not found in the registries; thus, the data covered almost 100%. This allowed the collection of detailed patient-level data and an evaluation of the effectiveness of health coaching, including all social and health care services and all severe distal endpoints. The intervention itself did not include additional visits to social and health care services (see Drummond 2001). Thus, the Finnish national registries enabled conducting ITT analysis in the long-term evaluation. CONSORT guidelines recommend that both ITT and PP analyses should be performed for all planned outcomes because they allow readers to interpret the effect of intervention (e.g. Schultz et al. 2010; Brody 2016; Ranganathan et al. 2016). In this study, to retain comparability between study groups, those patients who did not perform any activities related to the study were excluded from the PP analysis. Most of the excluded patients were so-called 'heavy users'; they died earlier in the control group than in the intervention group and thus did not cause any costs to the control group in the ITT analysis. According to the definition of PP analysis, it includes those patients who strictly follow study protocol. In this study, it was clear to exclude inactive patients from both study groups; therefore, it might better to use the term mITT. The short-term result was based on 'active participants', and therefore parallel reporting was used in the summary of the study. Further, ITT analysis is the most conservative, whereas the PP analysis estimates the true effectiveness of intervention but might exaggerate the results of study (Brody 2016; Mc Coy et al. 2017; Ranganathan et al. 2016).

The researcher's position in the TERVA research and development program in 2006–2009 was project manager. The researcher was very familiar with the TERVA program and had deep knowledge of project, thus enabling a realistic interpretation of the findings. However, the risk is that objectivity might diminish, especially in the qualitative part of the study (see Tuomi & Sarajärvi 2009). Health coaches were interviewed 10 years after the project end; therefore, the interview material was based mainly on memory. However, the health coaches' perceptions were consistent, and they were able to evaluate the health coaching service from the perspective of time.

The health coaching service would be more easily implemented as a part of the health care system today than it would have been 10 years ago.

7.5 Ethical considerations

The TERVA study protocol was approved by the Ethics Committee of Päijät-Häme Social and Care District (Dnro ETMK 56/2008 and Dnro 65/2008). Information and consent letters were sent to the participants, and written consent was obtained from all participant before enrolment. The purpose and procedures of the study were contained in the information letter. For studies II, III and IV, the National Institute for Health and Welfare approved the study protocol and gave consent for individual data collection from national registries (permission numbers: Dnro THL/66/5/5.05.00/2009 and Dnro THL/119/5.05/2015). The cause of death data was obtained from Statistics Finland (Dnro TK-53-1033-17). Written consent for health coaches' interviews was obtained from all health coaches. The purpose of the interviews and a description of the anonymous processing of interview material was stated in the information and consent letters.

Having the role of project manager, the researcher was also responsible for building the health coaching centre, recruiting the health coaches and coordinating the work streams of the program together with the project partners. The researcher was also responsible for ethical documentation, such as preparing the information and consent letters for ethical consideration, participating in patient recruitment, organizing patients' clinical measurements, participating in data collection and coordinating scientific steering group actions (Sub-study I). In Sub-studies II, III and IV, the researcher applied for permission from the National Institute for Health and Welfare and Statistics Finland and was responsible for data collection.

8 Conclusions and proposals for further research

8.1 Conclusions

The aim of this study was to evaluate the effectiveness of 1-year telephone health coaching on health benefits and social and health care costs among chronic disease patients in the short and long terms. In the short term, clinical outcomes remained modest, but the intervention increased QoL with acceptable costs. The results of the study were consistent with those of previous studies. This study provides new information on the effectiveness of health coaching in the long term. Specifically, national registries enabled reporting total social and health care costs at the individual level. The results showed positive effects of the intervention in all outcomes, but the differences were not statistically significant in the ITT analysis, probably due to the small sample size. Therefore, the results are not generalizable at the population level. However, from a health economics perspective, intervention is preferred. For patients who received health coaching, the intervention has remarkable potential for cost savings and for the prevention of secondary complications of chronic disease in the long term. Therefore, patients' capability of participating in intervention must be ensured. The effects might be higher in countries where self-care support for chronic care has not been arranged systematically, for example specialized diabetes nurse system is lacking. Based on the health coaches' interviews, the adoption of coaching skills took 1–3 years, and supporting coaching skills and quality assurance were prerequisites for adoption. Patients' behavioural changes also take time to be integrated into their daily lives.

This study confirmed that evaluating effectiveness is multidimensional and that effectiveness depends on the viewpoint of the discipline. From the viewpoint of rational decision making, understanding the nature of intervention is essential for evaluation and for decision makers to set realistic targets and evaluate them in a timely fashion in order to reveal the potential benefits. The results of this study suggest that evaluating the effectiveness of health coaching interventions should extend to at least 3 years using a multidisciplinary, multidimensional approach and should include patients' perspectives. Telephone health coaching might be preventive and cost-saving, especially in the long term and focusing on those patients who are willing and able to participate. Thus, it might become part of self-care support in the chronic care delivery system.

8.2 Proposals for further research

This study revealed many aspects for further research in order to better understand the mechanism and factors of the effectiveness of health coaching. Multidimensional, multidisciplinary and multimethod research is needed to identify effectiveness factors that may have a remarkable impact on the overall effectiveness of health coaching. Clarifying patients', professionals' and decision makers' experiences and attitudes would be useful to better understand and develop chronic disease management at the practice level. Further, decision makers often need information in the short term; therefore, it would be important to study effectiveness factors and methods of complex intervention that predict effectiveness in the long term. National registries in Finland enable long-term follow-up studies, and a further effectiveness study is planned by the researcher and the research group.

References

Abraha Iosief & Montedori Alessandro 2010. Modified intention to treat reporting in randomized controlled trials: systematic review. British Medical Journal 2010;340:c2697.

Absetz Pilvikki, Oldenburg Brian, Hankonen Nelli, Valve Raisa, Heinonen Heikki, Nissinen Aulikki & Uutela Antti 2009. Type 2 diabetes prevention in the real world: three-year results of the GOAL lifestyle implementation trial. Diabetes Care 32(8), 1418–1420. https://doi.org/10.2337/dc09-0039

Absetz Pilvikki, Valve Raisa, Oldenburg Brian Heinonen Heikki, Nissinen Aulikki, Fogelholm Mikael, Ilvesmäki Vesa, Talja Martti & Uutela Antti 2007. Type 2 diabetes prevention in the "real world": one-year results of the GOAL Implementation Trial. Diabetes Care 30(10), 2465–70. https://doi.org/10.2337/dc07-0171

Adair John 2016. Decision making and problem solving. London: Kogan Page.

Anderson Robert, Fitzgerald James, Wisdom Kimberlydawn, Davis Wayne & Hiss Roland 1997. A comparison of global versus disease-specific quality-of-life measures in patients with NIDDM. Diabetes Care 20(3), 299–305. https://doi.org/10.2337/diacare.20.3.299

Angermayr Lucia, Melchart Dieter & Linde Klaus 2010. Multifactorial lifestyle interventions in the primary and secondary prevention of cardiovascular disease and type 2 diabetes mellitus—a systematic review of randomized controlled trials. Annals of Behavioral Medicine 40(1), 49–64. https://doi.org/10.1007/s12160-010-9206-4

Axelsson Lilian & Engström Agneta 2001. The concept of effectiveness—a blind alley? A study of different interpretations in a Swedish county council. International Journal of Health Planning and Management 16 (1), 61–71. https://doi.org/10.1002/hpm.610

Barber Julie & Thompson Simon 2000. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. Statistics in Medicine 15;19(23), 3219–36. https://doi.org/10.1002/1097-0258(20001215)19:23<3219::AID-SIM623>3.0.CO;2-P

Barnett Karen, Mercer Stewart, Norbury Michael, Watt Graham, Wyke Sally & Guthrie Bruce 2012. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet 380 (9836), 37–43. https://doi.org/10.1016/S0140-6736(12)60240-2

Bell Simon, Airaksinen Marja, Lyles Alan, Chen Timothy & Aslani Parisa 2007. Concordance is not synonymous with compliance or adherence. British Journal of Clinical Pharmacology 64(5), 710–711. https://dx.doi.org/10.1111%2Fj.1365-2125.2007.02971_1.x

Bennet Heather, Coleman Eric, Parry Carla, Bodenheimer Thomas & Chen Ellen 2010. Health coaching for patient: does your practice "give patients a fish" or "teach patient to fish"? Family Practice Management 17(5), 24–29.

Benzo Roberto, Kirsch Janae, Hathaway Julie, McEvoy Charlene & Vickers Kristin 2017. Health coaching in severe COPD after a hospitalization: a qualitative analysis of a large randomized study. Respiratory Care 62(11), 1403–1411.

Berk Richard & Rossi Peter 1999. Thinking about program evaluation. Thousand Oaks, CA: SAGE.

Billot Laurent, Corcoran Kate, McDonald Alina, Powell-Davies Gavaine & Feyer Anne-Marie 2016. Impact evaluation of a system-wide chronic disease management program on health service utilisation: a propensity-matched cohort study. PLoS Medicine, 13(6), e1002035. https://dx.doi.org/10.1371%2Fjournal.pmed.1002035

Binmore Ken 2009. Rational decisions. Princeton, NJ: Princeton University Press.

Bissonnette Janice 2008. Adherence: a concept analysis. Journal of Advanced Nursing 63(6), 634–643. https://doi.org/10.1111/j.1365-2648.2008.04745.x

Bonell Chris, Fletcher Adam, Morton Matthew, Lorenc Theo & Moore Laurence 2012. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. Social Science & Medicine 75(12), 2299–306. https://doi.org/10.1016/j.socscimed.2012.08.032

Boos Leo & Jacquemart Charlotte 2000. A primer on decision making: how decisions happen. Schmalenbach Business Review 52(1), 92-94.

Brody Tom 2016. Intent-to-treat analysis versus per protocol analysis. In: Clinical trials, 2nd ed. Study Design, Endpoints and Biomarkers, Drug Safety, and FDA and ICH Guidelines. Academic 173–201.

Butterworth Susan, Linden Ariel & McClay Wende 2007. Health coaching as an intervention in health management programs. Disease Management and Health Outcomes 15(5), 299–307.

Byrnes Joshua, Elliott Thomas, Vale Margarite, Jelinek Michael & Scuffham Paul 2018. Coaching patients saves lives and money. American Journal of Medicine 131(4), 415–421.

Cardio Vascular Disease. Current Care Guidelines. Working group set up by the Finnish Medical Society Duodecim and the Finnish Cardiac Society. Helsinki: The Finnish Medical Society Duodecim, 2017 (referred May 27, 2020). Available online at: www.kaypahoito.fi

Campbell Michelle, Fitzpatrick Ray, Haines Andrew, Kinmonth Ann, Sandercock Peter, Spiegelhalter David & Tyrer Peter 2000. Framework for design and evaluation of complex interventions to improve health. British Medical Journal 321(7262), 694–696.

Campbell Neil, Murray Elizabeth, Darbyshire Janet, Emery Jon, Farmer Andrew, Griffiths Frances, Guthrie Bruce, Lester Helen, Wilson Phil & Kinmonth Ann 2007. Designing and evaluating complex interventions to improve health care. British Medical Journal 334(7591), 455–459.

Chapman Anna, Browning Colette, Enticott Joanne, Yang Hui, Liu Shuo, Zhang Tuohong & Thomas Shane 2018. Effect of a health coach intervention for the management of individuals with type 2 diabetes mellitus in China: a pragmatic cluster randomized controlled trial. Frontiers in Public Health 6:252.

Clarke Geraldine, Conti Stefano, Wolters Arne & Steventon Adam 2019. Evaluating the impact of healthcare interventions using routine data. British Medical Journal 2019; 365:2239.

Craig Peter 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. British Medical Journal 2008;337:a1655.

Dejonghe Lea, Becker Jennifer, Froboese Ingo & Schaller Andrea 2017. Long-term effectiveness of health coaching in rehabilitation and prevention:

a systematic review. Patient Education and Counselling 100(9), 1643–1653.

Dennis Sarah, Harris Mark, Lloyd Jane, Powell Davies Gawaine, Faruqi Nighat & Zwar Nicholas 2013. Do people with existing chronic conditions benefit from telephone coaching? A rapid review. Australian Health Review 37(3), 381–388.

Drummond Michael 2001. Introducing economic and quality of life measurements into clinical studies. Annals of Medicine 33(5), 344–349.

Drummond Michael, Schwartz Sanford, Jönsson Bengt & Luce Bryan 2008. Key principles for the improved conduct of health technology assessments for resource allocation decisions. International Journal of Technology Assessment in Health Care 24(3), 244–258. https://doi.org/10.1017/S0266462308080343

Drummond Michael, Sculpher Mark, Claxton Carl, Stoddart Greg & Torrance George 2015. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press.

Duodecim 2019. The Finnish Current Care Guidelines. https://www.kaypahoito.fi/en/ (Accessed May 27, 2020)

Elissen Arianne, Nolte Ellen, Knai Cecile, Brunn Matthias, Chevreul Karine, Conklin Anaalijn, Durand-Zaleski Isabelle, Erler Antje, Flamm Maria, Frølich Anne, Fullerton Birgit, Jacobsen Ramune, Saz-Parkinson Zuleika, Sarria-Santamera Antonio, Sönnichsen Andreas & Vrijhoef Hubertus 2013. Is Europe putting theory into practice? A qualitative study of the level of self-management support in chronic care management approaches. Bio Med Central Health Services Research 13, 117. https://doi.org/10.1186/1472-6963-13-117

Ellrodt Gray, Cook Deborah, Lee Jean, Cho Michaela, Hunt Dereck & Weingarten Scott 1997. Evidence-based disease management. Journal of the American Medical Association 278(20), 1687–1692.

Elster Jon 1996. Rationality and the emotions. The Economic Journal, 106(438), 1386–1397.

Ettelt Stefanie, Nicholas Mays & Allen Pauline 2015. Policy experiments: investigating effectiveness or confirming direction? Evaluation 21(3), 292–307. https://doi.org/10.1177/1356389015590737

Eurostat 2014. https://ec.europa.eu/eurostat/ (Accessed May 27, 2020)
Evans Lynda 2009. Cardiac rehabilitation – past to present. Journal of Research in Nursing, 14(3), 223–240.
Expert Groups on Health System Performance Assessment, (EGHSPA) 2016. SO WHAT? Strategies across Europe to assess quality of care, European Commission, Health and Food Safety. München, Luxembourg: Publication Office of European Union.
Finnish Government 2020. https://soteuudistus.fi/en/ (Accessed October 21, 2020)
Feldman Martha & March James 1981. Information in organizations as signal and symbol. Administrative Science Quarterly 26(1981), 171–186.
Fletcher Adam, Jamal Farah, Moore Graham, Evans Rhiannon, Murphy Simon & Bonell Chris 2016. Realist complex intervention science: applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. Evaluation 22(3), 286–303.
Gabor P 1976. Management theory and rational decision making. Management Decision 14(5), 274–281.
Gaede Peter, Lund-Andersen Henrik, Parving Hans-Henrik & Pedersen Oluf 2008. Effect of a multifactorial intervention on mortality in type 2 diabetes. The New England Journal of Medicine 358(6), 580–591.
Gardner Carrie 2015. Adherence: a concept analysis. International Journal of Nursing Knowledge 26(2), 96–101.
Grol Richard & Wensing Michel 2004. What drives change? Barriers to and incentives for achieving evidence-based practice. The Medical Journal of Australia 180(6), 57-60.
Gupta Sandeep 2011. Intention-to-treat concept: a review. Perspectives in Clinical Research 2(3), 109–112.
Häkkinen Unto & Peltola Mikko 2016. Sosiaali- ja terveyspalvelujärjestelmän ohjauksen edellyttämä tietopohja. Teoksessa Keskimäki Ilmo, Moisio Antti & Pekurinen Markku (toim). Julkisen talouden ohjaus ja sosiaali- ja terveydenhuollon ja koulutuksen rakenneuudistus. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 22/2016.

Hale Rachel & Giese Jeannie 2017. Cost-effectiveness of health coaching: an integrative review. Professional Case Management 22(5), 228–238. https://doi.org/10.1097/NCM.0000000000000223

Hämäläinen Päivi, Ahola Sanna, Halme Nina, Hastrup Arja, Hietanen-Peltola Marke, Häkkinen Pirjo, Ihanus Mirja, Järvelin Jutta, Kauppinen Tapani, Kilpeläinen Katri, Kuussaari Kristiina, Liukko Eeva, Kärkkäinen Jukka, Mäkelä Matti, Mölläri Kaisa, Nurmi-Kikkalainen Päivi, Partanen Airi, Perälä Marja-Leena, Rotko Tuulia, Räikkönen Outi, Suvisaari Jaana & Töytäri Outi 2016. Koottuja indikaattoreita kehittämistyössä huomioitavaksi. Sosiaali- ja terveydenhuollon laatu- ja vaikuttavuusindikaattoreiden arviointi palvelupakettien näkökulmasta. Työpaperi. Terveyden ja hyvinvoinnin laitos.

Harisalo Risto 2008. Organisaatioteoriat. Tampere: Tampere University Press.

Härter Martin, Dirmaier Jörg, Dwinger Sarah, Kriston Levente, Herbarth Lutz, Siegmund-Schultze Elisabeth, Bermejo Isaac, Matschinger Herbert, Heider Dirk & König Hans-Helmut 2016. Effectiveness of telephone-based health coaching for patients with chronic conditions: a randomised controlled trial. PloS One 11(9), e0161269.

Härter Martin, Dwinger Sarah, Seebauer Laura, Simon Daniela, Herbarth Lutz, Siegmund-Schultze Elisabeth, Temmert Daniel, BermejoIsaac & Dirmaier Jörg 2013. Evaluation of telephone health coaching of German health insurants with chronic conditions. Health Education Journal 72(5), 622–634. https://doi.org/10.1177/0017896912453990

Hawthorne Graeme, Richardson Jeff & Day Neil 2001. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Annals of Medicine 33(35), 358–370. https://doi.org/10.3109/07853890109002090

Hayes Eileen, McCahon Cynthia, Panahi Margaret, Hamre Taryn & Pohlman Karen 2008. Alliance not compliance: coaching strategies to improve type 2 diabetes outcomes. Journal of the American Academy of Nurse Practitioners 20 (3), 155–162. https://doi.org/10.1111/j.1745-7599.2007.00297.x

Holopainen Martti & Pulkkinen Pekka 2013. Tilastolliset menetelmät. Helsinki: Sanoma Pro.

Homer Caroline 2002. Using the Zelen design in randomized controlled trials: debates and controversies. Journal of Advanced Nursing 38(2), 200–207.

Honkasalo Miiko, Linna Miika, Sane Timo, Honkasalo Atte & Elonheimo Outi 2014. A comparative study of two various models of organising diabetes follow-up in public primary health care - the model influences the use of services, their quality and costs. Bio Med Central Health Services Research 14, 26 (2014).

Hunsson Sven 2007. Rationality and emotions. Theoria 73(1), 1–2.

Hutchison Andrew & Breckon Jeff 2011. A review of telephone coaching services for people with long-term conditions. Journal of Telemedicine and Telecare 17(8), 451–458.

Jarvala Tiina, Raitanen Jani & Rissanen Pekka 2010. Kansallinen Diabetesohjelma Dehko: Diabeteksen kustannukset Suomessa 1998-2007. Suomen Diabetesliitto ja Tampereen yliopisto.

Jeanes Emma 2019. A dictionary of organizational behaviour. New York: Oxford University Press.

Jones Rowan & Pendlebury Maurice 2010. Public sector accounting. Essex, UK: Pearson Education Limited.

Jonk Yvonne, Lawson Karen, O'Connor Heidi, Riise Kirsten Eisenberg David, Dowd Bryan & Kreitzer Mary 2015. How effective is health coaching in reducing health services expenditures? Medical Care 53(2), 133–140. http://doi.org/10.1097/MLR.0000000000000287

Kahneman Daniel 2012. Ajattelu, nopeasti ja hitaasti. Helsinki: Terra cognita.

Kapiainen Satu, Väisänen Antti & Haula Taru 2011. Terveyden- ja sosiaalihuollon yksikkökustannukset Suomessa vuonna. Raportti. Terveyden ja hyvinvoinnin laitos.

Karhula Tuula, Vuorinen Anna-Leena, Rääpysjärvi Katja, Pakanen Mira, Itkonen Pentti, Tepponen Merja, Junno Ulla-Maija, Jokinen Tapio, Van Gils Mark, Lähteenmäki Jaakko, Kohtamäki Kari & Saranummi Niilo 2015. Telemonitoring and mobile phone-based health coaching among Finnish diabetic and heart disease patients: randomized controlled trial. Journal of Medical Internet Research, 17(6), e153.

Karjalainen Leila 2010. Tilastotieteen perusteet. Ristiina: Pii-kirjat.

Kettunen Pekka 2017. Vaikuttavuuden arviointi sosiaali- ja terveydenhuollon palveluissa. Tutkimusraportteja 2/2017. Turun kaupunki.

Kind Paul 2001. Measuring quality of life in evaluating clinical interventions: an overview. Annals of Medicine 33(5), 323–327.

Kivelä Kirsi, Elo Satu, Kyngäs Helvi & Kääriäinen Maria 2014. The effects of health coaching on adult patients with chronic diseases: a systematic review. Patient Education and Counseling, 97(2), 147–157.

Klemola Katja 2015. Tuottavuuden, vaikuttavuuden ja kustannusvaikuttavuuden arviointi alueellisesti integroiduissa sosiaali- ja terveyspalveluissa – palvelujen käyttöön perustuva malli. Väitöskirja. Lappeenranta: University of Technology.

Komulainen Vuorela & Malmivaara 2014. Satunnaistetun kontrolloidun tutkimuksen periaatteita ja sudenkuoppia. Duodecim Lääketieteellinen Aikakausikirja 130(20), 1439–44.

Kontodimopoulos Nick, Pappa Evelina, Chadjiapostolou, Zinovia, Arvanitaki Eleni, Papadopoulos Angelos & Niakas Dimitris 2012. Comparing the sensitivity of EQ-5D, SF-6D and 15D utilities to the specific effect of diabetic complications. The European Journal of Health Economics 13(1), 111–120.

Konu Anne, Rissanen Pekka, Ihantola Mervi & Sund Reijo 2009. "Effectiveness" in Finnish healthcare studies. Scandinavian Journal of Public Health 37(1), 64–74.

Larsen Marie, Astrid, Krogstad Anne-Lene & Aas Eline 2016. Cost-utility analysis of supported self-management with motivational interviewing for patients with psoriasis. Acta Dermato-Venereologica, 96(5), 664–668. https://doi.org/10.2340/00015555-2331

Leal Jose, Ahrabian Djavid, Davies Melanie, Gray Laura, Khunti Kamlesh, Yates Thomas & Gray Muir 2017. Cost-effectiveness of a pragmatic structured education intervention for the prevention of type 2 diabetes: economic evaluation of data from the Let's Prevent Diabetes cluster-randomised controlled trial. British Medical Journal Open 7(1), e013592-013592.

Lewin Simon, Glenton Claire & Oxman Andrew 2009. Use of qualitative methods alongside randomised controlled trials of complex healthcare

interventions: methodological study. British Medical Journal 339(7723), 732–734.
Lin Wen-Chieh, Chien Hung-Lun, Willis Georgianna, O'Connell Elizabeth, Staunton Rennie Kate, Bottella Heather & Ferris Timothy 2012. The effect of a telephone-based health coaching disease management program on Medicaid members with chronic conditions. Medical Care 50(1), 91–98.
Lindner Helen, Menzies David, Kelly Jill, Taylor Sonya & Shearer Marianne 2003. Coaching for behaviour change in chronic disease: a review of the literature and the implications for coaching as a self-management intervention. Australian Journal of Primary Health, 9(3), 177–185.
Lumijärvi Ismo & Jylhäsaari Jussi 1999. Laatujohtaminen ja julkinen sektori: laadun ja tuloksen tasapaino johtamishaasteena. Helsinki: Gaudeamus 1999.
Mackenzie Mhairi, O'Donnell Catherine, Halliday Emma, Sridharan Sanjeev & Platt Stephen 2010. Do health improvement programmes fit with MRC guidance on evaluating complex interventions? British Medical Journal (Clinical research ed.)2010 Feb 1;340:c185. doi: 10.1136/bmj.c185. PMID: 20123834.
Mandelblatt Jeanne, Fryback Dennis, Weinstein Milton, Russell Louise, Gold Marthe & Members of The Panel on Cost–Effectiveness in Health, Marthe & Medicine MR 1997. Assessing the effectiveness of health interventions for cost–ffectiveness analysis. Journal of General Internal Medicine 12(9), 551–558.
Mattke Soeren, Seid Michael & Ma Sai 2007. Evidence for the effect of disease management: is $1 billion a year a good investment? The American Journal of Managed Care 13(12), 670.
McCoy Eric 2017. Understanding the intention-to-treat principle in randomized controlled trials. The Western Journal of Emergency Medicine 18(6), 1075–1078.
McGuire Alistair 2001. Theoretical concepts in the economic evaluation of health care. In Economic evaluation in health care: merging theory with practice. Edited by Michael Drummond and Alistair McGuire. Oxford: Oxford University Press.

Meklin Pentti 2001. Tavoitteiden saavuttamisen arviointi kuntataloudessa. Julkaisussa: Myllymäki A. & Vakkuri J. (toim) Tulos, normi, tilivelvollisuus: näkökulmia tilintarkastukseen ja arviointiin. Tampere. Tampere University Press.

Michie Susan, Abraham Charles, Whittington Craig, McAteer John & Gupta Sunjai 2009. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychology 28(6), 690–701.

Mikkeli Heikki & Pakkasvirta Jussi 2007. Tieteiden välissä? Johdatus monitieteisyyteen, tieteidenvälisyyteen ja poikkitieteisyyteen. Oppimateriaalit. Helsinki: WSOY.

Ministry of Social Affairs and Health 2019. https://stm.fi/sotepalvelut/jarjestelma-vastuut (Accessed May 27, 2020)

Moore Graham, Audrey Suzanne, Barker Mary, Bond Lyndal, Bonell Chris, Hardeman Wendy, Moore Laurence, O'Cathain Alicia, Tinati Tannaze, Wight Daniel & Baird Janis 2015. Process evaluation of complex interventions: Medical Research Council guidance. British Medical Journal 2015;350:h1258.

Morello Renata, Barker Anna, Watts Jennifer, Bohensky Megan, Forbes Andrew & Stoelwinder Johannes 2016. A telephone support program to reduce costs and hospital admissions for patients at risk of readmissions: lessons from an evaluation of a complex health intervention. Population Health Management 19(3), 187-195. https://doi.org/10.1089/pop.2015.0042

Mustonen Erja, Hörhammer Iiris, Absetz Pilvikki, Patja Kristiina, Lammintakanen Johanna, Talja Martti, Kuronen Risto & Linna Miika. 2020. Eight-year post-trial follow-up of health care and long-term care costs of tele-based health coaching. Health Services Research 55(2), 211–217. https://doi.org/10.1111/1475-6773.13251

Mustonen Erja, Hörhammer Iiris, Absetz Pilvikki, Patja Kristiina, Lammintakanen Johanna, Talja Martti, Kuronen Risto & Linna Miika 2020. 8-year post-trial follow-up of telephone health coaching among patients with type 2 diabetes and coronary artery disease: effect on mortality and morbidity. Submitted.

Möttönen Sakari 1997. Tulosjohtaminen ja valta poliittisten päätöksentekijöiden ja viranhaltijoiden välisessä suhteessa: kunnallisen tulosjohtamisen poliittisten päätöksentekijöiden ja viranhaltijoiden välistä tehtäväjakoa koskevat tavoitteet, niiden merkitys osapuolten väliseen valtasuhteeseen sekä tavoitteiden toteutuminen ja toteuttamismahdollisuudet valtasuhteen näkökulmasta. Suomen kuntaliitto.

Niiranen Vuokko & Lammintakanen Johanna 2011. Hallintotieteen rajapinnoilla – sosiaalihallintotieteen ja terveyshallintotieteen tutkimuksen alan ja kohteen näkökulmia. Teoksessa Virtanen Turo, Ahonen Pertti, Syväjärvi Antti, Vartiainen Pirkko, Vartola Juha ja Vuori Jari (toim). Suomalainen hallinnon tutkimus: mistä, mitä, minne? Tampere. Tampere University Press.

Niiranen Vuokko, Stenvall Jari & Lumijärvi Ismo 2005. Kuntapalvelujen tuloksellisuuden arviointi: tasapainotettu mittaristo kunnallisissa organisaatioissa. Jyväskylä: PS-Kustannus.

Nutt Paul 2003. Why decisions fail: avoiding the blunders and traps that lead to debacles. San Francisco, CA: Berrett-Koehler Publishers

Oakley Ann, Strange Vicki, Bonell Chris, Allen Elizabeth & Stephenson Judith 2006. Process evaluation in randomised controlled trials of complex interventions. British Medical Journal 332(7538), 413–416.

OECD 1993. Survey of Public Management Developments OECD 1990. General distribution.

OECD 2013. Health at a Glance. OECD Indicators.

Oksman Erja, Linna Miika, Hörhammer Iiris, Lammintakanen Johanna & Talja Martti 2017. Cost-effectiveness analysis for a tele-based health coaching program for chronic disease in primary care. Bio Med Central Health Services Research 17(1), 138. https://doi.org/10.1186/s12913-017-2088-4

Oliveira Arnaldo 2007. Decision-making theories and models: a discussion of rational and psychological decision-making theories and models: the search for a cultural-ethical decision-making model. Electronic Journal of Business Ethics and Organisation Studies, 12(2) 12–17.

Olsen Jeanette 2014. Health coaching: a concept analysis. Nursing Forum, 49(1), 18–29.

Øvretveit John 1998. Evaluating health interventions: an introduction to evaluation of health treatments, services, policies and organizational interventions. Buckingham: Open University Press.

Palmer Stephen, Tubbs Irene & Whybrow Alison 2013. Health coaching to facilitate the promotion of healthy behaviour and achievement of health-related goals. International Journal of Health Promotion and Education, 41(3), 91–93. https://doi.org/10.1080/14635240.2003.10806231

Panagioti Maria, Reeves David, Meacock Rachel, Parkinson Beth, Lovell Karina, Hann Mark, Howells Kelly, Blakemore Amy, Riste Lisa, Coventry Peter, Blakeman Thomas, Sidaway Mark & Bower Peter 2018. Is telephone health coaching a useful population health strategy for supporting older people with multimorbidity? An evaluation of reach, effectiveness and cost-effectiveness using a 'trial within a cohort'. Bio Med Center Medicine 16(1):80. https://doi.org/10.1186/s12916-018-1051-5

Panagioti Maria, Richardson Gerry, Small Nicola, Murray Elizabeth, Rogers Anne, Kennedy Anne, Newman Stanton & Bower Peter 2014. Self-management support interventions to reduce health care utilisation without compromising outcomes: a systematic review and meta-analysis. Bio Med Central Health Services Research 14:356. https://doi.org/10.1186/1472-6963-14-356

Parkerson George, Connis Richard, Broadhead Eugene, Patrick Donald, Taylor Thomas & Tse Chiu-Kit 1993. Disease-specific versus generic measurement of health-related quality of life in insulin-dependent diabetic patients. Medical Care 31(7), 629–639.

Paronen Elsa 2015. Rationaalinen päätöksenteko sosiaali- ja terveyspalvelujen uudistuksissa: tapaustutkimus kahdesta sosiaali- ja terveyspalvelujen uudistuksesta Pohjois-Savossa. Yhteiskuntatieteiden ja kauppatieteiden tiedekunta / Sosiaali- ja terveysjohtamisen laitos. Väitöskirja. Itä-Suomen Yliopisto.

Patja Kristiina, Absetz Pilvikki, Auvinen Anssi, Tokola Kari, Kytö Janne, Oksman Erja, Kuronen Risto, Ovaska Timo, Harno Kari, Nenonen Mikko, Wiklund Tom, Kettunen Raimo & Talja Martti 2012. Health coaching by telephony to support self-care in chronic diseases: clinical outcomes from The

TERVA randomized controlled trial. Bio Med Central Health Services Research 12:147. https://doi.org/10.1186/1472-6963-12-147

Peterson Martin 2009. An introduction to decision theory. New York: Cambridge University Press.

Pitkälä Kaisu, Savikko Niina & Routasalo Pirkko 2005. Geriatrisen kuntoutuksen tutkimus- ja kehittämishanke. Tutkimusraportti 10, Kuntoutuspolun solmukohtia. Helsinki: Vanhustyön Keskusliitto.

Pohjola Anneli, Kemppainen Tarja & Väyrynen Sanna 2012. Sosiaalityön vaikuttavuus. Rovaniemi: Lapin Yliopistokustannus.

Productivity Commission 2013. On efficiency and effectiveness: some definitions. Staff Research Note. Canberra.

Rajavaara Marketta 2006. Yhteiskuntaan vaikuttava Kela. Katsaus vaikuttavuuden käsitteisiin ja arviointiin. Sosiaali- ja Terveysturvan Katsauksia.Kansaneläkelaitos, tutkimusosasto.

Ranganathan Priya, Pramesh CS & Aggarwal Rakesh 2016. Common pitfalls in statistical analysis: intention-to-treat versus per-protocol analysis. Perspectives in Clinical Research 7(3), 144–146.

Räsänen Pirjo, Roine Eija, Sintonen Harri, Semberg-Konttinen Virpi, Ryynänen Olli-Pekka & Roine Risto 2006. Use of quality-adjusted life years for the estimation of effectiveness of health care: a systematic literature review. Finohta's Report 2006, 29.

Rieken Mieke, Strucmann Verena & Melchieorre Maria on behalf of the ICARE partners. ICARE 4EU. (2013). Improving care for people with multiple chronic condition in Europe. Eurohealth 19(3), 29–31.

Rollnick Stephen 2007. Motivational interviewing in health care: helping patients change behavior. New York: Guilford Press

Rossi Peter 2004. My views of evaluation and their origins. In: Alkin, Marvin C. (Ed). Evaluation roots. Tracing theorists' views and influences. Thousand Oaks, CA: Sage.

Routasalo Pirkko, Airaksinen Marja, Mäntyranta Taina & Pitkälä Kaisu 2009. Supporting a patient's self-management. Duodecim Lääketieteellinen Aikakauskirja 125, 2351–2359.

Russ Frederick 1996. Leadership, decision making and performance of sales managers: a multi-level approach. Journal of Personal Selling & Sales Management, 16(3), 1–15.

Salminen Ari 2011. Mitä on hallintotiede – tieteenalan identiteetin rakentaminen. Teoksessa Virtanen Turo, Ahonen Pertti, Syväjärvi Antti, Vartiainen Pirkko, Vartola Juha ja Vuori Jari (toim). Suomalainen hallinnon tutkimus: mistä, mitä, minne? Tampere. Tampere University Press.

Salminen Ari 1993. Hallintotiede: organisaatioiden hallinnolliset perusteet. Helsinki: Painatuskeskus.

Sanderson Colin 2006. Analytical models for decision making. Maidenhead: Open University Press.

Scheller-Kreinsen David, Blumel Miriam & Busse Reinhard 2009. Chronic disease management in Europe. Eurohealth 15 (1) 1–4;1, 1091.

Schulz Kenneth, Altman Douglas & Moher David 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. British Medical Journal 340, 698–702.

Scuffham Paul, Byrnes Joshua, Pollicino Christine, Cross David, Goldstein Stan & Ng Shu-Kay 2018. The impact of population-based disease management services on health care utilisation and costs: results of the CAPICHe trial. Journal of General Internal Medicine 34(1), 41–48. https://doi.org/10.1007/s11606-018-4682-5

Shapira Zur 1997. Organizational decision making. New York. Cambridge University Press.

Sharma Anjana, Willard-Grace Rachel, Hessler Danielle, Bodenheimer Thomas & Thom David 2016. What happens after health coaching? Observational study 1 year following a randomized controlled trial. Annals of Family Medicine14(3), 200–207. https://doi.org/10.1370/afm.1924

Sherifali Diana, Viscardi Virgnia, Bai Johnny-Wei & Ali Muhammad Usman 2016. Evaluating the effect of a diabetes health coach in individuals with type 2 diabetes. Canadian Journal of Diabetes, 40(1), 84–94.

Silvennoinen-Nuora Leena 2010. Vaikuttavuuden arviointi hoitoketjussa - mikä mahdollistaa vaikuttavuuden ja vaikuttavuuden arvioinnin. Hallintotiede, Kauppa- ja hallintotieteiden tiedekunta. Väitöskirja. Tampere University Press.

Simon Herbert 1979. Päätöksenteko ja hallinto. Espoo. Weilin Göös.
Simonen Outi 2012. Vaikuttavuustiedon hyödyntäminen erikoissairaanhoidon johtamisessa. Sosiaali- ja terveyspolitiikka, Terveystieteiden yksikkö. Väitöskirja. Tampere. Tampere University Press.
Simonen Outi 2011. Knowledge of effectiveness and its application in secondary healthcare management. International Journal of Productivity and Performance Management 60(8), 797–812.
Simonen Outi, Vatanen Elina, Konu Anne & Blom Marja 2009. Effectiveness in political—administrative decision-making in specialized healthcare. Scandinavian Journal of Public Health 37(5), 494–502.
Sinervo Lotta-Maria 2011. Kunnan talouden tasapaino monitulkintaisena ilmiönä. Kunnallistalous, Johtamiskorkeakoulu. Väitöskirja. Tampere University Press.
Sinkkonen Sirkka & Kinnunen Juha 1999. Terveystieteiden eriytymisprosessi, terveyshallintotieteen identiteetti ja suhde lähitieteisiin. Hoitotiede 11(6), 311–324.
Sinkkonen Sirkka & Kinnunen Juha 1994. Arviointi ja seuranta julkisella sektorilla. Kuopion yliopiston julkaisuja, E, Yhteiskuntatieteet, ISSN 1235-0494; 22. Terveyshallinnon ja -talouden laitos. Kuopion yliopisto,
Sintonen Harri 2001. The 15D instrument of health-related quality of life: properties and applications. Annals of Medicine 33(5), 328–336.
Sintonen Harri & Pekurinen Markku 2006. Terveystaloustiede. Porvoo, Helsinki: WSOY.
Smith Sarah 2005. Health care evaluation. Maidenhead: McGraw-Hill Education.
Steventon Adam, Tunkel Sarah, Blunt Ian & Bardsley Martin 2013. Effect of telephone health coaching (Birmingham OwnHealth) on hospital use and associated costs: cohort study with matched controls. British Medical Journal 2013;347:f4585 https://doi.org/10.1136/bmj.f4585
Stratton Irene, Adler Amanda, Neil HAW, Matthews David, Manley Susan, Cull Carole, Hadden David, Turner Robert & Holman Rury 2000. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. British Medical Journal 321(7258), 405–412.

Orchard Trevor, Dorman Janice, Maser Raelene, Becker Dorothy, Ellis Demetrius, LaPorte Ronald, Kuller Lewis, Wolfson Sidney & Drash Allan 1990. Factors associated with avoidance of severe complications after 25 yr of IDDM. Pittsburgh Epidemiology of Diabetes Complications Study I. Diabetes Care 13(7), 741–747.

The Diabetes Control and Complications Trial Research Group (DCCT) 1993. The effect of intensive treatment of diabetes on long-term complications in insulin-dependent diabetes mellitus, New England Journal of Medicine 329(14), 1035–1036.

The Look AHEAD Research Group 2013. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. The New England Journal of Medicine 369, 145–154.

The TERVA-protocol 2007. Health coach project in Päijät-Häme.

Thom David, Wolf Jessica, Gardner Heather, DeVore Denice, Lin Michael, Ma Andy, George Saba & Ibarra-Castro Ana 2016. A qualitative study of how health coaches support patients in making health-related decisions and behavioral changes. Annals of Family Medicine 14(6), 509–16. https://doi.org/10.1370/afm.1988

Tiede Michel, Dwinger Sarah, Herbarth Lutz, Härter Martin & Dirmaier Jörg 2017. Long-term effectiveness of telephone-based health coaching for heart failure patients: a post-only randomised controlled trial. Journal of Telemedicine and Telecare 23(8), 716–724. https://doi.org/10.1177/1357633X16668436

Tomer John 1992. Rational organizational decision making in the human firm: A socio-economic model. Journal of Socio-Economics 21(2), 85.

Trappenburg Jaap 2013. Self-management: One size does not fit all. Patient Education and Counseling 92(1), 134–137.

Tuomi Jouni & Sarajärvi Anneli 2009. Laadullinen tutkimus ja sisällönanalyysi. Helsinki: Kustannusosakeyhtiö Tammi.

Type 2 diabetes. Current Care Guidelines. Working group set up by the Finnish Medical Society Duodecim and the Finnish Cardiac Society. Helsinki: The Finnish Medical Society Duodecim, 2017 (referred May 27, 2020). Available online at: www.kaypahoito.fi

Tähtinen Juhani, Laakkonen Eero & Broberg Mari 2011. Tilastollisen aineiston käsittelyn ja tulkinnan perusteita. Turku: Turun Yliopisto.

Ueki Kohjiro & J-DOIT3 Research Group 2017. Effect of an intensified multifactorial intervention on cardiovascular outcomes and mortality in type 2 diabetes (J-DOIT3): an open-label, randomised controlled trial. The Lancet Diabetes & Endocrinology 5(12), 951–964.

UKPDS group 1998. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes *(UKPDS 34)*. Lancet 352(9131), 854–865.

Uzonwanne Francis 2016. Rational model of decision making. In: Farazmand A. (Ed). Global encyclopedia of public administration, public policy, and governance. New York. Springer.

Vale Margarite, Jelinek Michael, Best James, Dart Anthony, Grigg Leeanne, Hare David, Ho Betty, Newman Robert & McNeil John 2003. Coaching patients on Achieving Cardiovascular Health (COACH): a multicenter randomized trial in patients with coronary heart disease. Archives of Internal Medicine 163(22), 2775–2783.

Van Hout Ben, Maiwenn Al, Gilad Gordon & Rutten Frans 1994. Costs, effects and C/E-ratios alongside a clinical trial. Health Economics 3(5), 309.

van Loon Miriam, van Leeuwen K, Ostelo R, Bosmans J & Widdershoven G 2018. Quality of life in a broader perspective: Does ASCOT reflect the capability approach? Quality of Life Research 27(5), 1181–1189.

Vedung Evert 2003. Arvointiaalto ja sen liikkeelle panevat voimat. Stakes, FinSoc työpapereita 2/2003. Helsinki.

Virtanen Petri 2007. Arviointi: arviointitiedon luonne, tuottaminen ja hyödyntäminen. Helsinki: Edita.

Vuori Jari 2005. Terveys ja johtaminen: terveyshallintotiede terveydenhuollon työyhteisöissä. Porvoo, Helsinki: WSOY.

Wagner Edward, Austin Brian, Davis Connie, Hindmarsh Mike, Schaefer Judith & Bonomi Amy 2001. Improving chronic illness care: translating evidence into action. Health Affairs 20(6), 64-78.

Wagner Todd, Willard-Grace Rachel, Chen Ellen, Bodenheimer Thomas & Thom David 2016. Costs for a health coaching intervention for chronic care management. American Journal of Managed Care 22, (4)e141–e146.

Wayne Noah, Perez Daniel, Kaplan David & Ritvo Paul 2015. Health coaching reduces hba1c in type 2 diabetic patients from a lower-socioeconomic status community: a randomized controlled trial. Journal of Medical Internet Research 17(10), e224.

Weiss Carol 1972. Evaluation research: methods for assessing program effectiveness. Englewood Cliffs NJ: Prentice-Hall.

Wennberg David, Marr Amy, Lang Lance, O'Malley Stephen & Bennett George 2010. A randomized trial of a telephone care-management strategy. The New England Journal of Medicine 363, 1245–1255.

Willard-Grace Rachel, Chen Ellen, Hessler Danielle, Devore Denise, Prado Camille, Bodenheimer Thomas & Thom David 2015. Health coaching by medical assistants to improve control of diabetes, hypertension, and hyperlipidemia in low-income patients: a randomized controlled trial. Annals of Family Medicine 13(2), 130–138.

Wolever Ruth, Dreusicke M, Fikkan J, Hawkins TV, Yeung S, Wakefield J, Duda L, Flowers P, Cook C & Skinner E 2010. Integrative health coaching for patients with type 2 diabetes: a randomized clinical trial. The Diabetes Educator 36(4), 629–639.

Wolever Ruth, Simmons Ann, Sforzo Gary, Dill Diana, Kaye Miranda, Bechard Elizabeth, Southard Elaine, Kennedy Mary, Vosloo Justine & Yang Nancy 2013. A systematic review of the literature on health and wellness coaching: defining a key behavioral intervention in healthcare. Global Advances in Health and Medicine 2(4), 38–57.

World Health Organization 2003. Pitkäaikaisiin hoitoihin sitoutuminen. Näyttöä toiminnan tueksi. (Adherence to long-term therapies: evidence for action) Helsinki: Lääketietokeskus Oy.

Appendices

APPENDIX 1.

Table 1. Articles included in literature review.

Author(s), year	Health coaching target group	Intervention	Effectiveness	Methods, n, follow-up time	Main results
Clinical and behavioural outcomes					
Chapman et al. (2018)	Type 2 diabetes	Health coaching based on principles of Motivational Interviewing, 18 months	Glycated hemoglobin (HbA1c), secondary outcomes: encompassed a suite of physiological, psychological and self-care measures.	RCT (cluster): Community Health Stations, 21 (intervention, 18 (control), participants n=372 (intervention) n=339 (control), 18-month follow-up	No differential treatment effect was found for HbA1c or any specified secondary outcomes.
Dennis et al. (2013)	Chronic diseases: T2D, coronary artery disease, congestive cardiac failure, chronic obstruktive pulmonary disease, hypertension.	Two-way conversations by telephone or video. Behaviour change, goal setting and empowerment were essential features of health coaching.	Health behaviour, self-efficacy, health status, and satisfaction with service.	Review (30 studies)	Telephone health coaching can improve health behaviour, self-efficacy and health status. Evidence were limited in Quality of life, satisfaction with services, and improvements of health service use.
Karhula et al. (2015)	Type 2 diabetes and heart diseases	Telemonitoring and mobile-phone based health coaching	HRQol (Short Form 36), HBA1c, blood pressure, weight, waist circumference, and lipid levels	RCT (modified ITT), n= 337 (intervntion) n=134 (control), 12-month follow-up	Health coaching program did not improve patients HRQoL or clinical outcomes.

Author(s), year	Health coaching target group	Intervention	Effectiveness	Methods, n, follow-up time	Main results
Kivelä et al. (2014)	Chronic diseases (e.g. T2D, coronary artery disease, heart failure, dyslipidemic, overweight, rheumatoid arthiritis, cancer pain)	Delivered: telephone, internet, face-to-face, e-mail, frequency 3–14 times, lenght 3 weeks to 18 months, mostly nurses	Physiologigal, behavioral, psychological, social outcomes	Systematic review (13 studies)	Sigfnificantly improved results were reported in 11 of the 13 studies: statistically significant outcomes were found in the physiological, behavioral, psychological, and social outcomes.
Panagioti et al. (2018)	Multimorbidity eldery patient	Telephone healht coaching (six mothly phone calls, around 20 min.), delivered by 'health advisors'	Patient Activate Measure (PAM), Quality of life (WHOQoL-BREF), depression, (Mental Health Inventory, MHI-5), self-care (The summary of Diabetes self-care, SDSCA), cost and Qol	Trials within cohorts' design (ITT) n=504 (intervention) n= 802 (control), 20-month follow-up	ITT -analysis did not improve any outcome of PAM, QoL, MHI-5, SDSCA, in cost-effectiveness analysis lower levels of emergency care use was found but increased used of planned services with higher costs and QALY was gained.
Sherifali et al. (2015)	Type 2 diabetes	Telephone-based (15 to 60 minutes), combination of telephone and face-to-face, combinations of internet, phone and face-to-face strategies; 6 studies had coaching intervention durations 6 months or fewer, and 2 studies were longer than 6 months.	To evaluate effects of health coaching on HbA1c level	Review (8 studies: 1 RCT 2) ≥18 years of age with type 2 diabetes; 3) reported health-coaching intervention (in addition to usual care or self-management education/ support); 4) conducted by a health professional; and 5) report a mean change in A1C levels.)	The pooled effect of health coaching was a statistically significant reduction of A1C levels: the longer (>6months) interventions resulted in significantly greater decreases in HbA1C levels

Author(s), year	Health coaching target group	Intervention	Effectiveness	Methods, n, follow-up time	Main results
Tiede et al. (2017)	Heart failure patients	Telephone-based health coaching	Quality of life, changes in depression and anxiety, health-related control beliefs, control preference, health risk behaviour, healt-related behaviours	RCT: a post-only, 24 and 28 moths follow-up	Health coaching has no effect on QoL, anxiety and depression.
Tuluce & Kutluturkan (2018)	Chronic obstructive pulmonary disease.	12 face-to-face coaching sessions, 30 to 45 minutes/session, during 20 week	Respiratory functions, treatment adherence, self-efficacy, and quality of life.	Experimental non-randomized study, n=27 (intervention), n=27 (control), 20-week follow-up.	Statistically significant different were found self-efficacy and St. George Respiratory Questionnaire and degree of treatment adherence between the groups.
Vale et al. (2003)	Coronary heart disease	Telephone-based with e-mails during six months	To evaluate effect on total cholesterol,and physical, nutritional and psychological factors	RCT: ITT, n=398 (intervention), n=394 (control), 6-month follow-up	The intervention group achieved statistically significant reduction in total cholesterol due to intervention.
Wayne et al. (2015)	Type 2 diabetes	Health coaching intervention with and without the use of mobile phones, six months intervention.	To evaluate the change in HbA1c level from baseline to 6 months, secondary oucomes were weight, waist circumference, body mass index (BMI), satisfaction with life, depression and anxiety and quality of life (SF-12).	RCT: ITT and PP analysis, n= 67 (intervention) n= 64 (control group), 6-month follow-up	No significant between-group differences in change of HbA1c using intention-to-treat or per-protocol anlysis. Intervention group participants achieved significant decreases in weight and waist circumference.

Author(s), year	Health coaching target group	Intervention	Effectiveness	Methods, n, follow-up time	Main results
Willard-Grace et al. (2015)	Cardiovascular and metabolic risk factors	Delivered: face-to-face contact at least every 3 months and telephone calls at least once a month by trained medical assistant	Clinical outcomes: HbA1c, systolic blood pressure, low-density lipoprotein (LDL)	RCT: (ITT) n=224 (intervention), n=217 (control), 12-month follow-up.	HbA1c and LDL level improved, but not blood pressure compared with usual care.
Wolever et al. (2010)	Type 2 diabetes	Telephone-based 14 x 30 min during six months	To evaluate the effectiveness of integrative health coaching on psychosocial factors, behavior change, and glycemic control.	RCT: ITT, n=30 (intervention), n=26 (control), 6-month follow-up	Patient activation, perceived social support, and benefit finding all increased in the coaching group. HbA1c reduced significantly in the coaching group.
Health care utilization and costs					
Benzo et al. (2015)	Chronic Obstructive Pulmonary Diseases	Face-to-face and telephone	To evaluate the effect of health coaching on the rate of COPD readmissions.	RCT n=108 (intervention), n=106 (usual care), 12-month follow-up (max)	The health coaching intervention had a significant effect on the rates of COPD hospitalization at 1, 3, and 6 months post–hospital discharge, but this effect faded at 12 months. Disease-specific quality of life improved significantly due the health coaching at 6 and 12 months, no differences between groups in physical activity at any time point.

Author(s), year	Health coaching target group	Intervention	Effectiveness	Methods, n, follow-up time	Main results
Billot et al. (2015)	Diabetes, hypertension, chronic obstruktive pulmonary disease, congestive heart failure, and coronary artery disease.	1) care coordination across sectors (acute, ambulatory, and community care from public and private sector), and clinical specialities 2) health coaching (management of lifestyle risk, and medication and self-management)	To evaluate health service utilization: avoidable hospitalizations, secondary outcomes: avoidable readmissions, avoidable beddays, unplanned beddays, unplanned hospitalizations, and all-cause death.	Propensity -matched cohort n= 30 057 (intervention) n= 30 057 (matched controls), 15-month follow-up.	Intervention group was associated with an increased in avoidable hospital admissions compared with matched controls, no difference in the other outcomes or death.
Byrnes et al. (2018)	Cardiovascular disease (CVD)	Telephone-based with e-mails, six months coaching	To evaluate overall survival, hospital utilization and costs in long-term.	A prospective parallel-group case-control stydy with controls (propensity matched), n=512 (intervention), n=512 (control), 6.35-year follow-up.	Significant reduction in overall mortality patients who received four or more coaching session. Total costs of health insures was also lower in intervention group.
Hale & Giese (2017)	Not defined	Intervention varied	To evaluate health coaching effectiveness and cost-effectiveness	Integrative review (27 articles)	Majority of evidence supported effectiveness of healht coaching especially chronic disease such as hypertension, diabetes and hyperlipidemia. One of three cost-effectiveness article affirmed the effectiveness of health coaching .

Author(s), year	Health coaching target group	Intervention	Effectiveness	Methods, n, follow-up time	Main results
Hutchison & Breckon (2011)	Long-term conditions	One-to-one telephone coaching	To evaluate evidence of one-to-one telephone coaching	Review (34 articles)	32 studies reported outcomes in favour of telephone healht coaching, only 5 studies reported assessment of cost-effectiveness or costs
Härter et al. (2016)	Multiple conditions, heart failure and chronic mental illness	Telephone based, one phone call at least every six weeks (averige 12.9 calls) health coaching period one year, intervention descriped, health care professionals	To evaluate: 1. Time from enrollment until hospital admission 2. probability of hospital admission, number of daily difined medication doses, frequency and duration of inability to work 3. mortality within two years.	RCT (modified ITT) n=3229 (intervention) n=3223 control, 2-year follow-up	All groups: no differencies in hospital readmissions, Multiple condition group: hospital admission was higher in intervention group and sifnificant difference in mortality due to intervention , Heart failure group: significant reduction of number of hospital admission in intervention group and sifnificant difference in mortality due to intervention
Jonk et al. (2015)	High risk health plan enrollees (high cost, multiple comorbidities, and/or high adjusted clinical group (ACG) risk score)	Telephone, average 5-6 phone session during 6 months	To evaluate the effect of health coaching on inpatient emergency room, outpatient, and prescription drug expenditures	Quasi-experimental prepost design n=9178 intervention, 6-month follow-up	Health coaching led to significant reductions in outpatient and total expenditures for high-risk plan enrollees.

Author(s), year	Health coaching target group	Intervention	Effectiveness	Methods, n, follow-up time	Main results
Lin et al. (2012)	Medicaid members with chronic disease, aged 18 to 64, and 2 acute health service events of hospitalizations and/or ED visits within a 12-month period.	Telephone-based health coaching	To evaluate healthcare utilization (hospitalizations, emergency department (ED) visits, ambulatory care visits) and expenditures.	Matched comparison group 1 year before and 2 years after enrollment	Health coaching disease management program did not demonstrate significant effects on healthcare utilization and expenditures
Scuffham et al. (2018)	Diagnosis of heart failure, chronic obstructive pulmonary disease (COPD), coronary artery disease (CAD), diabetes, or low back pain, with predicted high cost claims.	Telephone-based and individual educational material	Hospital admissions, including same-day, medical and prostheses hospital claims.	Parallel-group randomised controlled trial, intention-to-treat analysis n=35 535 (intervention), n=8 883 (usual care), 12-month follow-up	Health coaching was not effective to reduce the total cost.
Sharma et al. (2016)	Uncontrolled diabetes, hypertension, or hyperlipidemia.	Healht coaches accompanied patient with clinical visits, met patient before visits, remained in exam room during visit, reviewd care plan after visits and contact by phone between visits. Coaching lasted 12 monhts.	To evaluate maintenance of improved cardiovascular risk factors (HbA1c, systolic blood pressure or LDL) following a 12 months health coaching intervention after 24 months.	RCT n= 157 (intervention), n= 133 (usual care), 12 month follow-up.	The most improved clinical outcomes persisted 1 year after the completion of the health coaching intervention.

Author(s), year	Health coaching target group	Intervention	Effectiveness	Methods, n, follow-up time	Main results
Steventon et al. (2013)	Heart failure, coronary heart disease, diabetes, or chronic obstructive pulmonary disease; and a history of inpatient or outpatient hospital use.	Telephone-based, one to four times in month, median lenght of coaching was 25.5 months.	To evaluate effect of telephone health coaching on hospital use and associated costs 12 months after enrollment.	Cohort study with matched controls n=2698 (intervention), n=2698 (matched controls) 12-month follow-up	Health coaching did not lead reductions in hospital admissions or secondary care costs.
Wagner et al. (2016)	Chronic diseases: T2D, hypertension, and/or hyperlipidemia.	Face-to-face and telephone: average 5-6 phone calls during six months, coaches attended medical visits with patient, met them before and after visits and called called or met them between visits.	To evaluate additional costs of healht coaches comparing savings in first year	RCT: n=224 (intervention), n=217 (usual care), 12-month follow-up	Health coaches was not pay themselves by reducing health care utilization in the first year.
Wennberg et al. (2010)	Patients with high predictive risk e.g. future hospital costs	Telephone based supporting with Web-links, videos and printed material, length and frequency of calls based on individual need (follow-up time was one year)		RCT n= 87 243 (usual support) 86 877 (enhanced-support),12-month follow-up	A telephone care-management program was successful in reducing medical costs and hospitalizations in populaton based-study.

APPENDIX 2.

Table 1. Cumulative costs/patient/year and patients at risk per year in intention to treat (ITT) analysis.

Costs/year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
ITT invervention	2705	5983	9859	15704	21638	27204	32806	39706
ITT control	2425	5515	10091	15815	22132	28880	34702	40916
No at risk/year								
ITT invervention	1004	970	941	897	854	821	793	764
ITT control	486	477	455	444	431	409	382	359

Table 2. Cumulative costs/patient/year and patients at risk per year in per protocol (PP) analysis.

Costs/year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
PP invervention	2561	5734	9168	14473	19955	24883	29899	35863
PP control	2279	5462	10246	16106	22461	29296	35266	41816
No at risk/year								
PP invervention	844	820	796	760	727	703	679	656
PP control	446	438	417	406	398	379	355	334

APPENDIX 3.

Table 1. Cox proportional hazard ratio (HR), all patient with event and p-values comparing the primary, secondary and other outcomes between intervention and control groups in per protocol (PP)-analysis. (Abbreviations: CVD=cardio vascular disease, AMI= myocardial infarction, UAP=unstable angina pectoris, CABG= coronary artery bypass grafting, PTCA= Percutaneous transluminal coronary angioplasty, PVD= Peripheral vascular disease, CHF=cardiac heart failure

Outcome	Patients with event	Intervention		Control		HR	(5% CI)	(95% CI)	p-value
Primary outcome	**no.**	**no.**	**(rate/100 person years)**	**no.**	**(rate/100 person years)**				
Death from CVD causes or stroke or AMI or UAP	324	203	3,34	121	3,73	0,87	0,71	1,11	0,30
Secondary outcomes									
Death from CVD causes or stroke or AMI	308	193	3,17	115	3,55	0,88	0,70	1,11	0,30
Death from any cause or stroke or AMI									
Death from any cause or stroke or AMI or UAP or CABG or PTCA or CHF or peripheral vasc. disease	557	352	5,79	205	6,33	0,89	0,75	1,06	0,22
Other outcomes									
Death, all causes	316	197	3,24	119	3,67	0,88	0,70	1,10	0,27
AMI, fatal or non-fatal	142	86	1,41	56	1,73	0,81	0,57	1,12	0,21
Stroke, fatal or non-fatal	85	50	0,82	35	1,08	0,75	0,49	1,16	0,19
Renal insufficiency	58	30	0,49	28	0,86	0,56	0,34	0,94	0,02
Peripheral vasc disease	77	50	0,82	27	0,83	0,98	0,62	1,58	0,90
Hospitalization due to CHF	210	130	2,14	80	2,47	0,85	0,67	1,12	0,27

Table 2. Cox proportional hazard ratio (HR), coronary artery disease (CAD) patient with event and p-values comparing the primary, secondary and other outcomes between intervention and control groups in per protocol (PP)-analysis.(Abbreviations: CVD=cardio vascular disease, AMI= myocardial infarction, UAP=unstable angina pectoris, CABG= coronary artery bypass grafting, PTCA= Percutaneous transluminal coronary angioplasty, PVD= Peripheral vascular disease, CHF=cardiac heart failure

Outcome	**Patients with event**	**Intervention**		**Control**		**HR**	**(5% CI)**	**(95% CI)**	**p-value**
Primary outcome	**no.**	**no.**	**(rate/100 person years)**	**no.**	**(rate/100 person years)**				
Death from CVD causes or stroke or AMI or UAP	104	60	3,98	44	4,93	0,80	0,54	1,19	0,28
Secondary outcomes									
Death from CVD causes or stroke or AMI	95	54	3,58	41	4,60	0,75	0,51	1,15	0,20
Death from any cause or stroke or AMI	138	77	5,10	61	6,84	0,73	0,52	1,02	0,07
Death from any cause or stroke or AMI or UAP or CABG or PTCA or CHF or peripheral vasc. disease	179	101	6,69	78	8,74	0,73	0,54	0,99	0,04
Other outcomes									
Death, all causes	94	52	3,45	42	4,71	0,74	0,49	1,10	0,13
AMI, fatal or non-fatal	44	23	1,52	21	2,35	0,65	0,36	1,17	0,15
Stroke, fatal or non-fatal	20	10	0,66	10	1,12	0,58	0,24	1,41	0,23
Renal insufficiency	16	6	0,40	10	1,12	0,35	0,13	0,97	0,04
Peripheral vasc disease	17	9	0,60	8	0,90	0,62	0,26	1,75	0,42
Hospitalization due to CHF	85	48	3,18	37	4,15	0,76	0,50	1,17	0,22

Table 3. Cox proportional hazard ratio (HR), type 2 diabetes (T2D) patients with event and p-values comparing the primary, secondary and other outcomes between intervention and control groups in PP-analysis. (Abbreviations: CVD=cardio vascular disease, AMI= myocardial infarction, UAP=unstable angina pectoris, CABG= coronary artery bypass grafting, PTCA= Percutaneous transluminal coronary angioplasty, PVD= Peripheral vascular disease, CHF=cardiac heart failure

Outcome	**Patients with event**	**Intervention**		**Control**		**HR**	**(5% CI)**	**(95% CI)**	**p-value**
Primary outcome	**no.**	**no.**	**(rate/100 person years)**	**no.**	**(rate/100 person years)**				
Death from CVD causes or stroke or AMI or UAP	220	143	3,13	77	3,28	0,94	0,71	1,24	0,68
Secondary outcomes									
Death from CVD causes or stroke or AMI	213	139	3,04	74	3,15	0,96	0,72	1,27	0,76
Death from any cause or stroke or AMI	295	194	4,24	101	4,30	0,98	0,77	1,24	0,88
Death from any cause or stroke or AMI or UAP or CABG or PTCA or CHF or peripheral vasc. disease	378	251	5,49	127	5,41	1,00	0,81	1,24	0,95
Other outcomes									
Death, all causes	222	145	3,17	77	3,28	0,96	0,73	1,27	0,80
AMI, fatal or non-fatal	98	63	1,38	35	1,49	0,91	0,60	1,37	0,64
Stroke, fatal or non-fatal	65	40	0,87	25	1,07	0,81	0,49	1,33	0,40
Renal insufficiency	42	24	0,52	18	0,77	0,67	0,36	1,24	0,20
Peripheral vasc disease	60	41	0,90	19	0,81	1,10	0,64	1,87	0,73
Hospitalization due to CHF	125	82	1,79	43	1,83	0,96	0,66	1,39	0,84

Articles

ARTICLE I

Patja Kristiina, Absetz Pilvikki, Auvinen Anssi, Tokola Kari, Kytö Janne, Oksman Erja, Kuronen Risto, Ovaska Timo, Harno Kari, Nenonen Mikko, Wiklund Tom, Kettunen Raimo & Talja Martti 2012. Health coaching by telephony to support self-care in chronic diseases: clinical outcomes from The TERVA randomized controlled trial. BMC Health Services Research 12:147. https://doi.org/10.1186/1472-6963-12-147

Reprinted with permission from BioMed Central.

ARTICLE II

Oksman Erja, Linna Miika, Hörhammer Iiris, Lammintakanen Johanna & Talja Martti 2017. Cost-effectiveness analysis for a tele-based health coaching program for chronic disease in primary care. BMC Health Services Research 17(1), 138. https://doi.org/10.1186/s12913-017-2088-4

Reprinted with permission from BioMed Central.

ARTICLE III

Mustonen Erja, Hörhammer Iiris, Absetz Pilvikki, Patja Kristiina, Lammintakanen Johanna, Talja Martti, Kuronen Risto & Linna Miika 2020. Eight-year post-trial follow-up of health care and long-term care costs of tele-based health coaching. Health Services Research 55 (2), 211–217. https://doi.org/10.1111/1475-6773.13251

Reprinted with permission from Wiley Publications.

ARTICLE IV

Mustonen Erja, Hörhammer Iiris, Patja Kristiina, Absetz Pilvikki, Lammintakanen Johanna, Talja Martti, Kuronen Risto & Linna Miika. Eight-year post-trial follow-up of morbidity and mortality of telephone health coaching. Submitted.

ARTICLE I

Patja Kristiina, Absetz Pilvikki, Auvinen Anssi, Tokola Kari, Kytö Janne, Oksman Erja, Kuronen Risto, Ovaska Timo, Harno Kari, Nenonen Mikko, Wiklund Tom, Kettunen Raimo & Talja Martti 2012. Health coaching by telephony to support self-care in chronic diseases: clinical outcomes from The TERVA randomized controlled trial. BMC Health Services Research 12:147. https://doi.org/10.1186/1472-6963-12-147

BMC Health Services Research

Health coaching by telephony to support self-care in chronic diseases: clinical outcomes from The TERVA randomized controlled trial

BMC Health Services Research 2012, **12**:147 doi:10.1186/1472-6963-12-147

Kristiina Patja (kristiina.patja@promedico.fi)
Pilvikki Absetz (pilvikki.ansetz@thl.fi)
Anssi Auvinen (anssi.auvinen@uta.fi)
Kari Tokola (kari.tokola@uta.fi)
Janne Kytö (janne.kyto@helsinki.fi)
Erja Oksman (erja.oksman@phsotey.fi)
Risto Kuronen (risto.kuronen@phsotey.fi)
Timo Ovaska (timo.ovaska@pfizer.com)
Kari Harno (kari.harno@kolumbus.fi)
Mikko Nenonen (mikko.nenonen@fimnet.fi)
Tom Wiklund (tom.wiklund@pfizer.com)
Raimo Kettunen (raimo.kettunen@phsotey.fi)
Martti Talja (martti.talja@phsotey.fi)

ISSN	1472-6963
Article type	Research article
Submission date	6 November 2011
Acceptance date	14 April 2012
Publication date	10 June 2012
Article URL	http://www.biomedcentral.com/1472-6963/12/147

Health coaching by telephony to support self-care in chronic diseases: clinical outcomes from The TERVA randomized controlled trial

Kristiina Patja[1*]
[*]Corresponding author
Email: Kristiina.patja@promedico.fi

Pilvikki Absetz[1]
Email: Pilvikki.absetz@thl.fi

Anssi Auvinen[2]
Email: Anssi.auvinen@uta.fi

Kari Tokola[2]
Email: Kari.tokola@uta.fi

Janne Kytö[3]
Email: Janne.kyto@helsinki.fi

Erja Oksman[4]
Email: Erja.oksman@phsotey.fi

Risto Kuronen[4]
Email: Risto.kuronen@phsotey.fi

Timo Ovaska[3]
Email: Timo.ovaska@pfizer.com

Kari Harno[5]
Email: Kari.harno@kolumbus.fi

Mikko Nenonen[6]
Email: Mikko.nenonen@fimnet.fi

Tom Wiklund[3]
Email: Tom.wiklund@pfizer.fi

Raimo Kettunen[4]
Email: Raimo.kettunen@phsotey.fi

Martti Talja[4]
Email: Martti.talja@phsotey.fi

[1] National Institute of Health and Welfare, Helsinki, Finland

[2] School of Health Sciences, University of Tampere, Tampere, Finland

[3] Pfizer Oy, Helsinki, Finland

[4] Joint Authority for Päijät-Häme Social and Health Care, Lahti, Finland

[5] Department of Social and Health Management, University of Eastern Finland (Kuopio Campus), Kuopio, Finland

[6] Hartola Health Care Center Hartola, Finland and and Finnish Medical Association, Helsinki, Finland.

Abstract

Background

The aim was to evaluate the effect of a 12-month individualized health coaching intervention by telephony on clinical outcomes.

Methods

An open-label cluster-randomized parallel groups trial. Pre- and post-intervention anthropometric and blood pressure measurements by trained nurses, laboratory measures from electronic medical records (EMR). A total of 2594 patients filling inclusion criteria (age 45 years or older, with type 2 diabetes, coronary artery disease or congestive heart failure, and unmet treatment goals) were identified from EMRs, and 1535 patients (59%) gave consent and were randomized into intervention or control arm. Final analysis included 1221 (80%) participants with data on primary end-points both at entry and at end. Primary outcomes were systolic and diastolic blood pressure, serum total and LDL cholesterol concentration, waist circumference for all patients, glycated hemoglobin (HbA_{1c}) for diabetics and NYHA class in patients with congestive heart failure. The target effect was defined as a 10-percentage point increase in the proportion of patients reaching the treatment goal in the intervention arm.

Results

The proportion of patients with diastolic blood pressure initially above the target level decreasing to 85 mmHg or lower was 48% in the intervention arm and 37% in the control arm (difference 10.8%, 95% confidence interval 1.5–19.7%). No significant differences emerged between the arms in the other primary end-points. However, the target levels of systolic blood pressure and waist circumference were reached non-significantly more frequently in the intervention arm.

Conclusions

Individualized health coaching by telephony, as implemented in the trial was unable to achieve majority of the disease management clinical measures. To provide substantial benefits, interventions may need to be more intensive, target specific sub-groups, and/or to be fully integrated into local health care.

Trial registration

ClinicalTrials.gov Identifier: NCT00552903.

Background

Diabetes and cardiovascular diseases represent large and costly chronic healthcare challenges [1]. Preventative measures can effectively reduce costs [2]. Despite differences between different conditions, the expectations on the patients are similar: to cope with multiple medications and co-morbidities, to alter behavior, to deal with social and psychological impacts of symptoms and to interact with medical care [3,4].

Health care providers have a difficult task in trying to manage chronic disease care in complex service systems that are poorly designed to motivate, equip and empower patients to behavior changes [5-7]. Resources should aim at maximized health gains, and this requires reorientation of services [8]. High expectations are put on information technology solutions that have been shown highly effective in promoting lifestyle changes [9]. So far, comprehensive efforts to assess the impact of incorporating a range of IT tools in chronic disease management have been targeting single disease groups, such as CHD [10,11], heart failure [12] or diabetes [13,14] Taylor et al. 2003, but studies with several disease groups and/or co-morbidities are lacking.

While technology can be an effective way to improve reach of disease management interventions, still the content is more important. Health coaching, a collaborative process characterized by motivational communication, patient-defined goals related to disease management, and patient acceptance of accountability for decisions made [15] can utilize different sets of self-management tools (SMTs) to promote adoption of an active role in self-care by the patient [16]. Health coaching can improve quality, effectiveness and cost-effectiveness of disease management [17]. The TERVA trial is the first large randomized controlled trial to simultaneously evaluate tele-coaching in a real-world health care setting in three patient groups: congestive heart failure (CHF), coronary artery disease (CAD) and type 2 diabetes mellitus (T2D). The aim of the trial was to assess the effect of health coaching on clinical outcomes (risk determinants) after one-year intervention.

Methods

Trial design

The TERVA study is a randomized, open-label, parallel groups trial comparing health coaching and usual care. The primary end points were defined as 10-percentage point difference between arms in increase in the proportion of participants reaching the target level in five global and two patient-group specific clinical parameters at 12 months (Table 1). The targets were set in accordance with Finnish evidence-based guidelines.

Table 1 Primary and secondary end points of TERVA trial

Primary end points
▪ Provider-measured BP ≤140/85 mmHg
▪ Total cholesterol ≤4.5 mmol/L
▪ LDL ≤2.5 mmol/L
▪ Waist circumference ≤94 cm for men and ≤80 cm for women – later revised as 90 cm for women and 100 cm for men based on national guidelines
For congestive heart failure an additional end-point:
▪ Improved or maintained NYHA class
For participants with T2D:
▪ HbA_{1c} ≤7%

Measures

Research nurses, unaware of the allocation, measured blood pressure and waist circumference in both arms. The laboratory results were extracted from the electronic medical records (EMR) at both entry and end of the intervention (at entry between 3 months before to 1 month after and at end 11 to 15 months from date of consent). NYHA-class was obtained from study questionnaires at entry and end of follow-up.

Identification and enrollment

Patients were enrolled from Päijät-Häme in the Southern Finland, a region with a population of 212,000. The target population was initially identified from primary care and hospital registries and records, followed by a detailed assessment of medical records (Table 2). Patients with more than one condition were enrolled in the following hierarchy: CHF - CAD - T2D, so CHF patients could have CAD and/or T2D, but not the other way around. All eligible patients were sent an information letter and a consent form in four batches during a 12-month period in 2007–2008 with one reminder for non-responders followed by a telephone call. Of the 2594 eligible patients 59.2% (1535) gave consent and were invited for an examination and interview by the research nurse, and 1225 (79.8%) completed it. The final analysis included 1221 patients (80%) having data on primary end-points both at entry and at end of follow-up. 1215 had both baseline and end of study measurements of waist circumference and blood pressure available (812 or 87% of committed patients in the intervention arm and 403 or 87% in the control arm). Laboratory measures of lipids at both time points were available in EMRs only for a fifth of the patients, and HbA_{1c} for 54% of the patients with diabetes. The age and sex distribution of the drop-outs did not differ from the analyzed patients (mean ages 65.0 vs. 64.8 years, 60.6% vs. 58.1% men). There were no substantial differences between participants and drop-outs in the primary end-points at baseline.

Table 2 Eligibility and exclusion criteria of TERVA trial

Eligibility criteria for enrollment included:
1. Residents in the region of Päijät-Häme aged 45 years or older
2. One of the following diagnose
a. Heart failure with NYHA II or III, and a history of hospital admission for heart failure within the last 2 years

b. History of myocardial infarction or cardiac revascularization procedure, and one of the following (treated or untreated): blood pressure above 140/85 mmHg, total serum cholesterol concentration

   >4.5 mmol/L, serum LDL concentration >2.5 mmol/L

c T2D on medication and serum HbA_{1c} >7% without clinically evident cardiovascular diseases e.g. MI, stroke, peripheral vascular disease

Exclusion criteria:

- Inability to cooperate or participate
- Pregnancy
- Life expectancy less than 1 year
- Patients with major elective surgery planned within 6 months
- Patient has had major surgery within the last 2 months

Randomization

A cluster design was used to accommodate the effects of individual health coaches with multiple patients. The randomization algorithm was based on computer-generated random numbers. A stratified randomization with permuted blocks was used to ensure balanced distribution within disease group and municipality between the arms. A Zelen type randomization (2:1 ratio for intervention/control arm) was performed prior to consent (Figure 1).

Figure 1 Flow Diagram. Distribution of study population from those filling inclusion criteria by the healthcare charts to those completed the intervention

Intervention

Health coaching was delivered from November 2007 by seven experienced certified nurses or public health nurses. They were trained for four weeks in a tele-coaching model initially developed by Pfizer Health Solutions (PHS) but modified for the Finnish health care system. Patients in the intervention group were called monthly, altogether 10–11 times. After a brief engagement call, there was one broader needs assessment call, followed by monthly coaching calls and finally an evaluation call. In between the coaching calls there was an opportunity for brief follow-up calls, but these were rarely used. The coaching call topics were based on 8 key recommendations of the program, with variations due to individual patient's preferences (Figure 2). The behavior change component integrated behavior change techniques from the Self-Regulation Theory and supported by evidence, i.e., self-monitoring, goal setting, action planning, and feedback [18]. After the first two months, quality assurance measures were taken in the form of listening to randomly selected 2–3 calls from each coach. Call length was also monitored. Calls were found to be long, typically up to 60 min, and they were based on a coach driven information provision model, and very little concrete goal setting and action planning was done. To improve quality, an explicit structure following the self-regulation model was developed jointly with the coaches, and the maximum number of topics to be tackled during one call, was limited to three. Also, coaches were further trained in Motivational Interviewing techniques of active listening, and using open questions, reflection and summaries [19], and they all received two individual supervision sessions in self-monitoring and developing their coaching practices. With these measures, quality (defined as use of structure and Motivational Interviewing techniques, and concrete actions as outcomes of the calls) was improved while call length decreased to approximately 30 min. Self-care books prepared in collaboration with the Finnish Heart Association and the Finnish Diabetes Associations supported the coaching, and the coach had access to the patients' EMRs. Both trial arms continued to receive routine care.

Figure 2 Pfizer Health Solutions has developed a tele-coaching intervention with 5 key functions and 8 recommendations to engage, inform, involve, and empower the patients in self-care.

Statistical methods

A sample size of 1250 patients was calculated to provide adequate statistical power ($1-\beta \geq 0.8$) for detecting a 10 percentage point difference between the intervention arms (with 6 coaches) with conservative assumptions ($\alpha = 0.05$ two-sided, 50% of the patients in the control arm would reach target, a 10% drop out rate and 10% of the subjects not evaluable at the end of the trial), as long as the intracluster correlation did not exceed 0.01 [20].

Data analyses were conducted using multilevel methods (generalized linear mixed models) to account for the clustered design. The trial data had a two-level structure, where the health coaches constituted an upper level, within which the individual patients were distributed allowing for correlation at individual level within a cluster (variance components at the two levels).

A modified intention to treat analyses by trial arm was employed including all patients with data at entry and at the end of the 12-month follow-up. No substantial imbalance at baseline was found in the primary end-point variables between the arms (Table 3).

Table 3 Baseline data available from patients who were allocated to the study (intervention = 1034, control = 501)

	Type 2 Diabetes		Coronary artery disease		Congestive heart failure	
	Intervention	Control	Intervention	Control	Intervention	Control
Number of patients	770	359	172	97	92	45
Age (years)	64.6 (9.4)	65.6 (9.5)	65.4 (9.4)	66.0 (8.6)	67.3 (7.9)	62.4 (7.7)
Sex (% male)	58.3	54.0	69.2	73.2	65.2	64.4
Age at self reported year of diagnosis	54.3 (11.1) 311	56.1 (11.8) 145	60.8 (10.2) 121	61.3 (10.1) 69	63.9 (7.9) 51	56.8 (9.3) 25
Self-reported duration of disease	10.3 (8.4) 302	9.2 (8.2) 142	4.7 (6.6) 120	4.8 (7.6) 68	4.4 (5.4) 51	4.2 (3.7) 25
Body mass index	32.3 (6.2) 727	31.9 (5.7) 338	28.3 (4.1) 164	28.8 (5.1) 91	30.2 (6.6) 81	29.9 (6.9) 42
Waist circumference (M/F)	109.9 (14.4)/ 106.8 (15.5) 429/298	110.1 (12.8)/ 104.8 (15.1) 182/156	100.3 (11.0)/ 95.1 (12.8) 116/47	102.9 (11.7)/ 91.5 (12.6) 66/25	106.7 (12.8)/ 94.1 (14.6) 55/25	108.0 (17.9)/ 89.6 (17.3) 25/16
Systolic blood pressure (mmHg)	143.4 (20.0) 727	143.1 (20.2) 337	137.6 (20.5) 163	139.9 (18.2) 91	132.9 (23.7) 81	127.6 (22.7) 42
Diastolic blood pressure (mmHg)	84.2 (11.1) 727	84.1 (10.8) 338	81.0 (12.2) 163	82.1 (10.5) 91	81.3 (14.1) 81	80.2 (11.5) 42
Serum total cholesterol (mmol/l)	4.4 (1.1) 225	4.4 (0.9) 121	3.8 (0.9) 66	3.8 (1.0) 46	4.0 (1.2) 23	4.0 (1.4) 12
Serum HDL cholesterol (mmol/l)	1.2 (0.4) 219	1.3 (0.4) 119	1.4 (0.4) 65	1.3 (0.4) 45	1.1 (0.4) 23	1.3 (0.5) 11
Serum LDL cholesterol (mmol/l)	2.3 (0.8) 210	2.4 (0.8) 115	1.9 (0.6) 65	1.9 (0.7) 45	2.1 (0.9) 21	2.2 (0.9) 11
Lipid lowering medication (%)	24.7 190	20.3 73	60.5 104	55.7 54	38.0 35	22.2 10
Daily smokers (%)	12.7 88/691	11.0 36/327	11.9 19/160	10.1 9/89	12.7 10/79	12.2 5/41

Hb1Ac (%)	7.5 (1.1)	7.7 (1.7)				
	415	224				
Oral antidiabetic drug	12.3	13.1				
and insulin (%)	95	47				
Oral antidiabetic drug	34.0	29.8				
(%)	262	107				
Insulin (%)	16.8	16.7				
	129	60				
SCORE[1]	8.0 (6.3)	7.6 (7.9)				
	215	115				
NYHA[2]					2.4 (1.4)	2.5 (1.4)
					52	30

[1] Score index includes: Sex, Age, Total Chol, HDL and BP, smoking status
[2] New York Heart Association index on angina pectoris

Ethical approval and trial number

Written informed consent was obtained from all participants prior to enrollment into the project. The study protocol was approved by the Ethics Committee of the PHSSHD and registered (ClinicalTrials.gov Identifier: NCT00552903).

Results

In the intervention arm, 48.1% of the patients (156/324) initially above the target level of diastolic blood pressure of 85 mmHg reached this value, while for the control arm the proportion was 37.3% (62/166). The 10.8% (95% confidence interval (CI) 1.5–19.7%) difference in proportion of patients who reached the goal was statistically significant and gave a number needed to treat of 10 (CI 5–66). Of the patients with a systolic blood pressure above the target level of 140 mmHg at baseline, 35.9% (143/398) in the intervention arm and 31.0% (58/187) in the control arm reached the target ($p = 0.24$).

For waist circumference, the target was below 100 cm for men and 90 cm for women. The difference was not statistically significant ($p = 0.08$ combined, 0.07 for males and 0.65 for females) (Table 4). For patients with T2D, the goal for HbA_{1c} there was no difference between intervention and control group (Table 4).

Table 4 Proportion (%) of those patients reaching targets in primary end points among those exceeding these values at baseline in the analysed population (intervention = 816, controls = 405)

	Type 2 Diabetes		**Coronary artery disease**		**Congestive heart failure**	
	Intervention	Control	Intervention	Control	Intervention	Control
Hb1Ac (<7%)	30.2 % (n=65/215)	29.7% (n=27/91)				
Waist circumference (<90cm women, <100cm men)	9.8% (n=46/470)	5.1% (n=12/234)	11.1% (n=8/72)	9.5% (n=4/42)	10.8% (n=4/37)	15.8% (n=3/19)
Systolic blood pressure (<140mmHg)	32.7% (n=107/327)	35.8% (n=53/148)	47.1% (n=24/51)	16.1% (n=5/31)	60.0% (n=12/20)	0% (n=0/8)
Diastolic blood pressure (<85mmHg)	45.5% (n=120/264)	37.7% (n=49/130)	56.4% (n=22/39)	26.1% (n=6/23)	66.7% (n=14/21)	53.8% (n=7/13)
Serum total cholesterol (<4.5mmol/l)	30.7% (n=23/75)	35.0% (n=7/20)	77.8% (n=7/9)	100% (n=3/3)	100% (n=1/1)	0% (n=0/1)
Serum LDL cholesterol (<2.5mmol/l)	43.4% (n=29/67)	47.4% (n=9/19)	75.0% (n=6/8)	100% (n=2/2)	100% (n=1/1)	(n=0)
NYHA class (similar or improved)					83.9% (n=26/31)	93.3% (n=14/15)
Target reached in at least one primary endpoint*	47.8% (n=276/578)	44.8% (n=125/279)	49.5% (n=55/111)	43.8% (n=28/64)	75.5% (n=40/53)	65.4% (n=17/26)

*Intervention: 50.0% (371/742), Control: 46.1% (170/369), $p = 0.217$

The goal for total cholesterol reduction was reached more often in control arm than in intervention arm ($p = 0.64$) as was the LDL cholesterol target (≤2.5 mmol/l) ($p = 0.68$). For patients with CHF, NYHA class remained similar or improved in both arms ($p = 0.39$). The proportion of patients achieving at least one of the defined primary objectives was 50.0% (371/742) in the intervention and 46.1% in the control arm (170/369, $p = 0.22$). Within the intervention arm, no substantial differences were found between subjects assigned to different nurses (intracluster correlation 0.01).

Discussion

The TERVA trial was carried out in a real life setting and aimed at increasing the proportion of intervention patients reaching at least one of the predefined targets (blood pressure, HbA_{1c}, waist circumference, NYHA class or total cholesterol) by 10% compared to controls. There was a **small**, non-significant improvement in the proportion of patients who reached at least one of the primary endpoints for both the whole study population, and for each of the disease area subgroups separately. However, the difference reached the predefined 10% difference between the groups only for the CHF patients. An encouraging finding is the high adherence, nearly 90% of the patients remained in the trial during the intervention (similar to the control arm). Further analysis of the intervention arm will define how well patients could achieve the goals that they actually set at the beginning of the intervention.

Chronic disease management is a complex process urging multiple simultaneous changes in self-care, in health behavior, and in the interaction with medical care [3,21]. A complex intervention such as ours that targets these multiple behaviors cannot be compared to single-behavior interventions such as smoking cessation, medication adherence, or physical activity interventions. Despite these methodological complexities, little differences were found between subjects assigned to different nurses, indicating consistency in delivering the intervention. Further, health behavior changes may have a delayed impact or may impact the risk of cardiovascular diseases independently of clinical outcomes [19]. These reasons may partly explain that we did not meet our study objectives. Another possibility is that the intensity of the intervention was too low to sufficiently cover multiple behaviors, as recent evidence suggests that telephony interventions targeting only physical activity or/and diet produce most favorable effects when the number of calls is 12 or more [9]. Several previous studies have assessed the effect of telephony interventions on similar outcomes as ours [6,7,14,22]. Also these trials have shown modest improvement in clinical and health behavior outcomes.

This study aimed to evaluate an intervention within the public health care system and occupationally active patients were underrepresented, as they are mostly covered by occupational health services [23], and retired patients with more severe disease are overrepresented. The T2D patients in the trial (selected based on HbA_{1c} >7% within 6 months prior to inclusion) represented approximately one third of the T2D patients in the region [24,25]. Of them 28% had HbA_{1c} >7% at the start of the intervention, which is comparable to the population-based studies of T2D patients [24], suggesting that the participants are representative of the target population. Davidson concluded in his review the key success factor in diabetes care being specially trained nurses or pharmacists and perhaps one reason for modest results was that those in treatment were receiving already specialist nurse care [4] and added value of telephony was limited.

We included three different disease areas with variable disease severity. The mean HbA_{1c} was only 7.5% in intervention arm and 7.7% in control arm, with 28% and 25% with baseline HbA_{1c} >7 respectively, and disease history of 9.2 and 10.3 years. The large proportion of T2D patients with HbA_{1c} at the target level at enrollment was due to the fact that the patients were originally screened from primary care EMRs, and had frequently improved by the time of enrollment, which could be up to 6 months later. Also, the end of study HbA_{1c} measurement could potentially be up to 10 months after the intervention. The abstraction of the laboratory data from EMRs instead of a strict measurement protocol was motivated by the pragmatic nature of the trial, but in the low proportion of subjects with such data at the end of the study reduced the power (despite reaching the target sample size) and could introduce bias, as assessments were not prescribed randomly. This limitation renders the findings related to laboratory data difficult interpret meaningfully. Further, the targets for primary end-points, for instance waist circumference, which were based on systematic reviews of behavioral risk factor and disease management interventions, may have been too stringent [26]. Finally, the intervention was not coordinated with other health care providers, but rather added on top of the existing services. Some specialist diabetes nurses expressed a concern that health coaching was challenging their professional role, but no assessments were carried out to objectively measure health professionals' perceptions of the coaching program. Therefore, we can only speculate on the effect of the perceived competition on the results. However, it should be emphasized that the changes that were detected under these circumstances, demonstrate effects achieved in a real life setting.

Conclusions

The results of this trial are inconclusive, as we did meet the primary end-point for diastolic blood pressure only with non-significant improvement in systolic blood pressure and waist circumference and no improvement in glycemic control, cholesterol or NYHA class. The overall lack of efficacy of health coaching may be related to the target population, coaching procedures and the duration of the follow-up time, and will be further explored in longer follow-up and sub-group analyses, as well as analysis of behavioral outcomes. Methodological factors and too strict primary targets may contribute to inability to meet all the predetermined primary objectives. Further, the primary analysis focused on efficacy, and analysis on resource utilization and cost-efficacy need to be performed to fully clarify the role of health coaching by telephone in this setting.

Competing interest

P. Absetz trained health coaches (paid by Pfizer Oy). A. Auvinen's and K. Tokola's institution received grants from Joint Authority for Päijät-Häme Social and Health Care, Sitra, TEKES, and Pfizer Oy. J. Kytö was Medical Advisor for Pfizer Oy. Tom Wiklund was Medical Director of Pfizer Oy and had stock options of Pfizer Inc. Timo Ovaska was employee of Pfizer Oy. M. Talja's institution received support for travel to meetings and fees for participation in review activities, and outside the submitted work for board membership, consultancy, employment, expert testimony and grants.

Authors' contributions

MT, AA, RKe, KP, RKu and TO participated in the design of the study (including conception, aims and procedures). EO, PA and TO took part in implementation of intervention. TW, KP, KH, MT, MN, RKe, RKu, PA and AA had oversight of the intervention throughout the study and the analysis, and planned the analysis. KT carried out data management and statistical analysis. All authors contributed to presentation and interpretation of findings. KP, AA and JK wrote the first draft. AA, KH, KP, JK, MN, MT, RKu, TW and PA took part in revision of the text for important intellectual content. All authors have approved the manuscript.

Acknowledgments

We thank Academy Professor Hannu Oja, School of Health Sciences, University of Tampere for advice on statistical methods. Hannu.oja@uta.fi.

Funding

Joint Authority for Päijät-Häme Social and Health Care

Sitra - the Finnish Innovation Fund

TEKES - the Finnish Funding Agency for Technology and Innovation

Pfizer Oy

References

1. Geneau R, Stuckler D, Stachenko S, McKee M, Ebrahim S, Basu S, Chockalingham A, Mwatsama M, Jamal R, Alwan A, Beaglehole R: **Raising the priority of preventing chronic diseases: a political process.** *Lancet* 2010, **13:**1689–1698.

2. Cecchini M, Sassi F, Lauer JA, Lee YY, Guajardo-Barron V, Chisholm D: **Tackling of unhealthy diets, physical inactivity, and obesity: Health effects and cost-effectiveness.** *Lancet* 2010, **20:**1775–1784.

3. Daar AS, Singer PA, Persad DL, Pramming SK, Matthews DR, Beaglehole R, Bernstein A, Borysiewicz LK, Colagiuri S, Ganguly N, Glass RI, Finegood DT, Koplan J, Nabel EG, Sarna G, Sarrafzadegan N, Smith R, Yach D, Bell J: **Grand challenges in chronic non-communicable diseases.** *Nature* 2007, **22:**494–496.

4. Davidson MB: **How our current medical care system fails people with diabetes: Lack of timely, appropriate clinical decisions.** *Diabetes Care* 2009, **2:**370–372.

5. Choe HM, Mitrovich S, Dubay D, Hayward RA, Krein SL, Vijan S: **Proactive case management of high-risk patients with type 2 diabetes mellitus by a clinical pharmacist: a randomized controlled trial.** *Am J Manag Care* 2005, **11:**253–260.

6. Costa BM, Fitzgerald KJ, Jones KM, Dunning Am T: **Effectiveness of IT-based diabetes management interventions: a review of the literature.** *BMC Fam Pract* 2009, **10:**72.

7. Krishna S, Boren SA, Balas EA: **Healthcare via cell phones: a systematic review.** *Telemed J E Health* 2009, **15:**231–240.

8. Wagner EH, Groves T: **Care for chronic diseases.** *BMJ* 2002, **26:**913–914.

9. Goode AD, Reeves MM, Eakin EG: **Telephone-delivered interventions for physical activity and dietary behavior change: an updated systematic review.** *Am J Prev Med* 2012, **42:**81–88.

10. DeBusk RF, Miller NH, Superko HR, Dennis CA, Thomas RJ, Lew HT, Berger WE, Heller RS, Rompf J, Gee D, Kraemer HC, Bandura A, Ghandour G, Clark M, Shah RV, Fisher L, Taylor CB: **A case-management system for coronary risk factor modification after acute myocardial infarction.** *Ann Intern Med* 1994, **120:**721–729.

11. Vale MJ, Jelinek MV, Best JD, Dart AM, Grigg LE, Hare DL, Ho BP, Newman RW, McNeil JJ: **COACH Study Group. Coaching patients On Achieving Cardiovascular Health (COACH): a multicenter randomized trial in patients with coronary heart disease.** *Arch Intern Med* 2003, **163:**2775–2783.

12. McAlister FA, Stewart S, Ferrua S, McMurray JJ: **Multidisciplinary strategies for the management of heart failure patients at high risk for admission: a systematic review of randomized trials.** *J Am Coll Cardiol* 2004, **44:**810–819.

13. Taylor CB, Miller NH, Reilly KR, Greenwald G, Cunning D, Deeter A, Abascal L: **Evaluation of a nurse-care management system to improve outcomes in patients with complicated diabetes.** *Diabetes Care* 2003, **26:**1058–1063.

14. Bu D, Pan E, Walker J, Adler-Milstein J, Kendrick D, Hook JM, Cusack CM, Bates DW, Middleton B: **Benefits of information technology-enabled diabetes management.** *Diabetes Care* 2007, **30:**1137–1142.

15. Hayes E, McCahon C, Panahi MR, Hamre T, Pohlman K: **Alliance not compliance: Coaching strategies to improve type 2 diabetes outcomes.** *J Am Acad Nurse Pract* 2008, **20:**155–162.

16. Barrett MJ: **Patient self-management tools: An overview.** *Californian Healthc Found* 2005, http://www.chcf.org/~/media/MEDIA%20LIBRARY%20Files/PDF/P/PDF%20PatientSelfManagementToolsOverview.pdf [read 6.11.2011].

17. Olsen JM, Nesbitt BJ: **Health coaching to improve healthy lifestyle behaviors: An integrative review.** *Am J Health Promot* 2010, **25:**e1–e12.

18. Michie S, Abraham C, Whittington C, McAteer J, Gupta S: **Effective techniques in healthy eating and physical activity interventions: a meta-regression.** *Health Psychology* 2009, **28:**690–701.

19. Rollnick S, Miller WR, Butler CC: *Motivational Interviewing in Health Care: Helping Patients Change Behavior*.: The Guilford Press;New York 2008.

20. Tokola K, Larocque D, Nevalainen J, Oja H: **Power, sample size and sampling cost for clustered data.** *Stat Probab Lett* 2011, **81:**852–860.

21. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, *et al*: **Framework for design and evaluation of complex interventions to improve health.** *BMJ* 2000, **321:**694–696.

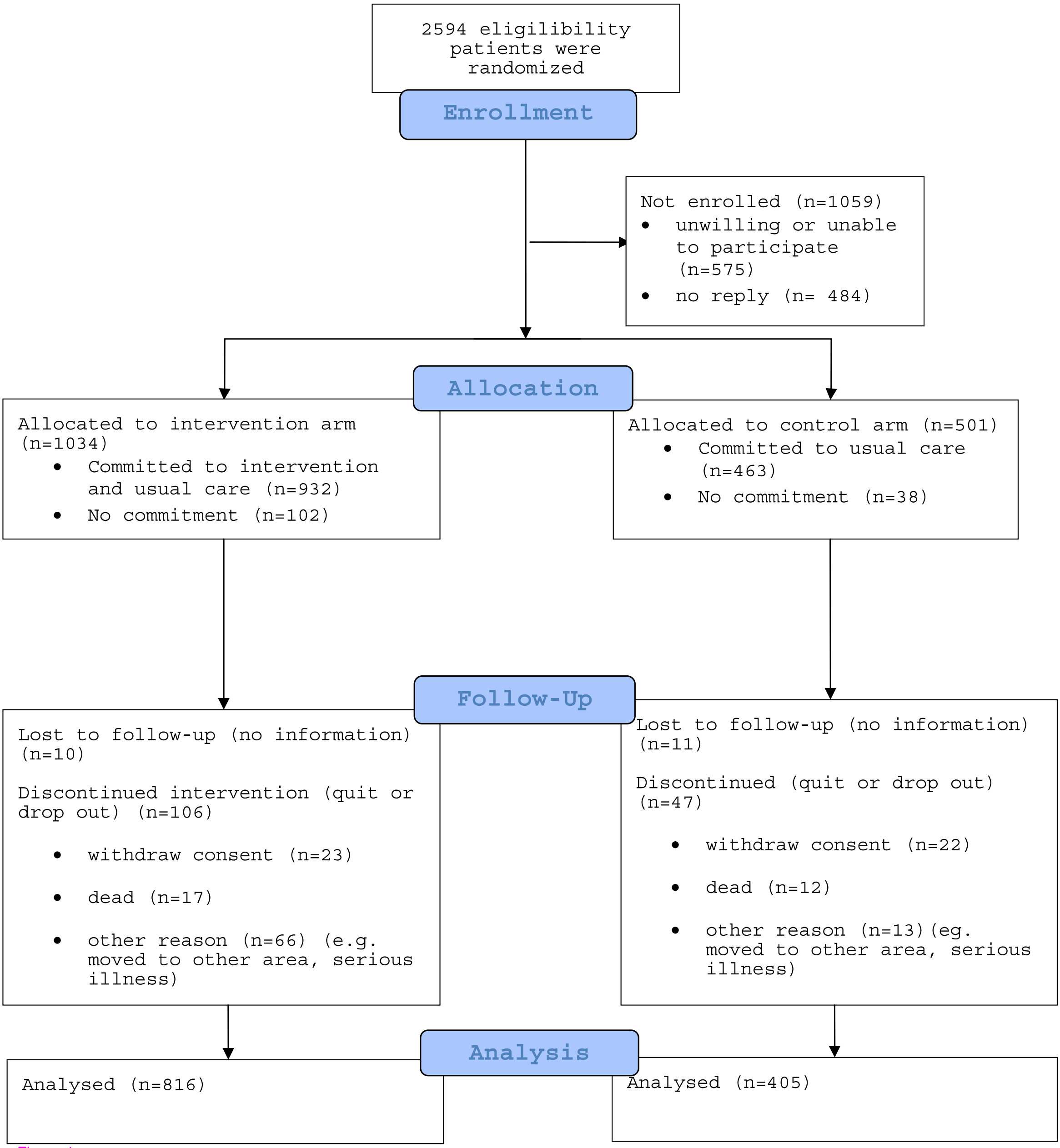
22. Kim SI, Kim HS: **Effectiveness of mobile and internet intervention in patients with obese type 2 diabetes.** *Int J Med Inf* 2008, **77:**399–404.

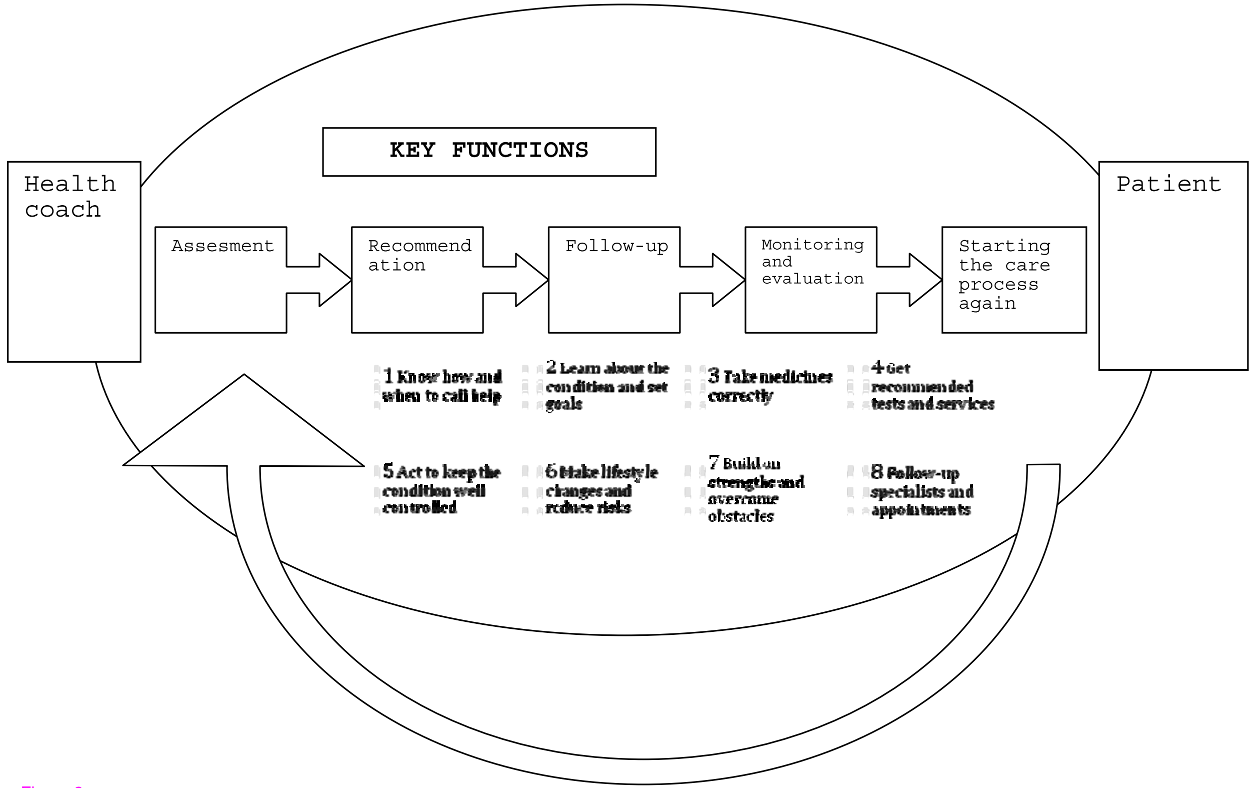
23. Häkkinen U, Alha P: *Health Service Utilization and its Socioeconomic Determinants. Health 2000 Survey*. Helsinki: National Public Health Institute; 2006.

24. Absetz P, Saaristo T, Viitasalo K, Peltonen M: **Implementation of prevention of type 2 diabetes – experiences from Finland**. In *Diabetes Prevention in Practice*. Edited by Schwarz P, Reddy P, Greaves C, Dunbar J, Schwarz J. Dresden: Tumaini Institute for Prevention Management; 2010.

25. Absetz P, Valve R, Oldenburg B, Heinonen H, Nissinen A, Fogelholm M, Ilvesmäki V, Talja M, Uutela A: **Type 2 diabetes prevention in the "real world": one-year results of the GOAL implementation trial.** *Diabetes Care* 2007, **30:**2465–70.

26. Oldenburg B, Absetz P, Chan CKY: **Behavioral interventions for prevention and management of chronic disease**. In *Handbook of Behavioral Medicine: Methods and Applications*. Edited by Steptoe A, Freedland KE, Jennings JR, Llabre MM, Manuck SB, Susman EJ.: Springer; Philadelphia 2010:969–988.


2594 eligibility patients were randomized
Enrollment
Not enrolled (n=1059)
• unwilling or unable to participate (n=575)
• no reply (n= 484)
Allocation
Allocated to intervention arm (n=1034)
• Committed to intervention and usual care (n=932)
• No commitment (n=102)
Allocated to control arm (n=501)
• Committed to usual care (n=463)
• No commitment (n=38)
Follow-Up
Lost to follow-up (no information) (n=10)
Discontinued intervention (quit or drop out) (n=106)
• withdraw consent (n=23)
• dead (n=17)
• other reason (n=66) (e.g. moved to other area, serious illness)
Lost to follow-up (no information) (n=11)
Discontinued (quit or drop out) (n=47)
• withdraw consent (n=22)
• dead (n=12)
• other reason (n=13)(eg. moved to other area, serious illness)
Analysis
Analysed (n=816)
Analysed (n=405)
Figure 1



KEY FUNCTIONS
Health coach
Assesment
Recommend ation
Follow-up
Monitoring and evaluation
Starting the care process again
Patient
1 Know how and when to call help
2 Learn about the condition and set goals
3 Take medicines correctly
4 Get recommended tests and services
5 Act to keep the condition well controlled
6 Make lifestyle changes and reduce risks
7 Build on strengths and overcome obstacles
8 Follow-up specialists and appointments

Figure 2

ARTICLE II

Oksman Erja, Linna Miika, Hörhammer Iiris, Lammintakanen Johanna & Talja Martti 2017. Cost-effectiveness analysis for a tele-based health coaching program for chronic disease in primary care. BMC Health Services Research 17(1), 138. https://doi.org/10.1186/s12913-017-2088-4

Oksman et al. BMC Health Services Research (2017) 17:138
DOI 10.1186/s12913-017-2088-4

BMC Health Services Research

RESEARCH ARTICLE Open Access

Cost-effectiveness analysis for a tele-based health coaching program for chronic disease in primary care

Erja Oksman[1*], Miika Linna[2], Iiris Hörhammer[2], Johanna Lammintakanen[3] and Martti Talja[1]

Abstract

Background: The burden of chronic disease and multimorbidity is rapidly increasing. Self-management support interventions are effective in reduce cost, especially when targeted at a single disease group; however, economical evidence of such complex interventions remains scarce. The objective of this study was to evaluate a cost-effectiveness analysis of a tele-based health-coaching intervention among patients with type 2 diabetes (T2D), coronary artery disease (CAD) and congestive heart failure (CHF).

Methods: A total of 1570 patients were blindly randomized to intervention (*n* = 970) and control (*n* = 470) groups. The intervention group received monthly individual health coaching by telephone from a specially trained nurse for 12-months in addition to routine social and healthcare. Patients in the control group received routine social and health care. Quality of life was assessed at the beginning of the intervention and follow-up measurements were made after 12 months health coaching. The cost included all direct health-care costs supplemented with home care and nursing home-care costs in social care. Utility was based on a Health Related Quality of Life (HRQoL) measurement (15D instrument), and cost effectiveness was assessed using incremental cost-effectiveness ratios (ICERs).

Results: The cost-effectiveness of health coaching was highest in the T2D group (ICER €20,000 per Quality-Adjusted Life Years [QALY]). The ICER for the CAD group was more modest (€40,278 per QALY), and in the CHF group, costs increased with no marked effect on QoL. Probabilistic sensitivity analysis indicated that at the societal willingness to pay threshold of €50,000 per QALY, the probability of health coaching being cost effective was 55% in the whole study group.

Conclusions: The cost effectiveness of health coaching may vary substantially across patient groups, and thus interventions should be targeted at selected subgroups of chronically ill. Based on the results of this study, health coaching improved the QoL of T2D and CAD patients with moderate costs. However, the results are grounded on a short follow-up period, and more evidence is needed to evaluate the long-term outcomes of health-coaching programs.

Trial registration: NCT00552903 [Prospectively registered, registration date 1st November 2007, last updated 3rd February 2009].

Keywords: Health coaching, Self-management, Cost-effectiveness, Chronic disease, 15D, Health-related quality of life

* Correspondence: erja.oksman@phsotey.fi
[1]Päijät-Häme Social and Health Care District, Keskussairaalankatu 7, 15 850 Lahti, Finland
Full list of author information is available at the end of the article

Background

In the European Union (EU), approximately 50 million people live with multiple chronic diseases, and this is one of the leading causes of growing healthcare costs. It is estimated that chronic diseases inflict 70–80% of total healthcare costs in EU countries [1]. In Finland (pop. 5.4 million in 2011), 37.2% of the population had at least one chronic disease or health problem in 2011 [2]. Therefore, how to manage the burden of chronic disease is a key question for policy makers.

Self-management support interventions are widely recognized as a promising approach to enhance health outcomes and contain costs in chronic care. Previous studies suggest that self-management interventions improve clinical outcomes, self-efficacy, quality of life and self-management behaviour [3–5]. They have also been successful in reducing hospitalization and healthcare costs, especially when the intervention has been focused on a single disease. The most promising results have been observed in respiratory and cardiovascular diseases [6].

However, the economic evaluation (cost effectiveness and cost-utility analysis) of self-management interventions is still scarce, although cost-effectiveness analysis has become a standard practice in evaluating, e.g. medical treatments [7–10]. This may be due to methodological challenges; self-management interventions are often complex interventions, and the standard experimental setting, a randomized controlled trial (RCT), is difficult to put into practice in real life [11, 12]. Furthermore, routinely collected administrative and clinical data typically lack the important measurements needed in the assessment of self-management interventions [13].

Health coaching is patient-oriented health promotion and education within a coaching context that emerged from the motivational interviewing concept [3]. The purpose of health coaching, as defined by Palmer et al. [14] is to motivate the patient to achieve a goal that enhances quality of life and improves health. A coach's role is to help patients weigh options, make choices and plan and identify challenges to help them change for the better [14].

Telephone-based health-coaching intervention was launched in November 2007 in the Päijät-Häme area in Finland. The number of inhabitants above the age of 65 years was increasing faster than in other parts in Finland, and costs of delivering secondary care were high, especially for chronic conditions, such as heart failure, coronary heart disease and diabetes. The health-coaching call centre was established in the city of Lahti as a public–private partnership, where the public partner was responsible for the primary care and secondary care in the region.

The objective of this study was to evaluate the cost effectiveness of 12 months of telephone-based health-coaching intervention (the TERVA trial) for chronically ill patients in Finland. This was tested using a two arm trial with three patient groups with sub optimally controlled disease: type 2 diabetes mellitus (T2D), coronary artery disease (CAD) or congestive heart failure (CHF). The primary outcomes of the TERVA trial, the short-term clinical outcomes at 12 months follow-up, have been reported earlier [15].

Methods

The total population of the area involved the health coaching program was approximately 112 000. Patients were recruited from electronic patient laboratory records in secondary care according to laboratory inclusion criteria (Glycated Hemoglobin (HbA1c) >7 or total cholesterol >4,5 or low density lipoprotein (LDL) >2.3 previous six months). In this phase we identified about 5 500 patients. After that research nurse identified those patients who were applicable for coaching according to inclusion and exclusion criteria from patients' medical records, 2594 patient fulfilled inclusion criteria and were invited to participate. The information and consent letters were sent to the patients. 1535 identified patients, gave consent and were randomized to either control (C) or intervention (I) groups. At the baseline, there were no significant differences in age, sex, self- reported duration of disease and age of diagnosed, blood pressure (systolic, diastolic), total cholesterol, high density lipoprotein (HDL), LDL, body mass index (BMI), waist circumference, daily smokers, lipid lowering medication, HbA1c, oral antidiabetic drug and insulin, oral antidiabetic drug, insulin and NYHA-class between intervention and control groups [15]. Randomization 2:1 ratio was intentional for practical reasons. Statistical power calculations were conducted to verify that the imbalance would not cause problems. The intervention group received monthly individual health coaching for 12 months in addition to routine social and healthcare. Patients with multiple morbidities received coaching for each disease according to their personal priorities. Patients in the control group received routine social and health care e.g. visited diabetes nurse and doctors in primary and secondary care. Patients with more than one disease were allocated to following hierarchy: CHF-CAD- T2D [15]. Of these 1535 participants, 998 patients with complete baseline and follow-up data were included in the cost-effectiveness analysis (83 patients in the CHF group (I 56, C 27), 192 in the CAD group (I 124, C 68) and 723 in the T2D group (I 505, C 218). A total of 537 patients were lost in the follow-up. The detailed recruitment and randomization process has been published previously [15].

Intervention

Eight experienced certified nurses and public health nurses were hired and trained in the motivational interviewing technique and coaching by telephone. Health coaches had access and the possibility to document patient health status into the primary and secondary care electronic health records (EHR), but they were not integrated in the care teams in the primary care centres. A more complete description of the health-coaching intervention process can be found in [15].

The health-coaching intervention included eight key recommendations developed by Pfizer Health Solution (PHS) and were adjusted for the Finnish healthcare system and Finnish evidence-based guidelines. The eight recommendations included: 1) know how and when to call for help, 2) learn about the condition and set goals, 3) take medicines correctly, 4) get recommended tests and services, 5) act to keep the condition well, 6) make lifestyle changes and reduce risk, 7) build on strengths and overcome obstacles and 8) follow up with specialists and appointments. Coaching was technology supported and utilized a traffic-light system for patients' progress in relation to the key recommendations. Patient's self-management booklets supported progress towards the key recommendations. Each disease had a separate booklet prepared in collaboration with the Finnish Heart Association and the Diabetes Association. The patients in the intervention group were called approximately 10–12 times during the intervention period.

Data

Health-related quality of life

HRQoL was measured by using 15D [16].15D is a generic, self-administered instrument for measuring HRQoL among adults (age over 16 years) with 15 dimensions: mobility, vision, hearing, breathing, sleeping, eating, speech, excretion, usual activities, mental function, discomfort and symptoms, depression, distress, vitality and sexual activity. Completing the questionnaire takes 5–10 min. Each dimension has five ordinal levels, and 15D can be used as a profile measure or a single index number on a scale of 0–1 (0 dead, 1 completely healthy). Typically, 15D is used to measure the effectiveness of a single intervention [16, 17] and performs well in comparison to SF-36 as a HRQoL-instrument [18].

The baseline HRQoL data were collected by sending the 15D questionnaire to patients in the intervention and control groups at the beginning of the health-coaching intervention and follow-up measurements were made when the coaching finished after 12 months.

Cost data

Data for the costs and use of social and healthcare services were collected from the National registries maintained by the National Institute for Health and Welfare (Dnro THL/119/5.05.00/2015). These registers included the hospital benchmarking database the National Discharge Registry (HILMO) and Care Registers for Social Welfare (SosiaaliHILMO). Using a unique patient identification code, patient cohorts were linked to the registers, and all use of social and healthcare during 1-year follow-up was included for each individual. Secondary care data included the use of hospital outpatient care (all types of visits) and hospital admissions (diagnosis-related groups [DRGs]). Social care data included all types of long- and short-term institutionalized care, housing and residential services and home care services.

Hospitalizations and hospital outpatient visits due to any cause were extracted from the Hospital Discharge Register based on the International Classification of Diseases 10th revision (ICD-10) codes, the Finnish version of the Nordic Classification of Surgical Procedures (NCSP) codes for diagnostic and treatment procedures and the respective NordDRG patient grouping classifications. The DRG cost weights for hospitalizations and outpatient visits were based on individual-level cost-accounting data from several hospitals. The unit cost estimates for social care encounters and bed days were derived from the national price list for unit costs of healthcare services in Finland [19].

The use of primary healthcare resources was collected directly from the patient administration system (PAS) containing patient-level data abstracts from the electronic patient records. The PAS data included contact types (such as a visit, phone call or electronic messaging), patient's age, the diagnosis (ICD-10) or the reason for encounter (ICPC-2) and the employee category of the healthcare professional in the contact. Extracting the patient-level data from the patient administration systems (with diagnosis and activity information) made it possible to group each individual encounter type by the Ambulatory and Primary Care Related Patient Groups (APR) grouper software, a grouping system equivalent to DRGs used in hospital care [20]. The batch grouper software assigned each individual patient encounter in one of the 44 APR groups. After grouping, each of the 44 APR groups in the sample was assigned a cost weight indicating the relative consumption of resources. Cost weights were based on large samples of time measurements in primary care contacts and procedures. All costs were deflated using the price index for public healthcare provided by Statistics Finland.

Statistical analysis

We report differences in the mean costs and outcomes and the corresponding cost-effectiveness ratio (ICER). ICER is defined by the difference in cost between the intervention and control, divided by the difference in their effect.

Uncertainty in the ICER estimates was accounted for by generating bootstrap 1000 replicates of the dataset, a method widely used in health economic evaluations [21, 22] to study the likelihood of effectiveness of an intervention in relation to the costs of care induced by the intervention [23]. Probabilistic sensitivity analysis was completed by calculating the cost-effectiveness acceptability curve (CEAC) derived from the bootstrap replicates. CEAC indicates the probability for cost effectiveness of the intervention at different levels of willingness to pay for the additional health outcome [24].

Results

The overall incremental ICER was €48,000 per QALY. The cost effectiveness of health coaching was highest in the T2D group (ICER €20,000 per QALY). The ICER for the CAD group was more modest (€40,278 per QALY), although the improvement in QoL was greatest in this group and also exceeded the threshold for a clinically significant change in 15D (>0.015 [25]). In the CHF group, the effect on QoL was slightly negative at an increased cost (Table 1).th=tlb=

Figure 1 presents the bootstrapped results among the whole study group displayed in a cost-effectiveness plane. There was considerable uncertainty in the ICER of the intervention.

The cost-effectiveness plane for HRQoL (15D) after health coaching showed that the intervention was more effective compared to care as usual but also more costly. Of the bootstrapped ICERs, 89% fell into the northeast quadrant, indicating increased QoL at an incremental cost; 9% of the points fell into the southeast quadrant, indicating increased QoL at a decreased cost. Only 2% of the data points fell into the northwest quadrant, and less than 1% fell into the southwest quadrant, suggesting a very small probability for a decrease in QoL at an incremental or decreased cost (Fig. 1).

Figure 2 shows the incremental CEACs for the whole participant group and for the disease-specific subgroups. At no willingness to pay for incremental QALY, the probability of health-coaching cost effectiveness was less than 10% among all participants. At a willingness to pay €46,000 per QALY, the probability that the intervention is cost effective was over 50%. If the decision maker were willing to pay €50,000 per QALY, the probability of cost-effectiveness is 55%. The CEAC for the T2D group showed over 50% probability of cost effectiveness at a willingness to pay €20,000 per QALY. At a willingness to pay €50,000 per QALY, the probability that the intervention is cost effective for the T2D patients was 75%.

Discussion

In this study, the cost effectiveness of 12 months telephone-based health-coaching intervention among three groups of chronically ill patients with unmet treatment goals was evaluated. The overall ICER was €48,000 per QALY. Further probabilistic sensitivity analysis showed a 55% probability of cost effectiveness if the decision maker were willing to pay €50,000 per QALY. Investments in programs for coaching patients may well be acceptable. Further disease-specific analyses indicated that the ICER for health coaching was lowest in the T2D group with a moderately low cost per QALY of €20,000. In the CAD group, the cost per QALY was higher (€40,278), and in the CHF group the effect on QoL was slightly negative at an increased cost.

Graves et al. [9] reported similar results ($29,375 per QALY, approximately €21,045 per QALY) for patients with T2D or hypertension after 1-year telephone-delivered intervention for physical activity and diet in a low socioeconomic area in Australia. Jacobs-van-der Bruggen et al. [26] analysed seven lifestyle interventions among patients with T2D and simulated the long-term outcomes. Health improvements were achieved at reasonable costs (≤ €50,000 per QALY), and average gained health-adjusted life years were 0.01–0.14 QALY per participant. Results in the CHF group somewhat contradicted the results of previous studies [5, 6]. However, the small number of patients (I 56, C 27) may have diluted the evidence in this subgroup or the coaching program did not support those people.

In this study, cost per QALY was found to be lowest in the T2D group. An improvement in QALY (0.008) was achieved with a small increase in the cost of care (€160 per patient). In the CAD group, both the improvements in QoL (0.018) as well as the increase in cost (€725 per patient) were higher. Possible explanations for the difference between these groups can be found in the medical history and the care received by the patients prior to the intervention. Most CAD patients were recruited for health coaching a few months after an acute percutaneous transluminal coronary angioplasty

Table 1 Incremental costs, quality of life and cost-effectiveness ratios in the disease subgroups and in the whole study group

	Costs (€), mean (95% CI)			QoL (15D), mean (95% CI)			ICER (€/QALY)
	Intervention	Control	Incremental cost	Intervention	Control	Incremental effect	
Type 2 diabetes	1948 (1673–2222)	1788 (1204–2371)	160 (−406–726)	0.008 (0.003– 0.014)	0.000 (−0.009–0.009)	0.008 (−0.002–0.018)	20 000
Coronary artery disease	2510 (1806–3214)	1785 (984–2585)	725 (−389–1839)	0.019 (0.007–0.030)	0.001 (−0.014–0.016)	0.018 (−0.001–0.037)	40 278
Congestive heart failure	4469 (1955–6983)	2214 (−105–4533)	2255 (−1669–6180)	0.013 (−0.007–0.032)	0.015 (−0.015–0.046)	−0.003 (−0.037–0.032	-
All	2256 (1940–2571)	1824 (1345–2302)	432 (−135–999)	0.011 (0.006 – 0.015)	0.002 (−0.006–0.009)	0.009 (0.000–0.018)	48 000

CI confidence interval, *QoL* Quality of Life, *ICER* incremental cost-effectiveness ratio, *QALY* quality- adjusted life years

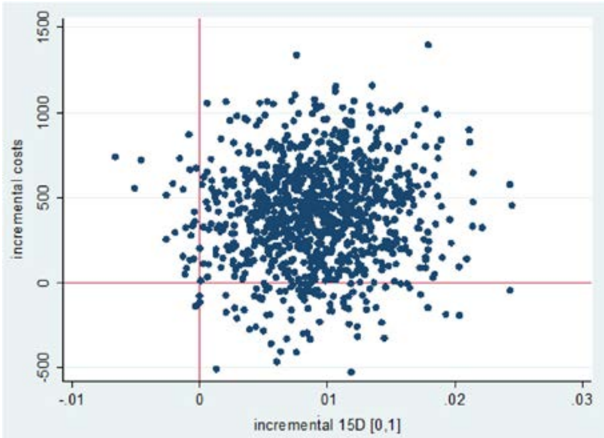


Fig. 1 Distribution of bootstrapped incremental costs and health-related quality of life

operation. Motivation to lifestyle changes and self-management are high after an acute cardiovascular attack [27]. The proximity of this severe incident may have activated the CAD patients in their self-care and healthcare service use and therefore fortified the effect of the intervention on the QoL increased cost in this group. Another reason for the increased cost in this group can be attributed to the standard follow-up visits after an acute cardiovascular attack or intervention in secondary care. Further, in care as usual, the diabetes patients receive treatment and self-care support from specially trained diabetes nurses, while the self-care support for CAD patients is not arranged as systematically in the present healthcare provision. This may explain the difference in the increased cost of care between the groups.

This study is among the few cost-effectiveness evaluations of health coaching for the chronically ill carried out in a real-life setting and using RCT design. Another strength of the present study was the use of national registries and local patient administration systems, including all social and healthcare services and their costs in the follow-up. Many studies published so far have relied on the self-reported use of services.

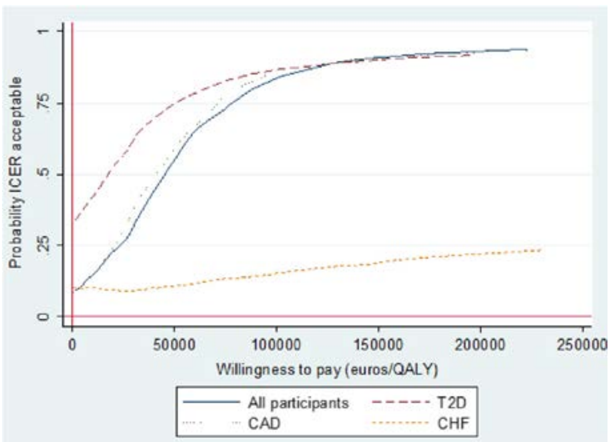


Fig. 2 Cost-effectiveness acceptability curves for all participants and diagnosis-based subgroups

One clear shortcoming in the study was the rather short follow-up period. Significant health behaviour changes take at least 6 months and may have delayed the impact in clinical changes [28, 29]. A new, long term follow-up study, using the cohorts in the present study and based on National registries, has been set up to clarify the effects by analysing the differences in distal end points (such as complications in T2D and major events in CAD) and the differences in cumulated health and social care costs.

Immediately after the TERVA trial, only the clinical results and direct cost data were available for the regional decision makers, and the health-coaching program was cancelled. This may be a common problem with the evaluation of self-management and other preventive interventions, which typically focus on short-term health outcomes [30]. In this study, a closer exploration using QALYs and subgroup analysis revealed that closing the coaching program may have been questioned on the basis of the cost-effectiveness analysis.

We conclude that the assessment of cost effectiveness in preventive actions is demanding and thus requires careful and balanced analyses to sufficiently inform the decision makers on preferred choices.

Conclusions

Decision makers in health care are actively seeking interventions leading to better health outcomes with a lower cost, but the evidence on cost effectiveness of self-management interventions is still scarce. In this RCT conducted in a real-life primary care setting, health coaching improved the QoL of T2D and CAD patients with moderate costs in the short-term follow-up. The results of our study suggest that health coaching should be targeted to selected patient groups. However, the follow-up period was probably too short to evaluate the cost-effectiveness of health-coaching intervention and a long-term evaluation is needed.

Abbreviations

APR: Ambulatory and primary care related patient group; BMI: Body mass index; CAD: Coronary artery disease; CEAC: Cost-effectiveness acceptability curve; CHF: Congestive heart failure; CI: Confidence interval; DRG: Diagnosis-related groups; EU: European Union; EVO: Special state funding; HbA1c: Glycated hemoglobin; HDL: High density lipoprotein; HILMO: National discharge registry; HRQoL: Health related quality of life; ICD-10: Classification of disease 10th revision; ICER: Incremental cost-effectiveness; ICPC-2: Reason for encounter patient administration system; LDL: Low density lipoprotein; NCSP: Nordic classification of Surgical Procedures; NYHA: New York heart association index on angina pectoris; PAS: Patient administration system; QALY: Quality adjusted life years; QoL: Quality of life; RCT: Randomized controlled trial; Sos. HILMO: Care register for social welfare; T2D: Type 2 diabetes; TERVA: Telephone-based health coaching intervention

Acknowledgements

Not applicable.

Funding

Corresponding author (EO) received the special state funding (EVO) grant from Päijät-Häme central hospital. No additional funding was obtained.

Availability of data and materials

Access to these databases is closed and permission to use the data was received from the national authorities.

Authors' contribution

EO is primary author and responsible for theoretical parts of manuscript. ML and IH were responsible for data management and statistical analysis and revised the manuscript. JL and MT revised the manuscript and MT was head of the TERVA –primary study. All authors took part the interpretation of findings and have approved the revised manuscript.

Competing interest

The authors declare that they have no competing interest.

Consent for publication

Not applicable. No individual participant data is present.

Ethical approval and consent to participate

The study was approved by the Ethics Committee of the Päijät-Häme Social and Health Care District [Dnro ETMK 56/2008, Dnro 65/2008]. Written informed consent was obtained from all participants prior to enrollment into the TERVA –project.

Author details

[1]Päijät-Häme Social and Health Care District, Keskussairaalankatu 7, 15 850 Lahti, Finland. [2]Healthcare Engineering, Management and Architecture Institute, Aalto University, Espoo, Finland. [3]Department of Social and Health Management, University of Eastern Finland, Joensuu, Finland.

Received: 23 June 2016 Accepted: 10 February 2017
Published online: 15 February 2017

References

1. Rieken M, Struckmann V, Melchieorre MG, on the behalf of the ICARE4EU partners, et al. ICARE 4EU: Improving care for people with multiple chronic condition in Europe. Eurohealth. 2013;19(3):29–31.
2. Eurostat 2011. Available at: http://ec.europa.eu/eurostat
3. Kivelä K, Elo S, Kyngäs H, et al. The effects of health coaching on adult patient with chronic diseases: A systematic review. Patient Educ Couns. 2014;97:147–57.
4. Trappenburg J, Jonkman N, Jaarsma T, et al. Self-management: One size does not fit all. Patient Educ Couns. 2013;92(1):134–7.
5. Ditewig J, Blok H, Havers J, et al. Effectiveness of self-management on mortality, hospital readmissions, chronic heart failure hospitalization rate and quality of life in patient with chronic heart failure: A systematic review. Patient Educ Couns. 2010;78:297–315.
6. Panagioti M, Richardson G, Small N, et al. Self-management support interventions to reduce health care utilization without compromising outcomes: A systematic review and meta-analysis. BMC Health Serv Res. 2014;14:356. Available from: bmchealthservres.biomedcentral.com/articles/10.11.86/1472-6963-14-356 [Accessed 13 Aug 2015].
7. Riippa I. Building evidence for cost-effectiveness interventions in chronic care. Acknowledging context and mechanisms. Aalto University School of Science. Department of Industrial Engineering and Management. Hema Institute. Aalto University publication series. Doctoral Dissertation 142/2015.
8. Basu R, Ory M, Towne S, et al. Cost-effectiveness of the chronic disease self-management program: Implications for community- based organizations. Front Public Health. 2015;3:27. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4410335/.
9. Graves N, Barnett A, Halton K, et al. Cost-effectiveness of telephone-delivered intervention for physical activity and diet. PLoS ONE. 2009;4(9):e7135.
10. Räsänen P, Roine E, Sintonen H, et al. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. Finohta's report. 2006;29:1–37.
11. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: New guidance. 2006. Medical Research Council. Available from: https://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/.

12. Weatherly H, Drummond M, Claxton K, et al. Methods of assessing of cost-effectiveness of public health interventions: key challenges and recommendations. Health Policy. 2009;93(2):85–92.
13. Riippa I, Kahilakoski O-P, Linna M, et al. Can complex health interventions be evaluated using routine clinical and administrative data? – a realistic evaluation approach. J Eval Clin Pract. 2014;20(6):1129–36.
14. Palmer S, Tubbs I, Whybrow A. Health coaching to facilitate to promotion of health behaviour and achievement of health-related goals. Int J Health Promot Educ. 2003;41:91–3.
15. Patja K, Absetz P, Auvinen A, et al. Health Coaching by telephony to support self-care in chronic diseases: clinical outcomes from The TERVA randomized controlled trial. BMC Health Serv Res. 2012;12:147.
16. Sintonen H. The 15D measure of health-related quality of life. Properties and applications. Ann Med. 2001;33:328–36.
17. Sintonen H. 15D Instrument. Available from: http://www.15D-instrument.net/15D
18. Hawthorne G, Richardson J, Atherton DN. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Ann Med. 2001;33(5):358–70.
19. Kapiainen S, Väisänen A, Haula T. Terveyden- ja sosiaalihuollon yksikkökustannukset Suomessa vuonna 2011, THL (National Institute for Health and Welfare) Raportti 3/2014. Available at: https://www.julkari.fi/bitstream/handle/10024/114683/THL_RAPO03_2014_web.pdf?
20. Honkasalo M, Linna M, Sane T, et al. A comparative study of two various models of organising diabetes follow-up in public primary health care – the model influences the use of services, their quality and costs. BMC Health Serv Res. 2014;12:26.
21. Bos IBCK, Hoving JL, van Tulder MW, et al. Cost effectiveness of physiotherapy, manual therapy, and general practitioner care for neck pain: economic evaluation alongside a randomised controlled trial. BMJ. 2003; 326(7395):911. PMID: 12714472.
22. Van Spijker BAJ, Majo MC, Smit F, et al. Reducing suicidal ideation: cost-effectiveness analysis of a randomized controlled trial of unguided web-based self-help. J Med Internet Res. 2012;14(5), e141. PMID: 23103835.
23. Briggs A, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health Econ. 1998;7(8):723–40. PMID: 9890333.
24. Van Hout BA, Al MJ, Gordon GS, et al. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ. 1994;3(5):309–19. PMID: 7827647.
25. Alanne S, Roine R, Räsänen P, Vainiola T, Sintonen H. Estimating the minimum important change in the 15D scores. Qual Life Res. 2015;24(3): 599–606. http://doi.org/10.1007/s11136-014-0787-4).
26. Jacobs-van der Bruggen M, van Baal P, Hoogenveen R, et al. Cost-effectiveness of lifestyle modification in diabetes patient. Diabetes Care August. 2009;32(8):1453–8.
27. Evans L, Probert H, Shuldham C. Cardiac rehabilitation-past to present. J Res Nurs. 2009;4(3):223–40.
28. Rollnick S, Miller WR, Butler CC. Motivational inteviewing in health care: Helping patient change behavior. New York: The Guilford Press; 2008.
29. Prochaska JO, Butterworth S, Redding CA, et al. Initial efficacy of MI, TTM tailoring and HRI's for multiple behaviors for employee health promotion. Prev Med. 2008;46:226–31.
30. Rabarison KM, Bish CL, Massoudi ML. Economic evaluation echances public health decision making. Front Public Health. 2015;3:164. doi:10.3389/2015.00164.

ARTICLE III

Mustonen Erja, Hörhammer Iiris, Absetz Pilvikki, Patja Kristiina, Lammintakanen Johanna, Talja Martti, Kuronen Risto & Linna Miika 2020. Eight-year post-trial follow-up of health care and long-term care costs of tele-based health coaching. Health Services Research 55 (2), 211–217. https://doi.org/10.1111/1475-6773.13251

DOI: 10.1111/1475-6773.13251

RESEARCH ARTICLE

HSR Health Services Research

Eight-year post-trial follow-up of health care and long-term care costs of tele-based health coaching

Erja Mustonen MSc[1] | Iiris Hörhammer PhD[2] | Pilvikki Absetz PhD[3,4] | Kristiina Patja MD, PhD[5] | Johanna Lammintakanen PhD[3] | Martti Talja MD, PhD[1] | Risto Kuronen MD, PhD[1] | Miika Linna D.Sc. (Tech)[2]

[1]Päijät-Häme Joint Authority for Health and Wellbeing, Lahti, Finland

[2]Healthcare Engineering, Management and Architecture Institute, Aalto University, Aalto, Finland

[3]University of Eastern Finland, Kuopio, Finland

[4]CEO, Collaborative Care Systems Finland, Helsinki, Finland

[5]Department of Public Health, University of Helsinki, Helsinki, Finland

Correspondence
Erja Mustonen, MSc, Päijät-Häme Joint Authority for Health and Wellbeing, Keskussairaalankatu 7, 15 850 Lahti, Finland.
Email: erja.mustonen1@outlook.com

Funding information
Päijät-Häme Joint Authority for Health and Wellbeing; Pfizer; Suomen Akatemia; Finnish Innovation Fund (Sitra)

Abstract

Objective: To evaluate the long-term effect of telephone health coaching on health care and long-term care (LTC) costs in type 2 diabetes (T2D) and coronary artery disease (CAD) patients.

Data Sources/Study Setting: Randomized controlled trial (RCT) data were linked to Finnish national health and social care registries and electronic health records (EHR). Post-trial eight-year economic evaluation was conducted.

Study Design: A total of 1,535 patients (≥45 years) were randomized to the intervention (n = 1034) and control groups (n = 501). The intervention group received monthly telephone health coaching for 12 months. Usual health care and LTC were provided for both groups.

Principal Findings: Intention-to-treat analysis showed no significant change in total health and long-term care costs (intervention effect €1248 [3 percent relative reduction], CI −6347 to 2217) in the intervention compared to the control group. There were also no significant changes among subgroups of patients with T2D or CAD.

Conclusions: Health coaching had a nonsignificant effect on health care and long-term care costs in the 8-year follow-up among patients with T2D or CAD. More research is needed to study, which patient groups, at which state of the disease trajectory of T2D and cardiovascular disease, would best benefit from health coaching.

KEYWORDS

coronary artery disease, costs, effectiveness, health coaching, type 2 diabetes

1 | INTRODUCTION

Chronic diseases contribute to 70-80 percent of health care costs. Specifically, the care of multimorbidity fails to meet the patients' complex needs, leading to insufficient care. Often, this leads to acute and unplanned use of health care services, especially in emergency units, and increasing hospitalization in secondary care.

The usual goal of chronic disease management programs is to improve patients' self-management skills in increasing treatment adherence, such as keeping appointments with health care professionals and taking prescribed medicines. Improved compliance reduces emergency visits and prevents expensive hospitalization.[1,2] Conventional disease management programs focus on the disease itself, emphasizing coordinated and comprehensive care pathways following

Trial registration: NCT00552903 (prospectively registered, registration date 1 November 2007, last updated 3 February 2009).

evidence-based clinical guidelines and encouraging patient compliance to treatments, but they focus less on the patient's individualized needs or behavior.[2,3]

Health coaching, a patient-centered approach aiming to empower patients in comanagement of their disease,[4] emphasizes and supports patient autonomy and learning instead of compliance. It is based on shared decision making and collaborative goal setting facilitated by motivational interviewing.[5-8] Based on Hale's [8] integrative review, health coaching is described throughout the literature as a partnership between the coach and the individual. More specifically, it is "a goal-oriented, client-centered partnership that is health-focused and occurs through a process of client enlightenment and empowerment".[6] Health coaching is usually provided by certified health coaches or health care professionals.[6] The role of the coach involves listening, understanding, facilitating, applauding, supporting, motivating, providing feedback, and helping the patient to weigh options, make choices, and identify and overcome challenges in the process of change for better.[4] Health coaching guides a learning process for improved disease management; therefore, if successful, it should lead to permanent changes in patient self-management skills and behavior. These changes in self-management skills and behavior take time to have an effect on health outcomes,[10,26] and therefore, the impact of health coaching on health care effectiveness and cost-effectiveness should be assessed in long-term follow-ups.

Evidence on the effectiveness of health coaching is conflicting, and it is based on studies with short-term follow-up only (up to 24 months).[8-11] Due to heterogeneity of target populations and outcome measures, no systematic reviews with meta-analyses have been completed.[12] Individual studies show either small significant effects or no effects.[10] Furthermore, evidence on the cost-effectiveness of health coaching remains limited: Utilization and cost of health care services has only been evaluated in the short term (usually 12 months), again with mixed outcomes.[7,13-18] However, due to the nature of the underlying mechanism of change—learning rather than compliance—it might take longer to evidence effects. Therefore, long-term evaluations of the effectiveness and cost-effectiveness of health coaching interventions are needed.

The TERVA trial (trial registration: NCT00552903) is a health coaching program that was implemented in the Päijät-Häme region in Southern Finland and tested as a randomized controlled trial in 2007-2009. Patients with suboptimally controlled T2D or CAD, including a subgroup of patients with congestive heart failure (CHF), were coached via telephone by trained health coaches during a one-year intervention period. The aim of the study was to evaluate the total health care and long-term care (LTC) costs among all participants and in the subgroups (T2D and CAD) for an 8-year follow-up of the TERVA trial.

What This Study Adds

1. Previous studies have shown mixed results on the cost-efficiency of telephone-based health coaching with rather short follow-ups
2. In an 8-year follow-up of all health care and long-term care costs, this study found no definitive evidence for cost-efficiency of health coaching among type 2 diabetes and coronary artery disease patients
3. Better practices for identifying patients most likely to benefit from health coaching should be developed

2 | METHODS

TERVA was a prospective, longitudinal randomized controlled trial with three disease groups randomized into intervention and control groups. Recruitment of participants from the health care services has been described in detail previously.[19] A total of 2,594 patients who fulfilled the eligibility criteria (age 45 years or older, with T2D, CAD or CHF, and unmet treatment targets) were randomized to either the intervention group or the control group with a 2:1 ratio. Of the eligible patients, 1535 (59.2 percent) gave consent: 1,034 in the intervention group and 501 in the control group. There were no significant differences between the groups at baseline.[19] Patients with more than one disease were allocated to the highest morbidity disease group using the following hierarchy: 1) CHF, 2) CAD, and 3) T2D. T2D group criteria were medication and serum HbA1c > 7 percent (53 mmol/mol) without clinically evident cardiovascular diseases, for example, MI, stroke, or peripheral vascular disease. In this article, groups 1 and 2 are combined as one, the CAD group.

2.1 | Usual care

In Finland, general practitioners and nurses at primary care clinics provide basic medical treatment, follow-up, and support for compliance. Patients with T2D have 2-6 planned annual visits to a doctor or nurse, depending on how well the disease is under control. Primary health care wards provide basic care in wards for patient with less severe conditions who are unable to cope at home. Patients with complications are treated for acute needs in secondary care, either at outpatient clinics or as inpatients in hospitals. The CAD patients' treatment planning is provided in secondary care, in addition to 1-2 primary care visits per year. Patients in need of LTC receive home-delivered care, care at service home facilities or nursing homes, or care as inpatients at primary care level. Standards for care are set in the Finnish Current Care Guidelines, which are independent, evidence-based clinical practice guidelines.[20]

2.2 | Intervention

A detailed description of the health coaching intervention was published earlier.[19] In addition to routine care as described above, patients in the intervention group received health coaching by telephone over 12 months.

The intervention included eight key recommendations: 1) know how and when to call for help; 2) learn about the condition and set goals; 3) take medicines correctly; 4) get recommended tests and services; 5) act to keep the condition well controlled; 6) make lifestyle changes and reduce risks; 7) build on strengths and overcome obstacles; and 8) follow-up with specialists and appointments. Self-management booklets were sent to patients to support progress toward the key recommendations, and a traffic light system was used to visualize patients' progress. Health coaches had access to all electronic health records (EHR) in primary and secondary care and could enter patient data into the EHR.

The intervention group was called by the coach 10-11 times for 12 months. Quality control on the length, frequency, and content of calls was administered. The coaches were tutored individually and in groups throughout the intervention by a psychologist (PA) specializing in lifestyle change and strength-based behavioral coaching.[21] Overall intervention cost per patient was €419 per 12 months.

2.3 | Data

Data for the utilization and costs of health care and LTC were collected from the beginning of the intervention (2007) to the eighth year of postintervention follow-up (2016) from the Finnish national registries maintained by the National Institute for Health and Welfare. In Finland, each citizen has a unique social security code enabling full linkages to the national registries providing comprehensive data about each individual's use of health care and LTC. Primary care data were collected from the primary health care EHR from 2007 until 2011, after which the EHR were integrated into national registries (AvoHilmo) that provided data for 2012-2016. Secondary care data included the National Discharge Registry: the use of hospital outpatient care (all types of outpatient visits) and hospital admissions related to diagnosis (diagnosis-related grouping, DRG). LTC data were collected from Care Registers for Social Welfare, and it includes all types of long- and short-term institutionalized care, housing and residential services, and home care services.

EHR data included structured data for contact types (such as a visit, a phone call, or electronic messaging); the patient's age; the diagnosis (ICD-10); the reason for encounter (ICPC-2); and the employee category of the health care professional in the contact. Extracting the patient-level data from the patient administration systems (with diagnosis and contact information) made it possible to group each individual encounter type by the Ambulatory and Primary Care Related Patient Groups (APR) grouper, a grouping system equivalent to the DRG used in hospital care.[22] The APR groups were supplemented with cost weights indicating the relative consumption of resources. Cost weights were based on large samples of time measurements in primary care contacts and procedures to compile a relative value scale. All costs were deflated using the price index for public health care provided by Statistics Finland.

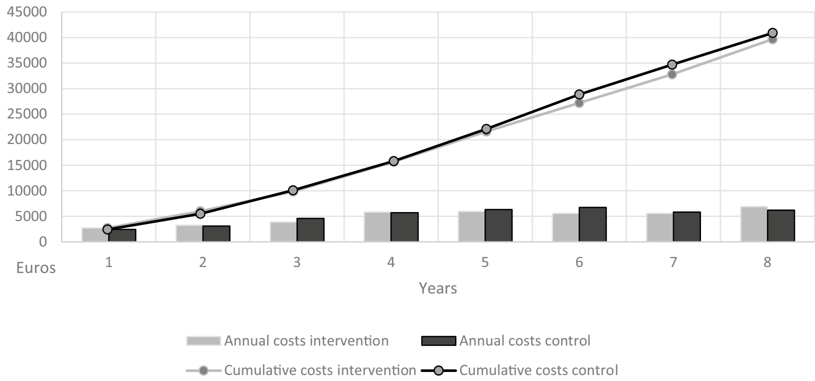
Hospitalizations and hospital outpatient visit due to any cause were extracted from the Hospital Discharge Register based on the International Classification of Diseases 10th revision (ICD-10) codes; the Finnish version of the Nordic Classification of Surgical Procedures (NCSP) codes for diagnostic and treatment procedures; and the respective NordDRG patient grouping classifications. The DRG cost weights for hospitalizations and outpatient visits were based on individual-level cost accounting data from several hospitals. The unit cost estimates for social care encounters and bed-days were derived from the national price list for unit costs of health care services in Finland.[23]

2.4 | Statistical analysis

Health care and long-term care costs were assigned to each patient over an 8-year follow-up period, and differences in mean costs between research arms were calculated. In the assessment of statistical significance of differences, we used nonparametric bootstrapping. Bootstrapping was used to draw a sample with replacement to calculate 1000 replicates of the mean difference in total costs (difference = mean costs in the intervention group − mean costs in the control group). Stata's bootstrapped 95% confidence intervals were used to indicate uncertainty in the mean difference estimator. The statistical significance of the difference of mean total costs per patient between the research arms was assessed using bootstrapped t test. Bootstrapping is a common method to account for the non-normality typical to cost data and for potential dissimilarity in cost distributions of the compared groups.[24] Intention-to-treat (ITT) strategy was applied, that is, all patients originally allocated to the intervention and control groups were included in the analysis. To assess the effect of the intervention among T2D and CAD patients, subgroup analyses were conducted. In addition to the main ITT analysis, per-protocol (PP) analyses were conducted excluding those of the randomized patients who did not perform any activities related to the study after giving their consent. The cumulation of cost over time was assessed by drawing cumulative cost curves for each research arm. Statistical analyses were performed using Stata version 15.0.

3 | RESULTS

The follow-up cost data were retrieved for 1033 patients in the intervention and 500 patients in the control group. One patient in each group was missing from the Finnish national registries, probably due to emigration. There were no significant differences in age and gender distribution between the research arms at baseline. The average age of participants was 65 and 65.4 years, and the proportion of females was 406 (39.3 percent) and 207 (41 percent) in the intervention and control groups, respectively. By the end of the eight-year follow-up, 26 percent (n = 269) of the patients in the intervention and 28 percent (n = 141) of the patients in the control group had become deceased.



Costs/year	1	2	3	4	5	6	7	8
ITT intervention	2705	5983	9859	15704	21638	27204	32806	39706
ITT control	2425	5515	10091	15815	22132	28880	34702	40916
No at risk/year								
ITT intervention	1004	970	941	897	854	821	793	764
ITT control	486	477	455	444	431	409	382	359

FIGURE 1 The cumulative and annual health care and LTC (long-term care) cost per patient over 8 years of follow-up

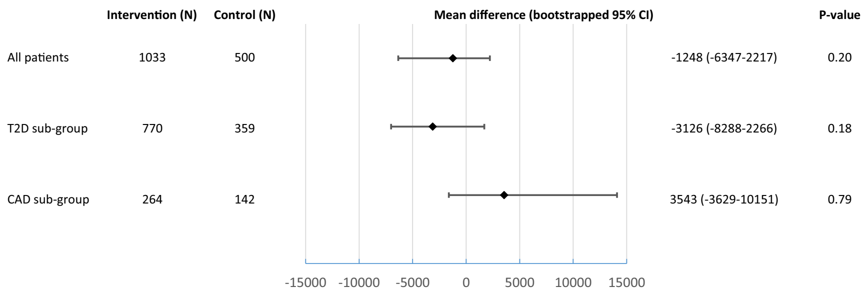


FIGURE 2 Mean difference in 8-year cumulative cost per patient and bootstrapped confidence intervals. Results among all participants and T2D and CAD subgroups

The cumulative cost curves per patient (Figure 1) showed that until a little more than two years after the beginning of the intervention, the cumulative cost was higher in the intervention arm than in the control arm. After this, however, the difference in cumulative cost changed sign, so that the cumulated cost was lower in the intervention arm. The difference grew steadily toward the end of the eight-year follow-up. The total costs accumulated per patient were €39 667 in the intervention group and €40 916 in the control group.

Figure 2 shows the mean differences in total cost per patient among all participants and T2D and CAD subgroups. For all participants, the total cost of care was €1248 (3 percent) lower in the intervention group than in the control group. The difference was, however, not statistically significant (95% CI from −6374 to 2217; $P = .20$). The subgroup analysis among T2D patients showed, in average, 7 percent lower costs (€−3126), while among CAD patients, costs were 10 percent higher (€3543) per patient in the intervention arm. Neither of these effects were, however, statistically significant.

To investigate where in the service system the changes in costs accrued, we calculated changes in the eight-year accumulated cost by different service types: primary care (visits and ward care), secondary care (outpatient and inpatient care), and LTC (home care, service homes, and nursing home). Among both T2D and CAD patients, the analysis revealed lower costs of secondary inpatient care and somewhat higher home care costs in the intervention group. Effects on other service type costs were mixed with mostly savings for T2D patients and increased costs for the CAD patients (Figure 3).

In the trial, there were patients in both intervention and control arms, who did not perform any activities related to the study after their consent and allocation to the intervention or control group. These patients were excluded from the PP analysis, resulting in 853

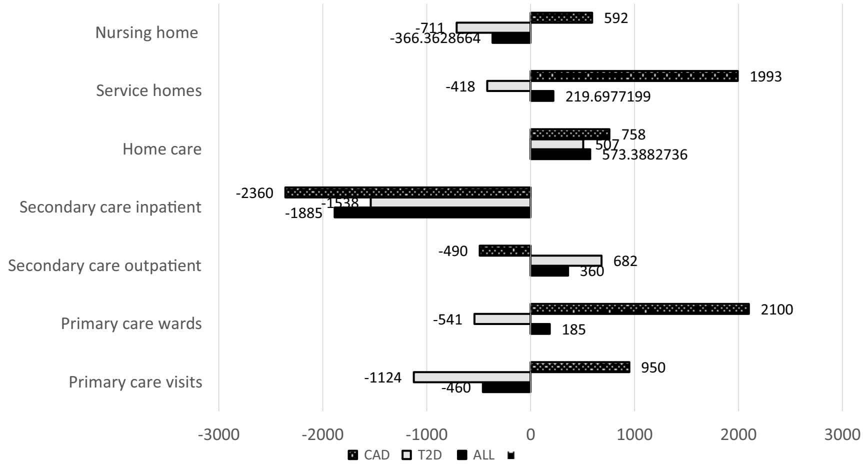


FIGURE 3 The incremental costs of tele-based health coaching per patient grouped by service types

patients in the intervention and 453 patients in the control arm. The proportion of the deceased was 23 percent (=197) and 26 percent (n = 119) in the intervention and control groups, respectively.

In the PP analysis, total costs were €35 863 and €41 816 per patient in the intervention and control groups, respectively. Until a little more than two years after the beginning of the intervention, the cumulative cost was slightly higher in the intervention arm than in the control arm. After this, however, the difference in cumulative cost changed sign to be lower in the intervention arm. The difference grew steadily toward the end of the eight-year follow-up (Figure S1). A statistically significant cost saving, €−5953 (14 percent), with a 95 percent bootstrapped confidence interval (CI) from €−9842 to €−1132 and *P* = .02 was found. PP analysis also showed a statistically significant cost saving of €−7287 (17 percent) per patient due to the intervention in the T2D subgroup (95% CI from €−12 528 to €−1760; *P* = .02), but no statistically significant effect in the CAD subgroup (Table S2).

4 | DISCUSSION

4.1 | Principal findings

We studied the eight-year cumulative health care and LTC costs of patients with T2D and CAD after a randomized controlled trial of a telephone health coaching program. At two years after the beginning of the intervention, the cumulative costs of the control group exceeded those of the intervention group and this difference remained until the end of the eight-year follow-up. However, the difference in the total cumulative costs per patient was not statistically significant. The average cost savings were greater in the T2D than in the CAD subgroup, but this result was neither statistically significant. Among both subgroups, cost savings were accrued in the secondary inpatient care, while effects on other health care and LTC costs were mixed.

4.2 | Comparison with other studies

To our knowledge, this study has the longest follow-up of the effects of health coaching on health care and LTC costs reported in the literature. Similarly, to previously reported health coaching interventions,[13,15-17] this intervention showed no reductions in health care costs in the first 12-month period.[19] As there is a delay from changes in patients' empowerment, learning, and behavioral changes to changes in physiological outcomes and following use of health services, a long-term follow-up of costs over 8 years after the intervention was conducted. In this study, after a little more than 2 years, the cumulative costs in the intervention group were steadily lower than in the control group. However, the difference in the accumulated 8-year costs was not statistically significant.

Three issues observed in our study may explain why costs in the intervention group were higher during a little more than two years after the beginning of the intervention. First, intervention highlighted the adequate and enough visits to health care for optimizing care and medication. Health coaches prepared patients for visits with health care providers and reflected with patients after the visits—building "a bridge between clinician and patient".[27]

This encouragement to collaboration with health care professionals may have at first increased patients' interest and need to consult their caregiver, and this might explain the increase of primary health care costs in the early stage of the follow-up. Second, building the health coaching program takes time, and the coaches keep developing their skills over the whole intervention period. In this study, all coaches had worked as nurses before the TERVA health coaching program and then trained to use the coaching methods and other skills required. Adaptation of new skills effectively took at least six months.[10] Third, patients were selecting multiple behavioral goals over the entire 12-month intervention and it must have taken even longer to gradually integrate the changes into their daily lives. With small but sustainable changes, clinical effects are also bound to be delayed.[26] This is contrary to studies assigning specific lifestyle goals to participants and implementing strategies for compliance, which may produce large effects at first, but these effects tend to diminish significantly over time.[25]

4.3 | Strengths and limitations

This study has several strengths. RCT design in a real-life clinical setting allows the strongest evidence of potential effects of an intervention in everyday clinical practice. Use of national registries allowed long-term follow-up of all the trial participants. The participants in our study represented two major noncommunicable diseases, T2D and CAD, both among the 10 most frequent causes of mortality in high- and middle-income countries.[28] We were able to conduct long-term follow-up and include the LTC costs, which to authors' knowledge have not been reported in any earlier study. LTC, such as residential facilities, cumulates cost over long periods of time and therefore contributes substantially to the total cost of care. There were no simultaneous interventions in the region.

Limitations exist, too. We were not able to blind the participants nor the health care professionals treating the patients. The intervention group was encouraged to be actively engaged in their treatment. This may have influenced the usual care they received. In our experience, some health care personnel perceived the intervention as threatening their areas of expertise, while others found that it added value to their clinical practice.

In this study, we were not able to fully assess patients' capabilities to participate in the coaching intervention. The inclusion of the participants was solely based on clinical inclusion criteria and EHR review. In both groups, there were patients who gave consent but did not participate in any other activities related to the study, for example, return study questionnaires or participate in the clinical measurements. These patients were found to have deceased earlier than those who performed at least some activities related to the study. Future research should attempt to define inclusion criteria that direct health coaching to those most potential to benefit from such interventions.

The intervention may have been too short for sustained effects to show. In the case studied, the early-stage observations and analysis on short-term 1-year follow-up showed increased cost in the intervention group, and the regional decision makers terminated the program after the one-year trial. Finally, the number of recruited patients may have been too small to observe statistically significant differences due to fairly large variation in individual costs. Despite the nonsignificant difference in cost of care, the intervention may still turn out to be preferable if we find marked improvements in long-term health outcomes.

Despite the steady improvements in diabetes care, approximately 50 percent of patients in Europe and the United States still do not achieve the targets of care.[29,30] While health coaching has been suggested as a feasible means to improve chronic care and avoid expensive complications, evidence of its cost-efficiency is still lacking. Randomized controlled trials with larger numbers of patients and on interventions more intense or exceeding one year may be needed to show strong evidence for the effect of health coaching. Careful attention must be paid to target the program to suitable patient segments and to execute the health coaching intervention appropriately.

ACKNOWLEDGMENT

Joint Acknowledgement/Disclosure Statement: Corresponding author (EO) received the special state funding (EVO) grant from Päijät-Häme Joint Authority for Health and Wellbeing. Finnish Innovation Fund (Sitra) and Pfizer Oy support our study for data acquisition and data analyses fees. Academy of Finland (IMPRO project, 312703) supported the authors (IH and ML) in this study.

CONFLICTS OF INTEREST

None of the authors have any conflicts of interest to disclose.

ORCID

Erja Mustonen https://orcid.org/0000-0002-7365-5855
Johanna Lammintakanen https://orcid.org/0000-0002-3333-1661

REFERENCES

1. Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: Translating evidence into action. *Health Affairs (Millwood)*. 2001;20(6):64-78.
2. Mattke S, Seid M, Ma S. Evidence for the effect of disease management: Is $1 Billion a year a good investment? *Am J Managed Care*. 2007;13(12):670-676.
3. Ellrodt G, Cook DJ, Lee J, et al. Evidence-based disease management. *JAMA*. 1997;278(20):1687-1692
4. Hayes E, MacCahon C, Panahi MR, Hamre T, Pohlman K. Alliance not compliance: coaching strategies to improve type 2 diabetes outcomes. *J Am Acad Nurse Pract*. 2008;20:155-162.
5. Palmer S, Tubbs I, Whybrow A. Health coaching to facilitate promotion of health behaviour and achievement of health-related goals. *Int J Health Promotion Educ*. 2013;41:91-93.
6. Olsen JM. Health Coaching: a concept analysis. *Nurs Forum*. 2014;49(1):18-29. https://doi.org/10.1111/nuf.12042.
7. Härter M, Dirmaier J, Dwinger S, et al. Effectiveness of telephone-based health coaching for patients with chronic conditions: A randomized controlled trial. *PLoS ONE*. 2016;11(9):e0161269. https://doi.org/10.1371/journal.pone.0161269.
8. Hale R, Giese J. Cost-effectiveness of health coaching. *Prof Case Manag*. 2017;22(5):228-238.
9. Dennis S, Harris M, Lloyd J, et al. Do people with existing chronic conditions benefit from health coaching? A rapid review. *Aust Health Rev*. 2013;37:381-388. https://doi.org/10.1071/AH13005.
10. Kivelä K, Elo S, Kyngäs H, Kääriäinen M. The effects of health coaching on adult patients with chronic diseases: A systematic review. *Patient Educ Couns*. 2014;97:147-157.
11. Tiede M, Dwinger S, Herbart L, Härter M, Dirmaier J. Long-term effectiveness of telephone-based health coaching for heart failure patients: A post-only randomized controlled trial (Article). *J Tele Telecare*. 2017;23(8):716-724.
12. Boehmer KR, Barakat S, Sangwoo A, Prokop LJ, Erwin PJ, Murad HM. Health coaching interventions for persons with chronic

conditions: a systematic review and meta-analysis protocol. *Sys Rev*. 2016;5:146. https://doi.org/10.1186/s13643-016-0316-3.
13. Steventon A, Tunkel S, Blunt I, Bardsley M. Effect of telephone health coaching (Birmingham OwnHealth) on hospital use and associated costs: cohort study with matched controls. *BMJ*. 2013;347:f4585. https://doi.org/10.1136/bmj.f4585.
14. Jonk Y, Lawson K, O'Connor H, et al. How effective is health coaching in reducing health services expenditure? *Med Care*. 2015;53(2):133-140.
15. Billot L, Corcoran K, MacDonald A, Powell-Davies G, Freyer A-M. Impact evaluation of system-wide chronic disease management program on health service. Utilization: A Propensity-Matched Cohort Study. *PLoS Med*. 2016; 13(6):e1002035. https://doi.org/10.1371/jornal.pmed.1002035.
16. Morello R, Barker A, Watts J, Bohensky M, Forbes A, Stoelwinder J. A telephone support program to reduce costs and hospital admissions for patients at risk of readmission: Lessons from an evaluation of a complex intervention. *Population Health Manage*. 2016;19(3):187-195.
17. Wagner T, Willard-Grace R, Chen E, Bodenheimer T, Thom D. Costs of a health coaching intervention for chronic care management. *Am J Managed Care*. 2016;22(4):e141-e146.
18. Oksman E, Linna M, Hörhammer I, Lammintakanen J, Talja M. Cost-effectiveness analysis for a tele-based health coaching program for chronic disease in primary care. *BMC Health Services Res*. 2017;17:138. https://doi.org/10.11867s12913-017-2088-4.
19. Patja K, Absetz P, Auvinen A, et al. Health coaching by telephony to support self-care in chronic diseases: clinical outcomes from The TERVA randomized controlled trial. *BMC Health Service Res*. 2012;12:147.
20. The Finnish Current Care Guidelines. http://www.kaypahoito.fi/web/english/home.
21. Rollnick S, Miller WR, Butler CC. *Motivational Interviewing in Health Care: Helping Patients Change Behavior*. New York, NY: The Guilford Press; 2008.
22. Honkasalo M, Linna M, Sane T, et al. A comparative study of two various models of organizing diabetes follow-up in public primary health care – the model influences the use of services, their quality and costs. *BMC Health Services Res*. 2014;12:26.
23. Kapiainen S, Väisänen A, Haula T. Terveyden- ja sosiaalihuollon yksikkökustannukset Suomessa. 2011. http://urn.fi/URN:ISBN:978-952-302-079-5
24. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Statist Med*. 2000;19:3219-3236.
25. The Look AHEAD Research Group. Cardiovascular effects of Intensive lifestyle intervention in type 2 diabetes. *N Engl J Med*. 2013;369:145-154. https://doi.org/10.1056/NEJMoa1212914.
26. Absetz P, Oldenburg B, Hankonen N, et al. Type 2 diabetes prevention in the "real world": Three-year results of the GOAL implementation trial. *Diabetes Care*. 2009;32(8):1418-1420.
27. Bennet H, Coleman E, Parry C, Bodenheimer T, Chen E. Health Coaching for Patients. Does your practice "give patients a fish" or "teach patients to fish"? Family Practice Management. 2010. http://www.aafp.org/fpm/20100900/p24.
28. World Health Organization. 2015. http://www.who.int/mediacentre/factsheets/fs310/en/index1.html.
29. Ali MK, Bullard KM, Saaddine JB, et al. Achievements of goal in U.S. diabetes care 1999–2010. *N Engl J Med*. 2013;368(17):1613-1624.
30. Stone M, Charpentier G, Doggen K, et al. Quality of care of people with type 2 diabetes in eight European countries. *Diabetes Care*. 2013;36(9):2628-2638.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Mustonen E, Hörhammer I, Absetz P, et al. Eight-year post-trial follow-up of health care and long-term care costs of tele-based health coaching. *Health Serv Res*. 2019;00:1–7. https://doi.org/10.1111/1475-6773.13251

ERJA MUSTONEN

This thesis provides eight-year follow-up results of effectiveness of heath coaching for chronic disease patients. The study was conducted using a multimethod and multidisciplinary approach which provide valuable findings for evaluating of effectiveness of complex intervention such as health coaching. The findings of the study can be also exploit on rational decision-making in relation to chronic disease management.

UNIVERSITY OF
EASTERN FINLAND

uef.fi

PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND
Dissertations in Social Sciences and Business Studies

ISBN 978-952-61-3688-2
ISSN 1798-5749