University of Oslo
Institute of Clinical Medicine
Faculty of Medicine
Department of Cardiothoracic surgery
Oslo University Hospital

KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Public Health
and Primary Care

Optimising patient discharge and follow-up after surgical aortic valve replacement to reduce readmissions and improve patient-reported outcomes

Stein Ove Danielsen

Jury:

Supervisor UiO: Dr. Irene Lie
Supervisor KUL: Prof. dr. Philip Moons
Co-supervisor: Prof. dr. Theis Tønnessen
Co-supervisor: Dr. Svein Solheim
Co-supervisor: Prof. dr. Marit Leegaard
Chair examining committee: Dr. Kristina Haugaa
Chair public defence: Prof. dr.
Jury member: Prof. dr. Lars Wallin
Jury member: Prof. dr. Paul Herijgers
Jury member: Dr. Kristina Haugaa
Jury member: Prof. dr. Theo van Achterberg

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Biomedical Sciences

December 2019

Table of content

1 Acknowledgements

The Ph.D. project was a clinical research project within the Department of Cardiothoracic Surgery at Oslo University Hospital in Oslo, Norway. The research project was possible to initiate with funding from the Bergesen Foundation and the Raagholt Foundation. This project has been made possible by the Norwegian ExtraFoundation for Health and Rehabilitation. The work was done between 2015 and 2019 within the region of Health South-East in Norway.

A special thanks to Irene Lie for making the opportunity to be educated in the work of science. It resulted in a PhD-Thesis based on the work associated with the Aortic Valve Replacement readmission -AVRre- Study Irene initiated. Professor and nurse Philip Moons at Katholieke Universiteit Leuven -KUL- in Belgium has provided substantial scientific supervision during the whole process. I am ever grateful for Philips support and mentorship. Philip and Irene have set the bar for further development in the work of science. Professor and cardiac surgeon Theis Tønnessen provided valuable support and inputs through the whole process. To Svein Solheim, Marit Leegaard and Leiv Sandvik; Thanks for the support and supervision. Thanks also to Irene Sandven for valuable support and supervision with paper II. The 24/7 Hotline staff, the ICU Nurses, provided valuable and necessary support to complete the intervention. Edith, Gro, Kirsti, Monica, Janne, Kenneth, Tor-Henning and Jonny; Thanks! Lena Korsnes, the head of the Cardiac Operating Room Dep. at Ullevål deserves a special thanks for her understanding she has demonstrated in her leadership of professional development for both individuals and the nursing profession.

I wish to especially thank the patients and their caregivers for the participation in this study. I hope I have paid their contributions with honour.

Last, and definitively not least, all the warm thanks for the patience and support from my family Christine, Eira and Thea: You are simply the best!

2 Abbreviations

AHA: American Heart Association

ANCOVA: Analysis of Covariance

AS: Aortic Stenosis

AVRre: Aortic Valve Replacement readmission

CABG: Coronary Artery Bypass Surgery

CONSORT: Consolidated Standards of Reporting Trials

COPD: chronic obstructive pulmonary disease

CPH: Cox Proportional Hazard

ESC: European Cardiology Society

EQ-5D-3L: EuroQol

GP: General Practitioner

HADS: Hospital Anxiety and Depression Scale

HF: Heart Failure

HRQoL: Health-Related Quality of Life

ICU: Intensive Care Unit

ITT: Intention-To-Treat analysis

LHL: Landsforeningen for Hjerte-og Lungesyke (Norwegian National Association for Heart and Lung Disease)

LMM: Linear Mixed Model

LOS: Length of stay

LVEF: Left ventricular ejection fraction

MI: Multiple Imputation

MRC: Medical Research Council

NOS: Newcastle-Ottawa Scale

OECD: Organization for Economic Co-operation and Development

OUH: Oslo University Hospital

PC: Project Coordinator

PP: Per-Protocol analysis

PREM: Patient-Reported Experiences Measures

PROM: Patient-Reported Outcome Measures

PROSPERO: International prospective register for systematic reviews

RCT: Randomised controlled trial

SAVR: Surgical Aortic Valve Replacement

SCG: Supra Coronary Graft

SCQ-16: Self-administered Comorbidity Questionnaire

SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

StaRI: Standards for Reporting Implementation Studies

TAVR: Transcatheter Aortic Valve Replacement

TFU: Telephone follow-up

VHD: Valvular Heart Disease

WHO: World Health Organization

30-DACR: Thirty-day all-cause readmission

3 List of papers

This thesis is based on the following original publications, which are referred to by their Roman numerals (Papers I-V):

I. Lie I, Danielsen SO, Tønnessen T, Solheim S, Leegaard M, Sandvik L, Wisløff T, Vangen J, Røsstad TH, Moons P. Determining the impact of 24/7 phone support on hospital readmissions after aortic valve replacement surgery (the AVRre study): study protocol for a randomised controlled trial. *Trials* 2017; 18:246

II. Danielsen SO, Moons P, Sandven I, Leegaard M, Solheim S, Tønnessen T, Lie I. Thirty-day readmissions in surgical and transcatheter aortic valve replacement: A systematic review and meta-analysis. *International Journal of Cardiology* 2018; 268, 85-91

III. Danielsen SO, Moons P, Sandvik L, Leegaard M, Solheim S, Tønnessen T, Lie I. Impact of Telephone Follow-up and 24/7 Hotline on 30-day Readmission Rates Following Aortic Valve Replacement -A randomized controlled trial. *International Journal of Cardiology* 2019; Jul 30. pii: S0167-5273(18)36036-4. doi: 10.1016/j.ijcard.2019.07.087. (Published online: July 30, 2019; Article in Press)

IV. Danielsen SO, Moons P, Leegaard M, Solheim S, Tønnessen T, Lie I. Facilitators of and barriers to reducing thirty-day readmissions and improving patient-reported outcomes after surgical aortic valve replacement -a process evaluation of the AVRre Trial. Submitted to peer-review journal (October 2019)

4 Introduction

The main objective of this thesis research was to investigate and optimise hospital discharge and follow-up of patients after surgical aortic valve replacement (SAVR). This investigation was carried out in the context of the Aortic Valve Replacement Readmission (AVRre) trial (ClinicalTrials.gov Identifier: NCT02522663). The AVRre trial tested the efficacy of a post-discharge telephone support intervention designed to reduce readmissions after SAVR and improve patient-reported health and quality-of-life outcomes. In support of this randomised controlled trial (RCT), a systematic review and meta-analysis of the medical literature was conducted to uncover and assess the worldwide magnitude and variability of thirty-day all-cause readmission (30-DACR) rates after SAVR and transcatheter aortic valve replacement (TAVR). Finally, a process evaluation was conducted on the intervention implementation and the patients' and staff's reactions to the intervention to determine whether it was carried out as intended.

4.1 Overview of surgical aortic valve replacement (AVR)

Aortic stenosis (AS) is the most common type of valvular heart disease (VHD), leading to SAVR treatment in Europe. [1] Many conditions can cause the tissues comprising the valve leaflets to become stiffer. Functionally in AS, the valve opening is narrowed, reducing blood flow. If the valve becomes so narrow (stenotic) that overall heart function is reduced, blood flow will be inadequate to the rest of the body. Severe AS is mainly the product of a degenerative change (calcification of the valve) or a congenital condition (bicuspid valve), resulting in AS. [1]

The prevalence of AS increases with age, and due to growth of the ageing population in Europe, is projected to continue to increase in coming years. [2] This nexus of demographics and disease trajectory is a cause for great concern, because Europe's population of people older than 65 years is estimated to nearly double from 2008 to 2060. [3] Today in

2019, for patients younger than 75 years with severe AS, they are treated with SAVR. However, a heart team should consider what is the best treatment plan for patients >75 years, evaluating whether SAVR, TAVR, or medical therapy is the best possible treatment. Optimally, the heart team should comprise cardiologists, cardiac surgeons, imaging specialists, anaesthetists and, if needed, general practitioners, geriatricians, and heart-failure specialists, cardiac electrophysiologists, or intensive-care specialists. [4] The European Society of Cardiology (ESC) recommends using the following flowchart as a guideline for the management of severe AS (Fig. 1).

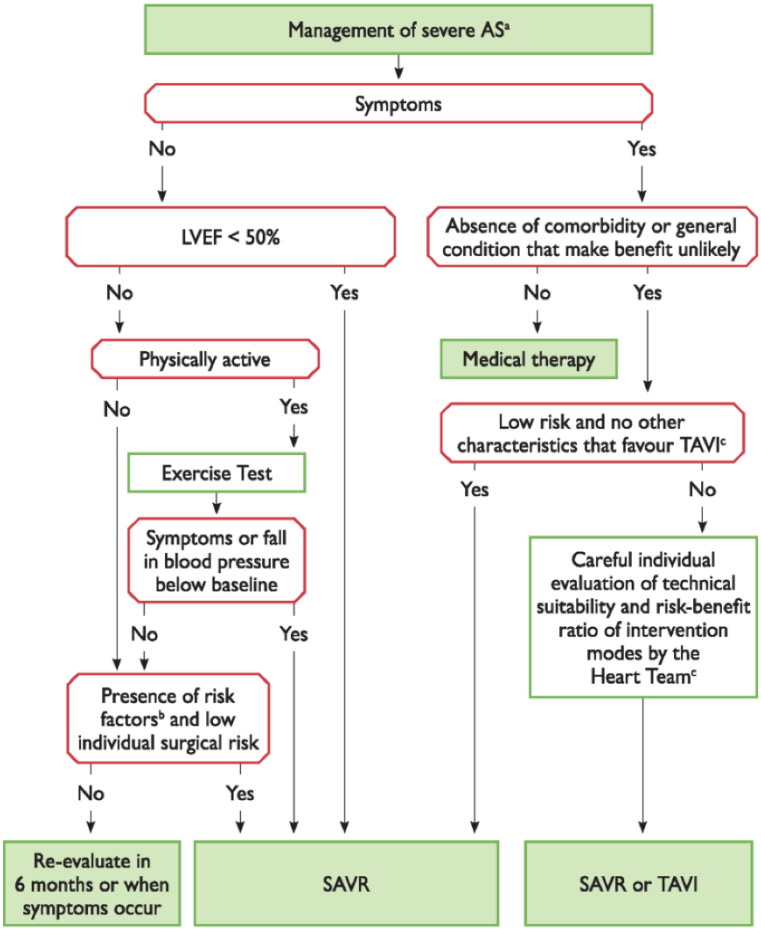


Fig. 1. Decision flow chart for the management of severe AS. (Reproduced with permission of the ESC, 2019. Original in [5]). Transcatheter aortic valve implantation (TAVI) is synonymous with TAVR. LVEF: Left ventricular ejection fraction

The 2016 annual report from the National Heart Registry in Norway reported that surgeons in 2015 conducted 1007 single-valve surgeries and 376 valve surgeries performed concomitantly with coronary artery bypass grafting (CABG) treatments. The proportions of females receiving the two kinds of valve surgeries were 46.2% and 26.6%, respectively. The surgical valve treatments of adults with valve diseases (including TAVR, starting from 2008) have increased annually from 645 in 1995 to 1736 in 2015. From 2004 to 2015, both SAVR and TAVR procedures have increased in Norway. [6]

In terms of treatment choices, the population of AS patients who are good candidates for invasive treatment has changed over time. TAVR has now emerged as the preferred choice of treatment — and the one superior to medical treatment — for patients who are not good candidates for surgery [7]. The nature of the SAVR procedure has also changed, originally comprising primarily mechanical valves to mostly comprising (> 80%) biological valves by 2010. [6, 8] The situation in Norway has followed the same trend from early 2000. [6, 8]

In-hospital clinical outcomes after cardiac surgery are well described. [9] However, patient-reported outcomes regarding perceptions of health and quality of life after hospital discharge for cardiac surgery are more sparsely reported. [10, 11] Heart failure (HF), cardiac rhythm disorders, and infections are common complications after discharge for SAVR, which often result in readmissions. [12, 13]

4.1.1 Transcatheter aortic valve replacement (TAVR)

The first TAVR procedure was performed in 2002 by Cribbier. [14] TAVR is widely believed to be superior to medical therapy for AS patients cleared for aortic valve replacement. [15] However, for moderate- to lower-surgical risk AS patients, sufficient evidence is lacking on whether replacing surgical treatment in favour of the minimally invasive TAVR is superior in terms of long-term survival and other clinical outcomes. [16] Robust evidence on the number of adverse events after TAVR, such as necessity of a permanent pacemaker, vascular

complications, and paravalvular regurgitations, ought to be clearly reduced before recommending that less symptomatic AS populations receive TAVR. Predictors of poor outcomes after TAVR include chronic lung and kidney disease (30-35% and 30-50% of the TAVR population, respectively), together with frailty [16] This lack of evidence for replacing surgery with TAVR is exemplified by the observation that the 30-day all-cause readmission (30-DACR) rate after TAVR was scarcely described before 2015. [17]

4.2 Hospital readmissions after AVR

In Norway, 30-day readmission is operationally defined as an unplanned and acute hospital admission for any cause to any hospital within 30 days after hospital discharge. [18] Thus, one can calculate over a population of patients and time period, what proportion is readmitted (i.e., percentage). The 30-DACR rate after SAVR is reported to be about 20%, based on a US population sampled between 1999 and 2011. [19] In Denmark, the 30-DACR rate in 2015 was reported to be up to 25% after valve surgery. [20] The 30-DACR rate after SAVR is unclear, reportedly ranging between 6.5-25.5%. [21, 22] In Norway, the risk-adjusted probability for 30-DACR reported in 2015 for elderly persons >67 years was 14.7%. [18] This value was determined on the basis of five diagnoses: asthma/COPD, heart failure, pneumonia, stroke, and bone fracture. Mean length of stay (LOS) in hospital for the first readmission in Norway within 30-days was 6.84 and mean days to the first readmission was 12.5. [18] The 30-DACR rates for SAVR populations, averaged over age (only patients > 18 years), are rarely described in the literature.

Hospital readmissions incur high costs. In the USA, for example, it is estimated that readmissions reached $17 billion, based on Medicare statistics (2005-2008) for patients >65 years) [23]. In Norway, the costs were 2 billion Norwegian Kroner (NOK) (reported in 2012). [24] In the USA since October 1, 2012, hospitals could be fined for excessive readmissions for certain kinds of diagnoses. [25] The Hospital Readmissions Reduction Program (HRRP) is

responsible for assigning economic penalisation and determining the threshold risk-standardised readmission ratio for certain conditions/procedures.

Increased efforts to prevent more readmissions have followed these financial disincentives, and research is well underway to monitor the patient health effects when readmission rates decline, especially how they affect mortality rate. [26] In Belgium, hospitals are penalised if a readmission occurs within 10 days after discharge. [27] Norway has no economic penalisation for readmissions. However, Norway offers positive economic incentives to hospitals for reducing hospital LOS. However, municipalities are penalised if they do not accept admitted hospital patients when the hospital has defined a patient as being discharge ready. [28] However, most SAVR patients in Norway are transferred to home from hospital. This could, in theory, represent a risk for early hospital discharge followed by readmission. [29] The Organization for Economic Co-operation and Development (OECD) reported for 2015 that the hospital discharge ratio (number of patients discharged from hospitals after at least one-night stay per 100,000 inhabitants[1]) was ~16% in Norway and Belgium. [28] Austria and Germany had the highest (~25% each), and Colombia, Mexico, Brazil, and Canada had the lowest (3-8%). [28] The mean hospital LOS reported by OECD in 2015 was 8 days across all OECD countries, ranging from 4 (Turkey) to 16 (Japan). [28]

Reports of the 30-DACR rate often come from registry studies. National- or hospital-level administrative or clinical databases are used to extract relevant readmission data. [30, 31] The National Patient Registry (Norsk pasientregister; NPR) in Norway is considered to be a high-quality patient registry, containing readmission data that are available to researchers. Researchers can gain access for minimal payment and with necessary ethical approval. [32] From 2009, hospitals in Norway have issued a unique NPR-identification number to every

[1] The hospital discharge ratio includes deaths in hospital following inpatient care. Same-day discharges are usually excluded (OECD (2019), Hospital discharge rates (indicator). doi: 10.1787/5880c955-en (Accessed on 26 October 2019).

admitted patient. [32] This number provides a way to record and track any new hospital stay a patient might have after discharge from the original treating hospital. This tracking procedure permits an accurate calculation of the 30-DACR rate. For the USA, any hospital readmission within 30-days of the initial discharge contributes to the 30-DACR rate. For Belgium, only readmissions to the same hospital where initial treatment was conducted contributes to the 30-DACR rate. [27] Examining medical charts and/or contacting patients (telephone interview, mail, survey) after discharge are ways of obtaining data for calculating the 30-DACR rate besides using registry data.

A 2015 annual report for Norway reported a readmission rate of 15% for adults > 67 years, [18] which is higher than Belgium's 2008 rate of 5.2% for adults > 17 years. [33] This means that 15% and 5.2%, respectively, of all discharged patients in these two scenarios get re-admitted to hospital within 30 days. This alternative way shows how different countries determine the 30-DACR rate. Within-country differences in the 30-DACR rate are also sometimes reported among hospitals after surgical treatment. For example, in US hospitals with high surgical volumes and lower mortality rates, fewer readmissions have been reported. [34] When interpreting and comparing readmission rates across different countries or hospitals with differing profiles, this diversity in procedures warrants caution. Thus, when publishing readmission rates, unequivocal and transparent reporting is paramount, especially with regard to how a readmission is defined, how admission data are collected and validated, and how they are analysed. Presently, there is no evidence-based guideline for consistently reporting 30-DACR rates.

The 30-DACR rate is often used as a quality indicator for hospital care performance [35], which might represent a valid proxy measurement for the quality of care after surgical treatment in a hospital. [34] Errors that interrupt the quality of healthcare delivery can be caused by structural or processual factors or be a natural consequence of the patient's co-

morbidity and clinical condition. [36] These could result in an adverse event like a hospital readmission. [36] Hence, in healthcare contexts, a readmission is often considered to be an adverse event. [36] Preventing readmissions are therefore an obvious goal for clinicians, as they are for administrators responsible for the readmission-related costs.

The number of preventable readmissions can be estimated, and the proportion of preventable readmissions can be as high as 79% (and as low as 5%). [37] We have found no reporting on the proportion of preventable readmissions after SAVR. Moreover, we do not know to what extent the 30-day interval is an appropriate period to assess when the objective is to optimise the discharge and follow-up after SAVR. Being readmitted to a hospital interrupts the expected care pathway and represents an extra burden for patients. Risk of iatrogenic errors are present in this situation, e.g., hospital-acquired infections or other complications affecting functional and/or cognitive status. [38]

4.3 Optimising discharge and follow-up after SAVR

Patient discharge is initiated at the hospital, and patient follow-up involves several steps before discharge results in a patient transferring to home, or more seldomly, to a healthcare facility having the appropriate level of care for SAVR patients (e.g., ordinarily a rehabilitation centre or a nursing resident home). Hospital discharge can be viewed as a journey in some ways, having multiple stops and transitions. It has been described this way:

> *'…hospital discharge is not an end point, but rather is one of multiple transitions occurring during the patient's care journey. The organisation and provision of this transitional care typically involves multiple health and social care actors, who need to co-ordinate their specialist activities so that patients receive integrated and, importantly, safe care.'* [39]

If a discharge increases patient satisfaction and quality of life and does not eventually lead to a hospital readmission due to prior hospital treatment within six weeks after discharge, it can be viewed as a successful discharge. [40]

The hospital discharge initiates the transition of care. In the health services, the transition of care is a concept having multiple definitions. Indeed, the World Health Organization (WHO) states that the concept incorporates more than just the act of clinical hand-over in healthcare, but should also comprise the views and values of the patients. [41] WHO refers to the American Geriatrics Society's definition of transition of care: *'a set of actions designed to ensure the coordination and continuity of health care as patients transfer between different locations or different levels of care within the same location'*. [41] As mentioned, the 30-DACR rate is considered to be a quality indicator of hospital service, and this rate, if too high, motivates investigations to improve discharge process and optimise follow-up after hospital stays. Providing necessary monitoring and management of patient symptoms after discharge are significant actions associated with the reduction of hospital readmissions. Promotion of self-management through patient education might also be a beneficial way to reduce hospital readmissions. [42]

Braet stated that if appropriate information is not provided, healthcare provider and management continuity can interrupt the transition of care and therefore disrupt the care continuum. [27] In Norway, the transition of care after SAVR is a primary concern of the university hospitals discharge management team, whose task includes transferring the patient to a local hospital. Then, the transition of care mostly ends with patients going home, with primary care being a responsibility of the general practitioner (GP). The GP is also responsible for patient follow-up. Some of the patients are directly transferred from hospital to a rehabilitation centre before going home. The transition of care for SAVR patients includes a prominent shift in roles in a rather short period, a shift from being a patient cared for by the

hospital to a private citizen being solely responsible for his own health. The 2014 national patient-reported survey of Norway on patients' experiences with hospital stays revealed that patients are often dissatisfied with the discharge process. [43]

Different discharge optimisation interventions have assessed how and whether they reduce the 30-DACR rate. Leppin et al. (2014) found that peri-interventions (around both in- and out-patient treatments) reduce the 30-DACR rate. [44] They also found that interventions conducted before 2002 were 1.6 times more effective than those conducted after 2002. [44] To explain this decline, it was hypothesised that improvements in care over time were either not recognised in the control group descriptions or simpler, fewer complex interventions were tested in years after 2002, such that they were inappropriate for the time period. More interventions measuring and reporting readmission rates differently, or more interventions with fewer human contacts also could have contributed to the finding of less effective interventions in reducing the 30-DACR rate after 2002. [44] There is also evidence that complex interventions with several components seem to be more effective than single-component interventions. [42, 44] Few surgical populations were included, and none were SAVR patients. Moreover, there was indication of publication bias. [44]

Hansen and colleagues defined interventions as either pre-, post-, or bridging interventions, and they found that no single intervention alone reduced the 30-DACR rate. [45] However, an RCT with a general medical adult population demonstrated that post-discharge telephone follow-up (TFU) had promising effects on reducing readmission rates. [45] Still, few RCTs on post-discharge TFU intervention RCTs of high methodological quality have been reported that show reduced 30-DACR rates. [45] To the best of our knowledge, no complex post-discharge TFU intervention with a bridging purpose (i.e., link between hospital and home by a 24/7 hotline) has been conducted that aimed to reduce the 30-DACR rate after SAVR.

According to the Donabedian model — the most cited and used framework for instituting quality improvement in healthcare — structural and processual factors of an intervention should also be analysed in order to improve healthcare quality. [46] Indeed, the intervention should not solely relate quality to the outcomes. [46] In this model, the structural factors include the hospital context where processual factors take place (e.g., the interactions among healthcare professionals and patients that occur during diagnosis and treatments), culminating in the outcomes of the intervention. The Donabedian model represents a logical approach to achieving quality improvement through which one also analyses the factors leading to the outcomes in order to establish excellent quality care. Using a mixed-methods approach, as in the AVRre trial, that includes sampling the participants and nurses' views on structural, processual and/or contextual factors, can deepen our understanding of the discharge process after SAVR. Hence, this rationale embodied in the Donabedian model motivated the research design for the AVRre telephone support intervention, allowing a broader and richer evaluation of its effects.

Patient-reported outcome measures (PROMs), metrics to explore how a patient experiences a disease or health condition, are now widely used. However, to be a scientifically valid measurement, PROM must be appropriate for the study context and aims, [47] and it must be transparently reported in the format recommended by the Consolidated Standards of Reporting Trials (CONSORT) extension of 2013. [48] Patient-reported experience measures (PREMs), metrics to explore how patients' experience healthcare services, are considered to be a valuable way to assess care quality. However, many PREM instruments still need more empirical evidence to overcome methodological issues related to its measurements and interpretations. [49] PREMs are considered to be useful for exploring patient perspectives when evaluating, for example, the applicability and usability of an intervention. [50] Therefore, both PROM and PREM instruments can produce valid results for evaluating the

effectiveness of an intervention. These kinds of measurements are also recommended for use in mixed-methods studies designed to evaluate a complex healthcare intervention, its implementation, and impact within an appropriate framework. [50]

4.4 Theoretical scheme: The Medical Research Council framework

The theoretical scheme used to frame and conduct this thesis research is described in the Medical Research Council (MRC) guidance on process evaluations of complex interventions in healthcare. [51] This highly cited framework recognises the value of process evaluation for RCTs, stating that it: *'…can be used to assess fidelity and quality of implementation, clarify causal mechanisms, and identify contextual factors associated with variation in outcomes.'* [52] We used the 2015 updated guidance, which elaborated and detailed the three themes for process evaluation described in the 2008 MRC guidance: implementation, mechanisms, and context. [51]

The logic flow in the process evaluation of an intervention as presented in the MRC framework has a similar structure to that described in the Donabedian model for quality assessment and improvement in healthcare. One key is recognising that it is imperative to analyse the processes prior to the care outcomes in order to construct a more complete picture of an intervention's relevance and potential. [46] The slightly modified MRC model (Fig. 2) used in this thesis research shows how the MRC framework was used to organise our investigation on how the TFU intervention impacts discharge after SAVR.

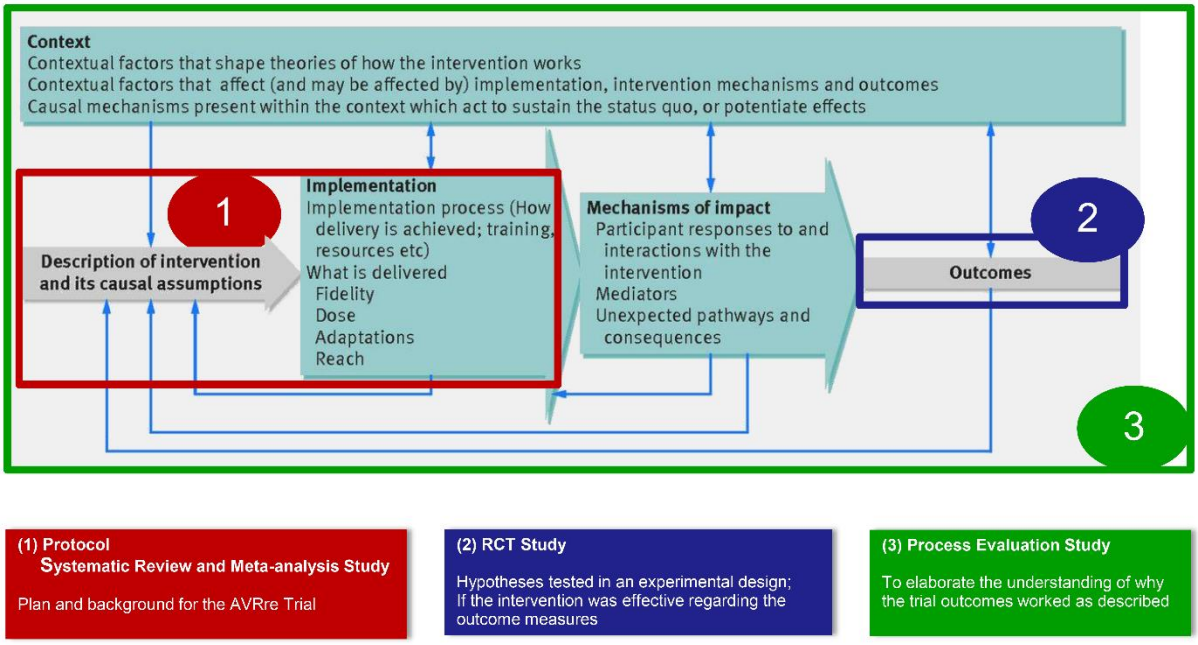


Fig. 2. Slightly modified MRC model used in organising, conducting, and evaluating this thesis research. (adapted from [51]; with permission from Graham F. Moore)

The MRC framework emphasises the need for a planned, prospective evaluation of RCT implementation; mechanisms of impact (patient reactions to the intervention); and contextual factors that influence the intervention. This approach is complemented by conducting a more traditional process outcome evaluation of the intervention in a RCT. The first step for the implementation aspect is to evaluate the development and piloting (including doing a feasibility check) phase of an intervention, which could reveal uncertainty related to procedural, clinical, or methodological issues. The outcome of this first step could highlight problems and lead to changes in implementation. [53] In the AVRre trial, we carried out the pilot and feasibility analysis together in an integrated approach, rather than conducting them separately as two unrelated steps. The second step for the implementation was to evaluate how the delivered intervention in the main trial was performed (i.e., its fidelity and dose). A planned evaluation strategy would allow tailored prospective and/or retrospective data collection for the process evaluation along with the outcome reporting, as the MRC

recommends. One reason for research waste in conducting RCTs can be traced to an inadequate development phase before the trials are fully tested in the main trials. [54, 55] Moreover, analysing qualitative findings related to quantitative results on interventions shows that such a mixed-methods approach can result in a deeper and broader understanding of an intervention. [56, 57] What is the 'gold standard' for evaluating a clinical intervention?

A RCT in healthcare science is often used to test the effectiveness of an intervention on selected outcomes. [58] Statistical analyses will provide answers as to whether the investigated intervention works or not on a targeted disease or an adverse medical event. Clinical trials producing nonsignificant results can lack sufficient statistical power to explore why results are negative. [59] Clinical trials are conducted in a real-world context, not a laboratory where one has more control over experimental variables to better pin down causal relationships between variables and outcomes. Applying a healthcare intervention in a clinical context is challenging. Why? Unlike in laboratory settings, in clinical settings it is difficult, even impossible, to control for potential confounding variables, to avoid experimenter or subject biases, to avoid random errors, or to choose the right outcomes. [59] Moreover, these experimental obstacles of RCTs are typically compounded by positive trial outcomes failing to be translated into clinical practice, [59] or if they are translated, by delays in getting the trial results into the hands of everyday practising clinicians. [60]

Modifying the pipeline to the clinic may be one way to get more positive clinical research results translated into clinical practice. That is, researchers might design clinical trials that integrate the participants'/patients' perspectives and views into the intervention and reported outcomes. [59] Providing trial participants an opportunity to express their experiences and feelings about their health and care during the RCT is not only appropriate within a mixed-methods design but might also go a long way towards achieving society's goals of making healthcare more patient-centred. [61] In short, greater participation of

healthcare consumers in the care system may improve overall healthcare. Analysis of Norwegian clinical medical guidelines developed between 2000 and 2009, for example, show that these were mostly developed without any patient involvement. Moreover, related literature searches show that they failed to include patient perspectives. [62]

Whether an intervention in an RCT design is complex or not, or even inherently complex, is an ongoing debate: '*We now think of a complex intervention as much more than the sum of its components parts. Its effects are likely to be modified by both the site and process of implementation.'* [53] According to Kernick, the issue of complexity in research emerged in the late 1980s, and many definitions of this concept have been proposed. [63] Briefly, one definition he provides captures the following essential elements:

> *'The [complex] system is different from the sum of the parts. In attempting to understand a system by reducing it into its component parts, the analytical method destroys what it seeks to understand. The corollary is that the parts cannot contain the whole and any one element cannot know what is happening in the system as a whole.'* [63]

Moreover, Kernick states that applying complexity theory in healthcare science might challenge the dominant positivistic view of science, in which there should be one correct answer to a problem, towards which all research will converge. However: *'Perhaps a more realistic perspective is to see complexity theory complementing existing approaches but alerting us to the importance of matching the research approach to the context and complexity of the environment to which it is applied.'* [63] Scriven characterises outcome research that is insulated from its 'how and why' as a kind of 'black box' evaluation. [64] By contrast, a 'clear box' evaluation provides a full explanation of how and why an intervention works. [64] However, the problem of a black box evaluation might not be overcome when considering that the increase in complexity can expand exponentially by adding a single

component to a complex intervention conducted within a health service that is a complex non-linear system. [65] All of these aforementioned considerations affected the design of the AVRre trial.

In the AVRre trial, after the initial design considerations, we reasoned that organising and conducting the study using the MRC framework could provide deeper and broader insight into the workings and potential clinical applications of the results. Using this scheme could also specifically inform the healthcare service about ways to optimise the discharge and follow-up care after SAVR. Prospective data collection using a mix of methods for evaluation purposes were integrated into the project from the beginning, especially with regard to including trial participants' perspectives into the new knowledge produced from the AVRre trial.

We reasoned that it would be appropriate and beneficial to use the MRC framework, which acknowledges the complexity of the intervention and the attendant problems that can emerge from conducting it within a complex non-linear health-services system. Using this organising scheme would also permit a better understanding of the AVRre trial outcomes and other potential important effects that might be translated quickly into clinical practice.

5 Aims of the study

5.1 Overall aim

The overall aim of this doctoral thesis research was to determine whether a remote post-discharge intervention could reduce hospital readmissions after aortic valve surgery and improve patient-reported health and quality of life. It included defining the current state of knowledge regarding 30-DACR rates after valve surgery and conducting specific process evaluations of AVRre trial reporting, implementation, and context. The AVRre trial was an RCT conducted in a university hospital in Norway.

5.2 Specific aims

I. To determine whether transparency was achieved in reporting the outcomes of the AVRre trial, according to the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines.
II. To determine what the current reported proportion, causes, and risk factors of 30-DACR rate are after SAVR and TAVR through a systematic review and meta-analysis of relevant medical literature.
III. To determine the effectiveness of a post-discharge 24/7-telephone support intervention after SAVR on 30-DACR rate, patient symptoms of anxiety and depression, and perceived health state in the AVRre trial.
IV. To determine whether the AVRre trial programme activities were implemented as intended through a formal process evaluation of trial implementation, patient responses, and contextual factors.

6 Methods

This thesis comprises four published articles in peer-reviewed journals, in which the main objective was to investigate an aspect of hospital discharge and follow-up after SAVR. Table 1 presents an overview of the AVRre trial design and its relationship to this thesis.

Table 1. Overview of essential elements of the AVRre trial and its relationship to peer-reviewed thesis articles.

	Paper I	**Paper II**	**Paper III**	**Paper IV**[a]
Design	Protocol	Epidemiological	Randomised controlled trial (RCT)	Process evaluation of RCT
Methods	Qualitative and quantitative	Qualitative and quantitative	Quantitative	Qualitative and quantitative
Recruitment and sample	Literature	Systematic literature search N=141,102	Prospectively from a tertiary hospital, adult patients admitted for SAVR N=288	Prospectively from a tertiary hospital, adult patients admitted for SAVR, (N=288); Prospectively and retrospectively, nurses from a university hospital, (N=5); Retrospectively prior to study 1, former cardiac patients, (N=5)
Data collection	Literature review	Literature review	Self-report questionnaires Medical chart review	Self-report questionnaires Field notes Interviews Focus group interview

Analysis	Power calculation	Meta-analysis Meta-regression Qualitative review	Descriptive statistics Kaplan Meier Survival analysis Univariate GLM analysis Cox Proportional Hazards analysis Linear mixed-model analysis	Descriptive statistics Qualitative analyses

[a]Paper IV is submitted and is under review (see Section 3), as of 30 October 2019.

[b]Total N patients for all papers included in the review.

AVRre = Aortic Valve Replacement Readmission Study; SAVR = Surgical aortic valve replacement; RCT = Randomised controlled trial; GLM = General linear model.

6.1 Design and study sample

Patients scheduled for SAVR at Oslo University Hospital (OUH) were included in the AVRre trial. The trial was conducted in the OUH Department of Cardiothoracic Surgery at Ullevål and Rikshospitalet in Oslo, Norway. The first patient was enrolled in the trial on 24 August 2015. Inclusion of participants for the trial ended March 2017. The AVRre trial cohort was followed for one year after AVR surgery. The AVRre trial was registered at ClinicalTrials.gov (NCT02522663) on 11 August 2015.

Figure 3 presents a flowchart showing how participants were selected for, allocated, and followed up in the AVRre trial. To be included in the trial, a patient had to meet the following criteria: adult (> 18 years); elective surgery as a single SAVR (mechanical or biological), SAVR + coronary artery bypass surgery (CABG), SAVR + supra-coronary graft (SCG), or SAVR + CABG + SCG; understand and write the native language (Norwegian) well; and be able to be contacted by phone and use a phone after hospital discharge. Patients were excluded from the trial if they had a stay in the intensive care unit (ICU) for more than 24 hours or experienced any complications that would have prevented them from being assessed for any of the inclusion criteria. A total of 482patients were assessed for study eligibility (Fig. 3).

All patients were informed about the study before their SAVR surgery and given time to consider whether or not to participate. Recruitment was done with the knowledge that, before major surgery, a patient is vulnerable with regard to making decisions. [66] So, we were careful not to pressure patients to participate. The Declaration of Helsinki [67] informed our implementation of the ethical approval of the trial (see section 8.3 Ethical considerations). After patients gave their consent to participate, they were randomly allocated to the control or intervention group (Fig. 3). The control group was assigned to ordinary scheduled discharge management care before they were discharged to home. For the control group, a primary-care

GP was the ultimate healthcare professional responsible for their follow-up. We used block randomisation (size varied from 8-12) in a 1:1 ratio, produced using a web-based algorithm provided by the Unit of Applied Clinical Research, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway [68].

The pilot study (n=10) for the AVRre trial was conducted from April to May 2015. We conducted interviews with former cardiac patients recruited by the Norwegian National Association for Heart and Lung Disease (Landsforeningen for Hjerte-og Lungesyke; LHL) in February and March 2015. A focus group with the hotline staff was held retrospectively to explore their experiences with the hotline. The semi-structured interview-guide is provided as supplemental material in paper IV.

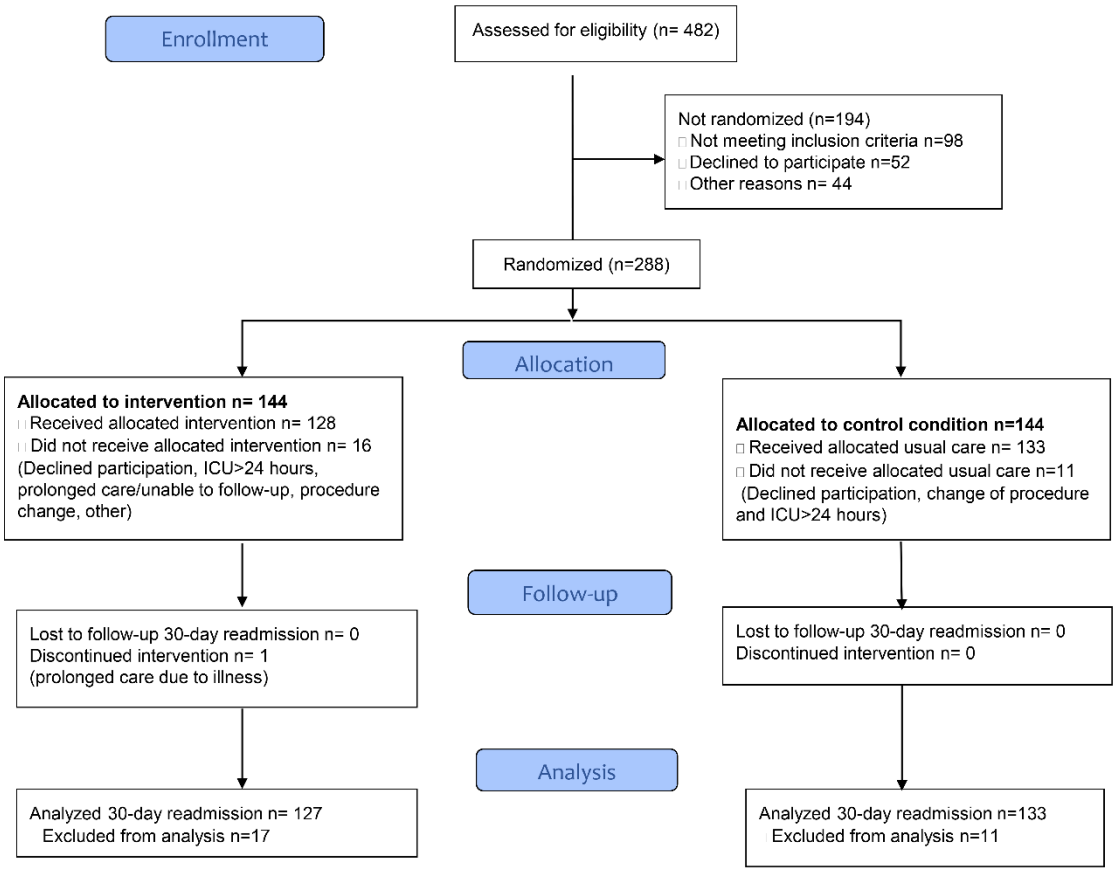


Figure 3. Flowchart for participant selection, treatment-type allocation, and follow-up in the AVRre trial

6.2 Telephone support intervention, training, and AVRre trial hotline manual

Participants in the AVRre intervention received post-discharge phone support in two ways: 1) they could freely call the 24/7 hotline staffed by experienced cardiac ICU nurses to access evidence-based health information when needed; and 2) they would receive a scheduled phone call on day 2 and day 9. The 24/7 hotline was available for the participants in the intervention group for 30 days after discharge from local hospital; they were explicitly told not to share the hotline number with other patients. The two scheduled phone calls could happen if the participant was discharged from the hospital to home or to a cardiac rehabilitation (CR) facility.

At the time of discharge from the university hospital where the participants received their SAVR, the project coordinator (PC) met the participants face-to-face and provided them verbal and written information about to which group they had been allocated (i.e., randomly). Only the PC and recruiting personnel knew the patients' group allocation before the day of discharge (or the day before). Participants allocated to the control group and hospital staff were not present when allocation information was given.

The information leaflet distributed to the participants reminded them of their participation in the AVRre trial. Only the intervention group received leaflets that contained necessary information like the hotline number. This leaflet also encouraged the intervention group participants to use the AVRre 24/7 hotline number, or a general medical emergency number, in case they were experiencing acute symptoms. The control group received information leaflets that reminded them about the importance of their follow-up (i.e., usual care; see questionnaires below) in the trial and that contained a note of gratitude for their participation. The PC followed the intervention patients' transition of care, was informed about discharge time from the local hospital, and then sent them an SMS to schedule the time of day for their first TFU call.

The TFU to the intervention group participants on day 2 and day 9 was a structured telephone call. That is, all intervention group participants were asked the same questions in the same order and were given the same information and reminders. The PC was prepared prior to the call with detailed information contained in the participant's medical and nurse charts about their health condition and in-hospital medical development after SAVR at the university hospital. The call also served as a reminder about their option to call the 24/7 hotline if they needed information or advice about managing their post-discharge self-care. The call also included advice about the positive effects of engaging in physical activity in the early CR phase. [69] Finally, the PC would answer any questions the patient might have about their present health condition. The hotline staff nurse was assigned to the phone service one week at the time. Concurrently, the PC had a paired phone to assist the hotline staff at any time. The participants were 'primed' to expect a possible short delay in the hotline response, if the nurse happened to be occupied with other tasks while on duty. An automatic recorded response would also state this possible delay, after which the participant could leave a recorded message. After a short time, the participant could expect to be called back if they had left a message; hotline staff and the PC gave these recorded messages priority.

Prior to the pilot and main AVRre trial implementation, to facilitate the implementation, we had meetings with key medical and nursing personnel that were involved in the care of SAVR patients in the hospital. In separate meetings, the cardiac surgeons and cardiologists were informed about their role in the AVRre trial, which was to be available for consultations with the hotline staff, if necessary. We conducted an orientation session in the emergency call centre to discuss experiences with listening, investigating, and responding to a phone call. We also studied and noted the way the emergency centre documented their work when carrying out their work by telephone. The local nurse and physician leaders were also informed about the trial, and we also met with the head of the Department of Patient Safety

and Quality at OUH to discuss and clarify hospital staff responsibilities related to the hotline service provided during the AVRre trial.

The hotline staff prepared for the intervention by attending one two-hour educational and training session before the pilot began. The PC discussed relevant background and outlined the rationale for the AVRre trial. A professional development nurse from the emergency call centre (113/911 quick-dial emergency numbers) also gave lectures on how to engage in active listening and shared experiences, gave useful advice, and answered questions from the attendees. In addition, at the end of this educational/training session, the hotline staff were given the opportunity to participate in role play, where they could practise answering and using the hotline manual under the guidance of the researcher conducting the session. After the session but before the pilot and the main study began, the hotline manual was made available to all staff involved in the AVRre trial for background reading and to prepare for the actual trial.

The evidence-based 24/7 hotline manual contained medical advice, elaborated information for the nurses related to the advice, and pertinent references. [29] The organisation of the themes in the manual was based on the experiences of former cardiac patients and on the universal convention of colour-coding red, yellow, and green in defining the emergency level of the calls. The manual was always available in the ICU ward, and the hotline staff also had a portable version with them when they were not present in the ward. More information about the manual can be found in paper I. [29]

Picture 1. Reproduction of the cover page of the portable 24/7 Hotline Manual used in the AVRre trial (Norwegian)

The PC was available to assist hotline staff on the paired phone whenever they needed relief from the hotline service for practical or other reasons. The paired phone also allowed the PC to monitor the number of hotline calls and the duration of the calls. The ICU nurses were accustomed to being on call for duty, as being on call was part of their ordinary work schedule. The PC was available for case consultations with the hotline staff at any time and held regular meetings during the main AVRre trial to discuss cases and how these were handled. We focussed on the most challenging calls and how they were perceived and interpreted by the staff.

Moreover, educational sessions with a specialist dealing with themes related to early rehabilitation were conducted in the main trial in order to support and empower the hotline staff. A cardiac surgeon conducted one educational session on handling dyspnoea issues; a cardiologist conducted one session on arrhythmias, especially focussing on atrial fibrillation; and a PhD cardiac rehabilitation physiotherapist conducted one session on post-SAVR patient physical activity and training.

6.3 Study procedures

The participants in the AVRre trial completed a baseline questionnaire before and up to one year after surgery (postal survey). The 30-DACR events were obtained by reviewing the medical charts. Table 2 shows the timeline for acquiring the data measurements in the AVRre trial.

Table 2. Timeline for data measurements in the AVRre trial.

Data type	Time after surgery					Time after trial
		Prospective				Retrospective
	Before surgery	1 month	3 months	6 months	1 year	
Demographic	X					
Clinical	X					
Co-morbidity	X					
HADS Questionnaire	X	X	X	X	X	
EQ-5D-3L Questionnaire	X	X	X	X	X	
30-DACR		X				
PROM and PREM survey			X			
Field notes	X	X	X	X	X	
Qualitative	X					X

AVRre = Aortic valve replacement readmission; 30-DACR = Thirty-day all-cause readmission from medical charts; HADS = Hospital Anxiety and Depression Scale; EQ-5D-3L = EuroQol-5D-3L; PROM = Patient-Reported Outcomes Measures; PREM = Patient-Reported Experience Measures; Qualitative interviews were conducted prior to surgery and after the main trial.

Paper I present the detailed study protocol that was ultimately used in the AVRre trial. The AVRre study sought to determine whether 24/7-phone support after discharge for SAVR reduces hospital readmissions within the 30 days after discharge from hospital (i.e., 30-DACR). For paper I, we followed the guidelines of the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) in reporting the details of the protocol paper [70]. The paper was submitted while recruitment was ongoing (November 2016). Using

the SPIRIT guidelines helps one to prevent selective reporting of study outcomes and offers transparency of the RCT for the benefit of the study population. [70]

The hotline manual described in paper I was based on thorough literature searches of relevant medical literature databases, including Medline, Cochrane, and Embase. We also acquired and studied the information leaflets of other cardiac surgery centres in Norway in order to enhance and refine our 24/7-telephone support manual used in the AVRre trial. Moreover, information about patient experiences after hospital discharge — especially during the first month — was obtained through focus group interviews with former cardiac patients (N=5) and through an interview with one participant organised through the Norwegian LHL.

A semi-structured interview guide [71] was developed and used during the interviews in order to consistently obtain data. Prior to the interviews, a mind map [72] was completed by the participants to enhance their recall during the interview. [73] Finally, the content of the hotline manual was appraised by two physicians and a nurse specialist with experience in early rehabilitation for cardiac surgery patients. Design of the manual was further informed by the Norwegian Medical Index for acute medical support. [74] Supporting material for paper I contains a translated excerpt from the hotline manual. Also included in the supporting material for paper I are the SPIRIT checklist we completed for the AVRre trial and examples of the informed written consent form used for the AVRre trial.

Paper II reported on the results of a systematic review and meta-analysis of papers in the medical literature to examine the overall incidence, causes, and risk factors of 30-DACR rate after SAVR and TAVR. On 30 March 2016, we prospectively entered the plan (PROSPERO no. 42016032670) (PROSPERO 2016 CRD42016032670; for conducting our systematic review and meta-analysis in PROSPERO, an international prospective register for systematic reviews in health and social care, welfare, public health, education, crime, justice, and international development, where there is a health-related outcome

(https://www.crd.york.ac.uk/PROSPERO/). We adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines in reporting the findings. [75]

We conducted a systematic search of the relevant medical literature databases after consultations with a senior research librarian at Oslo University Medical Library, whose supervision, professionalism, and experience ensured that we conducted an accurate systematic search in the selected databases. This rigour increased our chances that the search would capture relevant articles according to the aims of the study. The Patient/problem, Intervention, Comparison, Outcome (PICO) framework [76] was used to specify the search in relevant databases (details can be found in paper II). The systematic assessment led to the included papers and relevant numerical results for the analyses described in paper II (search strategy is shown as supporting material in paper II). The 30-DACR rates, study- and patient-level covariate data were collected and entered into a Microsoft Excel® spreadsheet. After identification of candidate papers from the literature search, two researchers with knowledge of the project independently assessed the full text of potential papers to be included in the review. Agreement for inclusion was reached through discussions between these two researchers. We used the Newcastle-Ottawa Scale (NOS) [77] for assessing the quality of the included papers. More details are provided in paper II.

Paper III reported on the outcomes of the AVRre trial. The CONSORT statement checklist from 2010 [78] was used and completed to ensure that we accurately reported the outcomes of the trial. Participants' demographic data, relevant clinical data, and data on their co-morbidities at baseline were collected from the medical charts and from the baseline questionnaire (Table 2). The baseline questionnaire used a self-report of co-morbidities, the Self-administered Comorbidity Questionnaire (SCQ-16). [79] Summaries of selected demographic, clinical, and co-morbidity data of the AVRre trial participants were presented in paper III.

The 30-DACR rate data were collected from the participants' medical charts for all hospital stays and from their responses in the questionnaire that was completed 3 months after the start of the intervention (Table 2). For the primary outcome variable, number and latency to readmission(s), we collected the following data: elapsed time to readmission; day of week for readmission; readmission at a university or local hospital; diagnoses (cause of readmission); and length of readmission stay. These data were additionally used in an ancillary analysis to estimate the proportion of avoidable and unavoidable readmissions in the study cohort. Two physicians and a nurse specialist (all members of the AVRre project group) independently estimated the proportions of avoidable and unavoidable readmissions for the study cohort (blinded for the participants' group allocation to intervention or usual care control).

PROMs, such as the Hospital Anxiety and Depression Scale (HADS) and the EuroQol (EQ-5D-3L) questionnaire, were used to assess the effects of the intervention on secondary outcomes. We used the Norwegian version of the HADS by obtaining a licence from the GL Assessment and the trusted translated version from Mapi Research Trust [80]. We also obtained permission to use the Norwegian version of the EQ-5D-3L. We measured the effect on health-related quality of life (HRQoL) [11] and on perceived health state. The latter is a measure equivalent to HRQoL and was used in this thesis when reporting results from the EQ-5D-3L questionnaire. EQ-5D-3L is a generic measurement of the respondents health states. It measures five dimensions of a respondent's perceived health state: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. [81] For each of the five dimensions, participants are scored on a 3-point scale (1, no problem; 2, some problems; 3, severe problems). Combining the dimension scores yields a 5-digit number, which equates to 1 of 243 possible combinations of health states. This score was converted to an index value using a value set derived by a time-trade-off (TTO)- and a VAS-based technique of a UK

population. [82] The EQ-5D-3L also uses a visual analogue scale (VAS) that ranges from 0 (absolute worst health state you could imagine) to 100 (perfect health state). Participants can select their overall current health condition on the VAS. Importantly, we assessed whether a ceiling or floor effect was present in the scores.

The HADS questionnaire was originally developed to assess whether patients in non-psychiatric hospitals might have anxiety and/or depression; it excludes questions about somatic symptoms associated with anxiety and depression in order to prevent interference with their somatic conditions. [83] The HADS questionnaire has seven items related to anxiety and depression. Each item is scored on a 4-point Likert scale (0-3 points), yielding a total item score of 0-21. [84] An item score of 8-21 points was considered to be indicative of symptoms of anxiety and/or depression. Using a cut-off score of 8 should give an acceptable balance between sensitivity and specificity (0.80). [83] Cronbach's alpha was used to measure the scale's internal consistency; a score > 0.7 is considered to be acceptable. [85]

The hotline staff always had easy access to a form used to register participant data relayed during the hotline calls: participant name, date and time of the call, elapsed time of the call, and an open form field where they could note keywords describing the content of the conversation. The form also included a section in which hotline staff could indicate (by tick mark) the caller's perceived symptoms or concerns and severity level of each symptom/concern (green, yellow, red) and the staff response to these concerns (i.e., whether or not they provided advice from the hotline manual). Picture 2 reproduces information on the call data registration form.

Set X for symptoms/clinical signs/questions from the patient. Additionally, log which red, yellow or green responses given:

		X	Codes (related to given advices)		
Symptoms/sign/question:			Red	Yellow	Green
Dyspnoea	1				
Heart rhythm	2				
Pain	3				
Prescribed medication	4				
Infection	5				
Psychological (anxiety, depression, cognitive)	6				
Activity	7				
Nutrition/Lifestyle	8				
Social network	9				
Sexuality	10				
Others	11				

Picture 2. Section in the call data registration form hotline staff used to register and assess participant information for each hotline call. This section allows staff to rate the caller's symptoms/concerns and symptom severity level (green, yellow, red); to note their response to the symptoms/concerns; and to describe advice they provided to the caller.

For paper IV, we used the MRC framework (see section 4.4) to guide the broader evaluation of the intervention. Moreover, the Standards for Reporting Implementation Studies (StaRI) guided the reporting of the process evaluation of the implementation, mechanisms of impact, and contextual influence of the intervention. [86] The qualitative design in the AVRre trial was informed by the methodological approaches described by Maxwell, [87] Malterud, [88] and Kvale and Brinkman. [71]

Three months after the start of the trial, the participants completed a follow-up questionnaire (i.e., the 3-month questionnaire). To get a better idea of the participants' experiences with the hotline and with their discharge, in general, we included three questions on how the intervention group participants used and experienced the hotline and questions on how all the participants experienced their hospital discharge. The questionnaire also contained an open-ended comment field in which participants could provide written feedback not captured in the direct questions. The first of the three questions for the intervention group was

a PROM question, which was to be answered by actual users of the hotline: ‘Were you satisfied or not with using the hotline?’ The second and third questions were PREM questions: ‘To what degree did having access to the hotline give you a sense of feeling safe?’ ‘To what degree did you think the hotline was a good offer?’ All three of these questions had six possible choices: not applicable, not at all, to a small extent, to some extent, largely, and to a very large extent. These questions were added to the questionnaire for evaluating purposes and were developed by the research group. The remaining questions were six PREM questions on the hospital discharge experiences of all participants; one question on whether a readmission had actually occurred (yes/no); and one question on whether it could have been prevented by the hospital (yes/no/don’t know), if indeed a readmission did occur. The six PREM questions had six possible choices, as indicated above. With the permission of the developers of these PREM questions (which were the same as those in a national survey reported in 2015), [43] we integrated them into the survey for evaluation purposes. The questionnaire was presented in Norwegian.

For the AVRre trial, all participants received a structured follow-up call from the project group PC on days 2 and 9 after hospital discharge to home. The PC collected the following data obtained during these calls: date and time, elapsed time of the calls, and contents of the conversations. These data were entered into a secure Excel® spreadsheet for later analysis. No personal identifiable data were recorded. In addition, the PC systematically took field notes from all the other encounters (regular meetings, educational sessions, and consultations) related to the hotline service in the prospective intervention, and these were preserved for clarification of what occurred during the encounters.

In the AVRre trial, research interviews were conducted prior to the pilot to inform and refine the content of the hotline manual. A convenience sampling of former SAVR patients was used in order to better understand their experiences with early rehabilitation after

hospital discharge. [71] Two small focus groups (2+2) and one single interview were conducted. The interviews were digitally recorded and transcribed verbatim. Details of this procedure are described above in the *paper I* paragraph of this section.

A small pilot study (N=10) was also conducted to evaluate the logistics, recruitment, randomisation, hotline and telephone follow-up, and method to inform patients of their group allocation in the actual AVRre trial. A visit to the medical emergency call centre and discussions with key medical and nurse leadership in the hospital were conducted to inform, facilitate the implementation, and determine factors that could potentially challenge or undermine the conduct of the intervention.

We also conducted a retrospective focus-group interview with hotline staff members (N=5). The participants were notified prior to the focus group through an email, which contained a reminder describing what they should try to recollect about their preparations and what was to take place during the intervention. A semi-structured interview guide was used to facilitate and guide the focus group interview, which was digitally recorded. This approach allowed us to collect information about the implementation and about the participants' reactions from the hotline staff's perspective, which also could provide important clues about potential mediators of an effect or unexpected outcomes related to the intervention. [89]

6.4 Data analysis

6.4.1 Quantitative analysis

In paper I, we presented the power calculation of the sample size we would need for the AVRre trial. The sample size was based on published data about hospital readmissions of patients > 65 years old in Norway. [90] We expected that the readmissions of participants in the intervention group would decrease by 10% compared to that in the control group, with a power of 80% and a risk of type I error of 5%. This yielded a sample size estimate of 143 in the two arms of the trial.

In paper II, we conducted a meta-analysis with sub-group analyses and a univariate meta-regression analysis of 30-DACR rates reported in the medical literature. For the meta-analysis, we used the DerSimonian-Laird method, [91] pooled the 30-DACR rates, and calculated the overall incidence of 30-DACR after SAVR and TAVR. These rates were presented in a Forest plot (Fig. 4). I^2 statistics were used to evaluate the heterogeneity between the studies. A sub-group analysis based on the collected participant covariates was conducted to evaluate heterogeneity. This was extended into a univariate meta-regression test using a random effects model to analyse whether the heterogeneity estimates were affected by the covariates. We conducted a sensitivity analysis to assess the robustness of the overall results. [92] All analyses were performed with STATA version 14.0 [93] and MedCalc version 16 [94] statistical software. NOS was used to assess the quality of the included studies. [77] Publication bias was assessed by visual inspection of Funnel plots and estimated using the Eggers test. [95]

In paper III, demographic and clinical data (i.e., categorical data) were presented as proportions (real numbers and percentages), whereas continuous data were presented as means or medians with standard deviations (SDs). Pearson chi-square tests and Fisher's exact tests were used to evaluate differences between the intervention and controls groups for the categorical data, whereas independent t-tests or Mann-Whitney U test were used to assess group differences for the continuous data. An intention-to-treat (ITT) analysis (N=282), along with a per-protocol (PP) analysis (N=260), were performed to evaluate the primary outcome (30-DACR). The Pearson chi-square test (between the groups) was used to evaluate the effect of the intervention on 30-DACR. To determine the time to readmission within the 30 days after hospital discharge, we conducted a Kaplan Meier Survival analysis, followed by log rank tests to evaluate any group difference. [96] A Kaplan Meier Survival plot was made to visualise the groups' time to readmission.

Analysis of covariance (ANCOVA) was performed to assess the intervention effect on the secondary outcomes at each of the assessment times (see Table 2), adjusting for the baseline scores using the covariates. This form of regression analysis is suitable for detecting an intervention effect with appropriate power. [97] A Linear Mixed Model (LMM) analysis was applied to measure the between-group differences in the secondary outcomes on repeated measures (up to one year after SAVR). The baseline score, time variable, and group were designated as fixed factors, whereas the intercept was designated as a random effect. LMM was an appropriate statistical analysis to use for our longitudinal data in the AVRre trial, because it allowed us to analyse both fixed and random effect factors in the modelling. [98]

Missing data are unavoidable in clinical and longitudinal studies and can cause analysis problems. Because most statistical tests assume that the dataset is complete [99], analysing incomplete datasets (e.g., leaving out entire cases with some missing data) can bias the results. [100] To address this issue, we analysed the missing value patterns of participants' data and performed multiple imputation (MI) with 20 iterations in each model for the secondary outcomes. [100]

A Cox Proportional Hazards (CPH) regression model was applied to explore predictors of 30-DACR after SAVR. The CPH model is often used to investigate the effect of multiple variables when a specific event will take place within a specific time span. [101] The chosen model was adjusted for other variables, using an appropriate number of covariates for the final model.

The assessment of the proportions of avoidable and unavoidable readmissions in the study population was assessed by two physicians and a nurse, who were blinded to the participants' group allocation, but they did have relevant clinical data available. This approach was deemed appropriate according to a recommendation for such assessments. [102] The readmissions were classified as either avoidable, unavoidable, or disagreement/questionable.

We chose not to resolve disagreements or readmissions deemed questionable due to an expected margin of error caused by individual physicians' preferences and different local healthcare systems.

In paper IV, we presented the descriptive statistics as numbers, percentages, and standard deviations. Fisher's exact test was used for comparative analyses of categorical variables with small numbers of cases.

6.4.2 Qualitative analysis

In paper II, we presented the results of the systematic literature review of the risk factors for and causes of 30-DACR after SAVR. The summaries of these factors and causes are presented as percentages in tables and in the corresponding text.

In paper IV, we qualitatively analysed the content of prospectively collected project staff field notes, memos, registration forms, and questionnaire narratives and the transcripts of a retrospectively conducted interview of participating staff. NVivo software, version 10 and 11 Pro, [103, 104] was used to organise the transcribed text from interviews, written questionnaire narratives, and field notes. Organising the text material and coding their content themes into meaningful text units are the two first steps in doing systematic text condensation, a qualitative analysis method described by Malterud. [88]

In the first step of this qualitative analysis, the texts were thoroughly read in order to gain an overview of the texts' content. In the second step, meaningful text units were retrieved and coded. NVivo was used in the second step to organise the codes and match them up with their associated text units. This is an important step, as it provided us with an overview of the data before proceeding to the third step of the analytical process. In the third step, we did text condensation, wherein the codes were abstracted into categories (meanings). In the final step, overarching themes were constructed; these represent the main findings (descriptions and/or concepts). [105] The technique of critical reflection was applied throughout all steps of the

analysis. [87] This involved maintaining close proximity with the relevant theory, by which the researcher moved back and forth between the analysis steps and the theory. This procedure was operationalised in that the researcher actively tried to challenge the validity of similarities between the codes and their categories. This analysis strategy requires time to achieve an appropriate level of critical reflection.

In our qualitative research study, the analytical process was not linear, in which data collection occurs entirely prior to the analysis phase. By contrast, in our study, analysis started and was conducted in parallel while the data were being collected. This enabled the researchers to carefully avoid, for example, confirmation bias and also to be open to various possible narratives, such as ones in which information tangentially related to the research questions could be considered. Such efforts were taken in order to prevent construction of one's own preconceptions and to simply prevent repeat acquisition of already-known information just to achieve appropriate scientific qualitative analysis. The critical reflection technique was supplemented through critical discussions of the preliminary and final findings with another researcher, who critiqued and challenged these findings. In our opinion, this procedure bolstered the trustworthiness of the findings obtained through qualitative analyses. It also lent support to the notion that qualitative approaches are essentially equivalent to statistical analytical approaches when it comes to validity assessments.

7 Brief summaries of the results

7.1 Paper I

Determining the impact of 24/7 phone support on hospital readmissions after aortic valve replacement surgery (the AVRre Trial): Study protocol for a randomised controlled trial

Paper I was a protocol paper that was timely published to ensure transparency in the reporting of the AVRre trial outcomes. Therein, we reported on the detailed protocol for the AVRre study (Table 1). We presented reasons why we believed that instituting a complex around-the-clock intervention within a university hospital-based setting would be an effective strategy for reducing the high readmission rates to hospital after SAVR. The paper presented the primary and secondary outcomes we would evaluate and presented the printed manual for conducting the telephone support.

In paper I, we also presented the power calculation for a reasonable sample size we would need to detect a 10% decrease of readmissions in the intervention group compared to the control group. We concluded that the knowledge gained from the AVRre trial would provide valuable insights for adjusting aspects of the healthcare system now and would likely highlight areas that could be improved in caring for SAVR patients after hospital discharge.

7.2 Paper II

Thirty-day readmissions in surgical and transcatheter aortic valve replacement: A systematic review and meta-analysis

In *paper II*, we reported on an investigation that determined the overall proportion of the 30-DACR rate and causes of and risk factors for 30-DACR after SAVR and TAVR (Table 1), as reported in the medical literature. The meta-analysis pooled the total numbers of patients. The proportion of 30-DACR following SAVR was 17% (95% CI: 16-18%), and for TAVR it was

16% (95% CI: 15-18%). Causes of 30-DACR after SAVR and TAVR were similar to those reported in the literature, with heart failure, arrhythmia, infection, and respiratory problems being the most frequently reported causes. A comprehensive list of risk factors for 30-DACR after SAVR has not been reported in the literature. The independent risk factors most frequently associated with 30-DACR after TAVR were diabetes, respiratory illness, atrial fibrillation, kidney illness, and using the transapical approach for inserting a new valve.

By examining subgroups in the reviewed papers, we found a higher proportion of readmissions in multicentre studies (SAVR, 20%; TAVR, 18%) versus single-centre studies (SAVR and TAVR, both 12%). Also, we found a higher proportion of readmissions in multicentre studies in the USA (18%) versus other countries (14%). Retrospective studies (17%) also had a higher incidence of readmissions compared to prospective studies (SAVR, 14%; TAVR, 11%). Only 6 prospective studies were included versus 26 retrospective studies.

Examining heterogeneity using meta-regression in univariate mode, we found that a higher proportion of readmissions in multicentre versus single-centre studies; both populations were significantly associated with the readmission rate (SAVR, *P*= 0.013; TAVR, *P*= 0.038). Furthermore, we found a weak association between a higher readmission rate in the TAVR population in the USA versus other countries (*P*= 0.091).

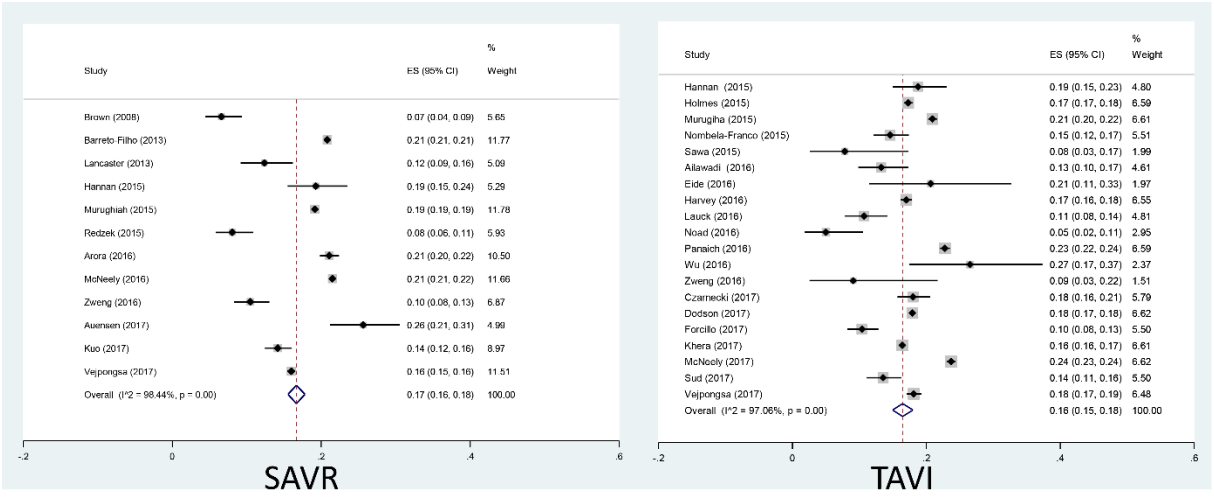


Figure 4. Forest plots summarising a meta-analysis of the proportion of 30-day all-cause readmission rate after surgical aortic valve replacement (SAVR) and transcatheter aortic valve replacement (TAVR)

A quality assessment of the included papers revealed that most studies did not include a transparent validation statement of the readmission statistics.

7.3 Paper III

Impact of telephone follow-up and 24/7 hotline on 30-day readmission rates following aortic valve replacement – A randomised controlled trial

In *paper III*, we reported on an investigation (Table 1) that determined the effectiveness of the AVRre telephone support intervention in reducing the 30-DACR rate, symptoms of anxiety and depression, and improving the SAVR patients perceived health state. The results revealed that the intervention had no significant effect on the 30-DACR rate (P= 0.274). However, the intervention was effective in reducing symptoms of anxiety within one month after discharge (P= 0.031), but this reduction did not persist up to one year after SAVR surgery. The 24/7 telephone intervention also had no effect on reducing symptoms of depression (P= 0.758) or on improving the patients' perceived health state (EQ-5D-3L VAS, P= 0.636) up to one year after surgery.

Total unplanned 30-DACR rate was 22.3% in this cohort, and 83% of all readmissions occurred within 14 days after hospital discharge. The most frequent cause of readmission was cardiac rhythm disturbance (34%), in which atrial fibrillation was prominent. Interestingly, 14% of the readmissions were caused by pericardial effusion. Independent risk factors for 30-DACR after SAVR were symptoms of anxiety before surgery (P= 0.003) and pleural drainage after surgery but before hospital discharge (P= 0.027). We also observed that a high proportion of readmissions were unavoidable in this sample, estimated overall to be 75%.

The 24/7 hotline service in the trial was used by 46% of the participants, and women used the telephone service significantly more often than men (*P*= 0.046). Callers were more frequently readmitted than non-callers in the intervention group, a significant finding (*P*= 0.001).

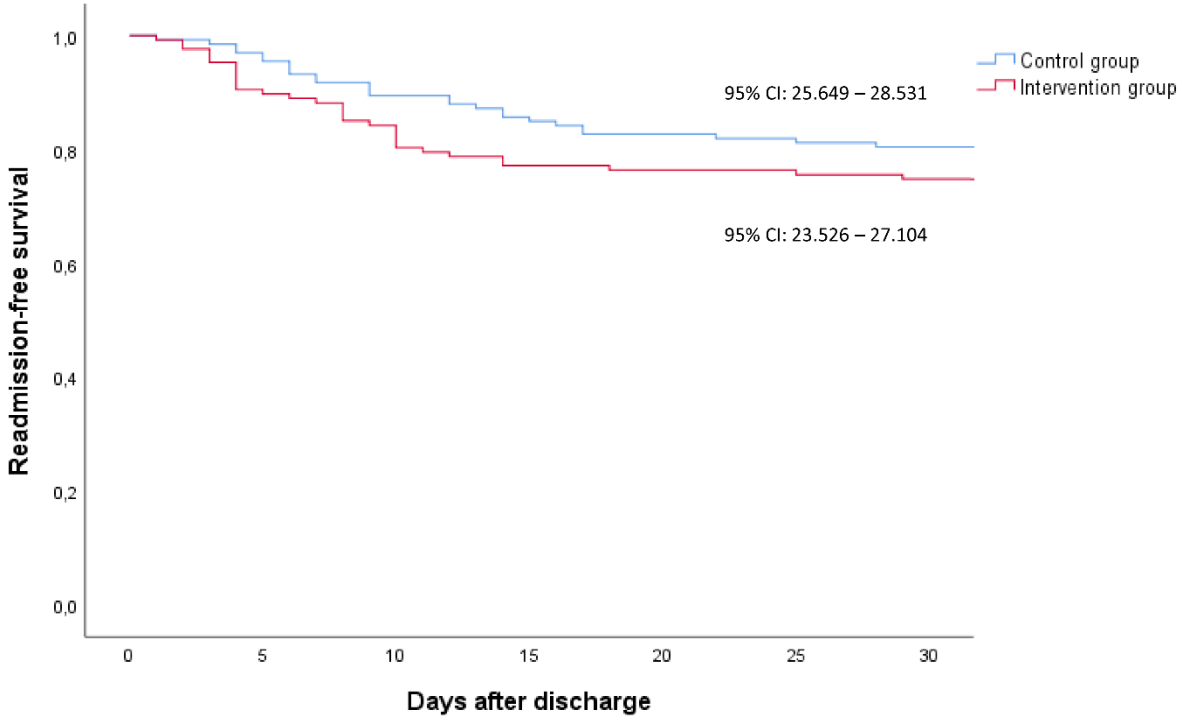


Figure 5. Kaplan-Meier survival curves demonstrating elapsed time to readmissions for the two groups in the AVRre trial.

7.4 Paper IV

Facilitators of and barriers to reducing thirty-day readmissions and improving patient-reported outcomes after surgical aortic valve replacement: A process evaluation of the AVRre trial

In *paper IV* [submitted], we extended the evaluation of the AVRre trial by conducting a process evaluation of the telephone intervention (Table 1). The findings revealed that SAVR participants were highly satisfied with the hotline service, and 91% perceived it as being a

trusted option. The TFU aspect of the intervention was also perceived as being trustworthy and valuable.

Our process evaluation found that a possible barrier to optimal implementation of the telephone support was that staff were insufficiently prepared in their training and education prior to the start of the trial. However, this is somewhat at odds with the prospective follow-up of the hotline staff conducted during the trial, in which the staff perceived it as being highly valuable and useful and that it facilitated the high-fidelity delivery of the intervention. Moreover, we found that the AVRre trial participants revealed that despite our efforts in conducting two telephone follow-up calls, they perceived a ‘gap in the care continuum’, ‘need for individualised care’, and ‘a need for easy access to health information’. We also found that discharge management of local hospitals had readmission rates from 0-50%, which affected the 30-DACR incidence.

The robust integration of user experiences into the AVRre trial produced a more complete picture of the impact of the intervention and the discharge and follow-up care after SAVR. This demonstrated the utility of a mixed-methods evaluation approach in a clinical trial with an RCT design, in addition to the direct outcome evaluation analyses.

8 Discussion

Hospital readmissions after SAVR incur high financial and emotional costs. Most SAVR patients in Norway are transferred to home from hospital soon after surgery, which could represent a risk for readmission. [29] A 2015 annual report for Norway reported that the hospital readmission rate was 15% for adults > 67 years. [18] The 30-DACR rate is often used as a quality indicator for hospital care performance [35], which might represent a valid proxy measurement for the quality of care after surgical treatment in a hospital. [34] Preventing readmissions is an obvious goal for clinicians, as it is for administrators managing readmission-related costs. Moreover, readmissions to hospital interrupt the expected care pathway and represent extra health and emotional burdens for patients. Thus, a primary aim of this doctoral thesis research was to test a post-discharge intervention that might reduce readmissions. Specifically, we sought to determine whether a remote intervention could reduce hospital readmissions after SAVR surgery and also improve patient-reported health and quality of life and symptoms that could affect them, such as anxiety. This aim was carried out in the context of the AVRre trial, an RCT conducted in a university hospital in Norway.

The AVRre trial used a mixed-method design to explore the discharge and follow-up of patients after SAVR treatments in the hospital to aid efforts in optimising the care. The aims were to provide a transparent protocol for the purposes of the AVRre trial and an answer to the overall incidence of the 30-DACR rate after SAVR and TAVR. Furthermore, we aimed to test whether a post-discharge telephone intervention would reduce the 30-DACR rate and improve patient-reported outcomes after SAVR. Finally, we aimed to provide a broader understanding of the clinical trial in the AVRre study in the context of the MRC process evaluation model.

An overview of the AVRre trial protocol was provided in paper I. In paper II, we reported through a systematic review of published studies that there is a high overall incidence

of 30-DACR after SAVR (17%) and TAVR (16%). To the best of our knowledge, the AVRre trial was the first to investigate the overall incidence of the 30-DACR rates after SAVR and TAVR using a systematic review and meta-analysis. The results of this review suggest that improvement is needed in most healthcare systems internationally to reduce the negative patient and financial impacts related to readmissions. Moreover, paper II revealed that new more rigorous prospective studies are needed that consistently report the 30-DACR rate.

In paper III, we reported results from the prospective AVRre trial, showing that the overall 30-DACR rate was 22.3%. Unfortunately, we found that the trial's telephone support intervention failed to significantly reduce the 30-DACR after SAVR. The intervention also failed to persistently improve patient-reported outcomes. Although symptoms of anxiety within one month after surgery did improve significantly, the improvement was not long-lasting, however, as the follow-up assessment one year after SAVR failed to show differences in anxiety symptoms. Symptoms of anxiety before surgery and pleural drainage before hospital discharge increase the risk of 30-DACR. We estimated the overall proportion of unavoidable readmissions to be 75%.

In paper IV, the SAVR participants reported being satisfied overall with the intervention, felt secure, and perceived the telephone support as being trustworthy. The intervention was implemented as planned, with the process evaluation providing evidence that the intervention was carried out with high fidelity. Although robust follow-up interviews with the AVRre staff during the trial favourably influenced implementation fidelity, more preparatory education and training might have further increased staff satisfaction with and fidelity to the intervention. The trial participants reported that the discharge was not optimal and could benefit if more follow-up was done during the transition of care for SAVR patients. The 30-DACR rate was found to be dependent on the context of local hospitals discharge management.

Overall, the research demonstrated that a mixed-methods approach is appropriate for this kind of clinical RCT. Moreover, it demonstrated that a process evaluation of the trial implementation and the impact of the intervention is useful for gaining a broader and deeper understanding of the results in this kind of clinical trial.

8.1 Methodological considerations

Although some positive results were observed, the AVRre trial had some limitations. Each paper will be considered in turn.

Study design and population

Paper I

The protocol for the AVRre trial was submitted for publication shortly before the trial was completed with no alterations, and outcome reporting was done according to the original plan (ClinicalTrials.gov Identifier: NCT02522663). There are several potential bias pitfalls regarding validity of the outcomes from clinical trials, in general, when the protocols are published in advance, such as unblinding biases, crossover biases, or bias related to the Hawthorne effect. [106] We considered these in the context of the AVRre trial.

Information on the internet and sharing of such information among trial participants is simple and fast through social media today. This situation presents challenges to the design of and interpretation of clinical trials, as more today than before, sharing information is so ubiquitous and easy. Concerns also might be raised because of the untimely publishing of the protocol paper. However, even though the 'protocol' manuscript was submitted while participant recruitment was still taking place, no data were entered to prevent ascertainment bias. [107] All research protocols for the AVRre study — available for scrutiny by readers of the 'protocol' paper — confirm that the outcome measures related to the efficacy of the intervention were not amended at any time. Protocol papers are intended to be prospectively

submitted to minimise scientific misconduct (e.g., reduce publication bias) and prevent selective publication and selective reporting of research outcomes. In other words, outcomes should only be a result of carrying out sound scientific practice.

We designed the AVRre trial in order to evaluate the telephone intervention. However, the trial could have been enhanced to focus data collection more specifically on other important factors (e.g., on local hospital discharge management systems). We may have underestimated the complexity of local hospitals' discharge practices in influencing the primary outcome, under-designing the trial with respect to confounding and/or interacting factors. However, there was limited theory on factors that affect the SAVR population and the 30-DACR outcome and on interventions that might reduce the 30-DACR rate. This lack might also have handicapped the refinement of our intervention according to the study population. Having highlighted these issues, present theory still provided a reasonable basis to attempt an intervention that might reduce the 30-DACR rate. Post-SAVR follow-up, including monitoring and managing symptoms early after discharge, reduces likely readmissions. [42]

Paper II

Conducting the systematic literature review and meta-analysis was the appropriate approach to accurately determine the overall proportion of the 30-DACR rate reported in the field after SAVR and TAVR. The review and analysis were performed in several steps to promote high-quality procedures and outcomes for the study. An early and vital step was performing a systematic review in which data collection was established by using a search strategy that efficiently searched the appropriate databases. Cooperation with a senior research librarian guaranteed that the strategy would systematically identify appropriate articles. Using this approach could be considered to be a validation step and a strength of this study. The number of published articles indexed by major literature databases has been increasing rapidly. [108] This fact motivated us to carefully consider which databases to search in order to do an

efficient, exhaustive, and accurate search. [109] We and others agree that a crucial initial step in a meta-analysis and systematic review, at least for less experienced researchers, is to cooperate with a senior research librarian. [110]

Another step in the review included the assessment of relevant candidate articles. This required spending sufficient time to ensure that all relevant articles were captured and systematically assessed. Since there is a risk of inadvertently overlooking relevant articles because of publication bias, [111] one cannot rule out the possibility that one or more articles were missed. However, we have no indication there was such bias in the present study, which is a strength of paper II.

Some reasons a relevant article may be overlooked is that there is no consensus on how an article should be titled in which the 30-DACR was the main objective of the study. This uncertainty can lead to identification mistakes (false negatives) when screening article titles and abstracts. However, our systematic procedure for scrutinising potential articles (two researchers involved was robust, helping us avoid duplicate inclusions. For example, by using the two-reviewer approach, we discovered that a candidate article's results were based on the same registry data recorded within the same time frame as data already included in another article. Therefore, this article was excluded from our analysis.

Another reason that the true 30-DACR reported in the literature might be underestimated is that completed studies that observed a higher proportion of readmissions in a single-centre observational cohort study or in an experimental study — one yielding a unfavourable statistical reduction in readmissions — are often less likely to be published in peer-reviewed journals. [112] In these cases, we would have had no opportunity to evaluate such articles for inclusion. Publication bias is present if the missing literature is systematically different from those we included. [111] However, we used several statistical tests (Funnel plot and Eggers test) for evaluation, and a sensitivity analysis confirmed that potential missing

publications were not systematically different from our sample. Therefore, we could conclude that the outcomes were less likely to be biased. As mentioned in paper II, we also did not assume a publication language bias was present due to the language limitations we used in the search strategy.

We also sent e-mails to researchers of publications asking if we could have a copy of the 30-DCAR rate data of the study population, and two authors failed to answer. We also considered whether overlooked 'grey literature' might also have introduced a publication bias leading to an inaccurate estimation of an effect size or a proportional size outcome. 'Grey literature' is defined as research *'that … is produced on all levels of government, academics, business and industry in print and electronic formats, but which is [sic] not controlled by commercial publishers'*. [113] Thus, these research results are produced but may not be published in the traditional commercial or academic publishing and distribution channels. There are several resources online for conducting a reasonable grey-literature search. [113] Even though we searched for relevant articles in the grey literature, and the tests for publication bias were negative, we cannot completely rule out the possibility that we might have missed relevant articles or data reports in the grey literature that could have pushed the true 30-DACR rate we observed in either direction (i.e., higher or lower).

An accurately conducted meta-analysis provides a cumulative analysis that can display patterns containing important insights for clinicians as well as for researchers. [111] The display or disclosure of patterns can generate new hypotheses and important clinical and methodological suggestions. The demonstration of patterns of proportions of readmission rates in a meta-analysis could therefore be viewed as a method to display more than a simple analysis. However, we used a meta-regression analysis in our study to investigate the relationship between readmission rate and study-level covariates. [111] Reporting the causes and risk factors for 30-DACR after SAVR and TAVR might provide a basis for tailoring new

interventions aimed at reducing the 30-DACR rate. It was interesting to note that in paper II, heart rhythm disturbances, heart failure, and infections are common complications after both kinds of treatments, a pattern that might suggest new clinical improvement projects to reduce readmissions after an invasive (conventional or mini-invasive valve replacement) heart valve procedure in the hospital. TAVR patients presently tend to be older and have higher number of co-morbidities. [12] This likely reflects a similar high proportion of early readmissions after TAVR, which is unexpected since the TAVR procedure is less invasive compared to conventional SAVR. [17]

Paper III

The experiment reported in paper III was an RCT, the 'gold standard' of experimental design when investigating the effects of a clinical intervention, in our case the post-discharge telephone support intervention. Using an RCT design is a strength of the AVRre trial. Why? One reason is that the randomisation of participant allocation to the treatment and control groups of an RCT generally prevents subject selection bias by distributing possible confounding variables fairly equally between the two experimental groups. Thus, statistical analyses of group performance can potentially detect a true intervention effect and not a spurious one that could be related to the presence of systematic confounders in one group or the other. [58]

In the AVRre trial, 96 candidate patients (20%) declined to participate or did not participate for other reasons. Potential participant self-selection might have led to selection bias in this study and, thus, threaten its internal validity. However, patients declined for different reasons, and we had no relevant data to analyse and compare this non-participating population with the participating ones on variables that may shed light on a potential self-selection bias. This was due mainly to the original approval conditions stipulated by the ethical committee regarding the scope of the AVRre trial; they did not give the AVRre trial

coordinators permission to obtain extra data without additional informed consent. The most common reasons for non-participation was that patients felt they would be too fatigued after surgery to participate in follow-up, or felt that for practical reasons, follow-up after discharge would be too burdensome (e.g., excessive traveling distance, inability to respond to the questionnaires in the required time frame). These kinds of factors increase the risk of selection bias being present. Having said this, however, we failed to find any systematic differences between the included study population and non-participants in terms of other kinds of self-reported treatments, age, or gender that would suggest our included study population was not representative of the general SAVR population. Thus, we can conclude, at least modestly, that we have some evidence that the AVRre trial was not harmed by selection bias.

The cumulative published research on interventions aimed at reducing the 30-DACR, in general (including a few mixed-cardiac surgery populations), have reported ambiguous positive effects of a reduction. It is noteworthy that results of intervention trials published after 2002 have reported less of an effect on reducing the 30-DACR rate than those published earlier. [44] Results from the AVRre trial, on the other hand, are the first to indicate which factors are predictors of 30-DACR after SAVR. Specifically, we reported in paper III that in-hospital pleural drainage and pre-surgery symptoms of anxiety are independent predictors of 30-DACR.

Based on the power calculation, we expected that the intervention would produce a 10% reduction in the 30-DACR rate between the two groups. It could be suggested that a 10% difference in a clinical trial might be too high and that the AVRre trial did not have sufficient statistical power to detect such a significant difference. Unfortunately, there were no previous studies on a SAVR population and 30-DACR rate after a TFU intervention that we could specifically use to inform us in calculating the sample size. However, studies using a TFU had reported a 10-30% reduction in the 30-DACR rates. [114, 115] If we had planned to conduct a

larger pilot study to guide our calculation of the needed sample size of the main trial, we might have been able to be more specific in our understanding on the trial's power to detect statistical significance between the groups. However, at the time, the main trial seemed to have sufficient power to detect a significant group difference for the primary outcome within the expected margin based on chi-square statistics. As discussed in paper III, the study lacked sufficient power to conduct a sub-group analysis on differences between the groups' proportions of unavoidable versus avoidable readmissions.

What other limitations and strengths can we note for the research reported in paper III? Blinding of group allocation can prevent bias in a RCT. [116] We did not design the AVRre trial to use 'true' blinding of group allocation in order to avoid performance bias. Having said that, the medical staff of the treating hospital and all participants were blinded until shortly before discharge from the university hospital. Participants were also encouraged to keep secret their group allocation and not openly display their information leaflet (section 6.2) until they were home. This request apparently was successful, since the hotline only received calls from intervention participants. Still, we cannot rule out the possibility that the group allocation was inadvertently revealed to discharging staff for some participants, which possibly could have influenced the standard discharge management (i.e., that was used for the control group). Also, the PC had updated knowledge about the context where the intervention was being conducted at the university hospital, but we detected no problem associated with the risk of information bias. However, the physicians and nurses involved in the discharge of the AVRre trial patients were aware of the study purpose and, thus, might have selectively increased their efforts to better inform the intervention group participants and caregivers at the time of discharge (i.e., Hawthorne effect). However, we have no direct evidence to support this speculation.

Are there other limitations and strengths that can be noted for the research reported in paper III? Performance bias can be introduced in a clinical trial in which several clinicians are determined to deliver one single intervention. [117] This can especially occur when the 'treatment' is based on the nurses' interpretation of the caller's message(s), or because of the communication skills of the caller or staff, previous professional experiences of the staff, and the actual skilled/unskilled use of the hotline manual. We cannot rule out the possibility that differences in how the individual AVRre-participating nurses delivered the hotline service might have impacted the fidelity of the intervention and the patients' expectations.

Another factor possibly affecting the delivery of the intervention was the presence of a learning effect over time manifest in the PC and the hotline staff. Thus, delivery of the intervention later in the trial might have improved, to some extent. In other words, intervention participants who took part earlier in the AVRre trial may have had a slightly different experience than those who took part later in the trial, because the staff would have gained mastery over the delivery over time and with practice. In the real-world clinical context, as our trial reflected, a certain level of individual differences in care performance are expected. However, we made efforts to minimise performance bias by ensuring information coherence in using a standardised manual for the delivered advice, conducting preparatory hotline staff training, recruiting a rather homogenous volunteer group of experienced ICU nurses in the field of cardiac surgery, and by providing robust and close support and supervision during the trial. We are confident that these precautions led to the delivery of the telephone intervention in as uniform a way as possible in a real-world context, minimising performance bias. We could have designed an intervention in which the participants received pre-recorded calls, although this approach poses its own problems. Still, the 'pre-recorded' approach might have given a more accurate understanding of whether performance bias related to intervention delivery was present.

In an effort to strengthen the AVRre trial, we chose not to severely constrain the age range requirement. Thus, we aimed to include adult patients electing SAVR who were > 18 years and not limit patient ages to > 65 years. This decision might have been even more appropriate with regards to statistical analyses due to the SAVR population's age composition (i.e., mean age 67, range 26-85 years). Prior trials with the 30-DACR rate as primary outcome have largely not included younger patients. [118] Not limiting participation in our intervention to patients older than 65 year enabled us to provide as complete a picture as possible about post-discharge care related to readmissions, without missing potentially clinically important knowledge on patients < 65 years. This is perceived as a strength. Combining the main sub-groups (see inclusion criteria, section 6.1) of SAVR patients into one group for analysis might have underestimated the intervention effect on the 30-DACR rate, if different sub-groups respond differently. More specifically, there was a risk of introducing selection bias due to potentially different clinical characteristics and mixes of co-morbidities across the sub-groups. Having said this, however, we did observe an equal distribution of important clinical variables between the sub-groups within each arm of the trial, and we observed a similar distribution of readmissions in the various sub-groups.

Paper IV

The MRC model was chosen to frame the AVRre trial, and it guided the extended process evaluation of the intervention reported in paper IV. This choice provided the opportunity to conduct a broader assessment of the outcomes and to identify other possible important findings related to the intervention and the participants' discharge experiences. Ideally, the process evaluation of the intervention should have been formally embedded from the start in the design using a defined model for its purpose. [53] Although the MRC-inspired process evaluation was not part of the original trial design, the evaluation was still mostly done prospectively, which strengthens this study.

The plan for data collection ensured that we had enough data to conduct the process evaluation in a valid manner. The retrospective interview with the hotline staff was conducted soon after trial completion. These interviews could also serve as a way to compare or check the field note results that were obtained during the intervention. On the other hand, by using a retrospective focus group, we might have inadvertently introduced confirmation bias, [119] in which prior interpretations from the field notes and the healthcare professional conversations during the entire trial possibly could have influenced the direction of the focus group. We were aware of this possible source of bias threat when planning the focus group, conducting it, and analysing the responses. However, employing two researchers to discuss emerging themes in the final stages of the qualitative analysis enhanced the trustworthiness of the results. Our research design using the planned broad data acquisition in a systematic and mostly prospective way robustly strengthens the study's validity and could also generate new hypotheses for further explorations and more tailored interventions. The design allowed us to conduct a thorough assessment of the strengths and limitations of the AVRre clinical trial enhanced by the transparency of the study (presented in paper I), together with the evaluations presented in paper III and IV.

The MRC framework emphasises the importance of doing a thorough developmental phase in planning for a clinical trial. [120] The developmental phase of the AVRre trial had several activities to facilitate the intervention definition and refinement, including conducting meetings with important personnel and presenting information within the university hospital, intensely developing the hotline manual, conducting pre-trial focus groups and interviews with former patients (user involvement in the planning phase), educating and training intervention staff, and conducting a pilot study. All these pre-main trial activities were recorded by the PC for evaluation purposes and strengthened the post-trial evaluation as informed by the MRC framework and presented in paper IV.

The main trial started with the hotline being served by the nurse staff in the cardiac ICU at the university hospital. However, after gaining experience within a few weeks, we recruited a group of volunteer nurses from the ICU to staff the hotline. The only change was that now fewer ICU nurses (8 in the group) would provide the same hotline service and give the planned intervention to the participants. We failed to analyse the ICU department's relationship with the intervention, which is a limitation. On the other hand, this trimming down of staff also strengthened the study in the sense that it facilitated achieving greater fidelity for the intervention by using a smaller, dedicated hotline staff. Some might perceive this change as a threat to the trial's validity, because we changed the logistics of intervention delivery. However, this staff change occurred early in the trial and in reality, did not affect the planned intervention in any way for the participants.

We conducted a one two-hour educational session and training opportunity for all the hotline staff to prepare for the intervention. Retrospectively, we concluded that this short session was a limitation of the study design. More education and training preparation might have improved the roll-out of the AVRre trial and ultimately might have improved the outcomes of the trial. In the follow-up of the hotline staff during the main trial, this suggestion is tempered by the fact that the follow-up led to a high-fidelity implementation of the intervention. Thus, we believe that the educational session and training was a true strength of the study. This is an important feature to consider when designing a similar future clinical trial. We also suggest that this aspect of the design likely saved time and effort prior to conducting the main trial. However, training must be carefully planned and evaluated to fit the actual study population and the healthcare system (context) in which the intervention unfolds.

The pilot study (N=10) was appropriate for its planned purpose and provided important information for refining the main trial and strengthened the overall study. One important input led to a change in how and when the participants were informed of their group

allocation. We found that a majority of the five pilot intervention participants had not opened the sealed envelope when calling them on day 2 after discharge from hospital. That led us to change how we informed the patients. Instead of giving them an envelope containing instructions, we directly relayed information face-to-face to the participant at the time of their discharge from the university hospital. This procedural change likely produced greater patient adherence in the follow-up of the study. Even though it was informative, conducting a larger pilot that would include testing, assessment, and refinement of the theoretical foundations of the intervention and how well the primary outcomes fit the intervention might have led to stronger positive outcomes. Information gained during a larger pilot study could also have contributed to a better sample-size calculation for the AVRre trial.

We might have underestimated the contextual influences on the intervention in the design, which limited our possibilities to analyse its impact on the outcomes. More attention on the influence of local hospitals, primary care (GPs), and caregivers could have been an area for even more elaborated explorations before and after the trial. However, we had already integrated validated questions related to some important features of contextual matters. These questions provided valuable knowledge about understanding the outcomes. The MRC model includes evaluation of contextual factors and has the potential to increase the value of clinical trials and reduce research waste. [53]

The value of integrating user experiences in the evaluation of the intervention was appropriate to achieve the aim of paper IV, a broader understanding of the ‘why’s’ in clinical trial outcomes, which is often warranted in medical research. [57] Applying a novel mixed-methods approach enhanced the extended evaluation purpose in this study, as presented in paper IV. Performing a longer longitudinal follow-up of the user experiences than what our study actually did might have bolstered the study.

Study outcomes

Paper II

In conducting our systematic review, we were guided by the considerations of Rao. [121] According to Rao, if researchers interpret that a set of papers to be reviewed contains too much clinical and methodological heterogeneity, then they might evaluate whether it is appropriate to conduct a meta-analysis. [121] Instead, they could consider doing only a systematic review. We determined that the pertinent studies we identified had little clinical and methodological heterogeneity; thus, we decided to perform a meta-analysis.

In paper II, we used I^2 statistics to determine the percentage of variance in a meta-analysis that is attributed to heterogeneity in the studies. [122] Our assessment using I^2 statistics revealed that there was a high degree of statistical heterogeneity (> 90%). Higgins and colleagues proposed using the following threshold I^2 values when quantifying the magnitude of heterogeneity: low, 25%; moderate, 50%; high, 75%. [123] We proceeded with sub-group and meta-regression analyses and were able to better understand potential causes of this heterogeneity; there was sufficient power (> 10 studies) to determine this. This in-depth analysis strengthened the study reported in paper II. Meta-regression was an appropriate analytical step for this purpose. If the I^2 ratio is large, then it is reasonable to analyse the heterogeneity further [111], which we did by using meta-regression in a random-effects model. Given that the I^2 is a measure of inconsistency across the findings of the included papers, the R^2 values from the regression analyses revealed how much of the heterogeneity could be explained by the study-level covariates. [111]

We used the NOS to evaluate the quality of the included articles. [77] A Scientific Statement From the American Heart Association (AHA) written by Rao and colleagues stated that there is no uniform agreement on how to evaluate the quality of different types of studies. [121] Since most cohort studies are observational, we chose NOS because it is widely used for

assessing cohort articles. [77] As we included different study types, we could have used other assessment tools in addition to NOS for assessing the observational cohort studies. Alternatively, we could have used the Grades of Recommendation, Assessment Development and Evaluation Working Group (GRADE), which is a promising tool for quality assessment of scientific evidence. [124] However, GRADE is an imperfect tool, and some write that it needs more empirical evidence to support using it for its intended purpose of making valid recommendations. [125] The AHA does not explicitly recommend specific tools to assess studies for meta-analyses. [121] However, one strategy that is acceptable to the AHA is using two researchers independently to carefully assess the weaknesses and strengths of the different included studies, instead of using NOS. [121] This approach revealed an interesting finding related to the quality of how validated hospital readmission statistics in individual studies are described. We were concerned about the lack of explicit/transparent statements in the included studies on how the readmission statistics were validated, which prompted us to assess the quality overall to be moderate to high. This finding mostly fits the retrospective cohort studies based on registry data that we included. This finding might also be considered to be a weakness of our meta-analysis, in addition to our warning about transparent descriptions on how thirty-day readmission data based on registry data are validated. This is also linked to our challenge of obtaining validated 30-DACR data from the NPR in Norway for analysing the primary outcomes of the AVRre trial reported in paper III.

Paper III

As described in paper I, we intended to extract readmission data from the NPR. In paper II, however, we found that very few studies done using registry data had transparent statements on how these data were validated. This lack changed our plans somewhat. Before paper III was published and prior to the AVRre main trial initiation, we purchased from the NPR,

anonymised historical data on the Norwegian 30-DACR rate after SAVR. Our goal was to verify that the primary outcome (30-DACR rate) data could be collected from the NPR with a high degree of validity. However, the NPR data was not considered to be a valid measurement of the 30-DACR after SAVR for our study population.

We measured the degree of co-morbidities by using medical diagnoses from participants' medical charts to calculate the Charlson Comorbidity Index (CCI) score, [126] an often used and valid method for pre-risk evaluation. [127-129] We decided not to include SCQ-16 scores, because many parts of the participants' questionnaire were incomplete or only partially completed and because its questions were answered inconsistently (i.e., there were many systematic errors). Indeed, participants reported that they were uncertain on how to fill out the SCQ-16. An informal comparison between the diagnoses in the medical charts and the answers on the self-reported SCQ-16 confirmed that including SCQ-16 scores likely would have introduced information bias. The risk of information bias via the SCQ has been reported previously. [130] Medical chart and other information collected at baseline provided an accurate description of baseline demographic and clinical variables for pre-risk scoring and other assessments; these data strengthen the AVRre study.

Intention-to-treat analysis (ITT) analysis is a more conservative approach for analysing the effects of a clinical trial, but it can lead to an underestimate of the intervention effect and also yield a type II error. [131] By contrast, per-protocol (PP) analysis can lead to an overestimation of the effect and yield a type I error. [132] A type II error will, in the worst case, postpone an effective intervention, which is less harmful than exposing patients to an unnecessary and potentially harmful intervention by committing a type I error. In line with CONSORT recommendations, we performed both ITT and PP analyses, which yielded similar negative results on the primary outcome and confirmed that our intervention likely did not act to reduce readmissions after SAVR. For clinical purposes, we chose PP as the main way to

analyse the primary outcome and to determine the actual efficacy of the intervention. However, the preferred method is ITT, because with this method, losing statistical power due to reduced sample size is avoided and equal distribution of confounding variables is maintained between the groups. Thus, this approach avoided analysing a biased dataset. [132]

Not all randomly allocated participants (N=288) underwent SAVR. Thus, we chose PP to analyse data from participants who took part in the entire intervention (N=260) in order to estimate the effect of the 24/7 telephone hotline on 30-DACR rates. However, for sensitivity analyses, we conducted ITT analysis on the (N=282) participants in the cohort who underwent SAVR treatment. This enabled us to assess the effect of the assigned treatment, as the dataset for these participants was complete for the primary outcome. Performing both types of analyses is recommended, especially in cases where some data is missing due to lack of adherence to the protocol and/or due to loss to follow-up. [133] As the dataset related to the secondary outcomes had some missing data, we conducted MI for the PP and ITT analyses.

In the PP analysis, approximately 10% of the EQ-5D-3L scores and 6% of the HADS scores had missing data. For the secondary outcomes, total missing data was 12% (33/282) at T1; 15.5% (43/282) at T2; 13.5% (38/282) at T3; and 18% (51/282) at T4. In addition to the missing data that was > 10%, we found that at each measurement point, the control group had more missing data than the intervention group. We chose, therefore, to conduct MI to replace the missing data.

LMM analysis for the repeated measurements up to one year after surgery was chosen as the most appropriate method to measure the longitudinal results from the HADS and EQ-5D-3L questionnaires. Repeated measurements (in longitudinal studies) for each case are often dependent, but this can vary among cases. To mitigate this problem in statistical testing, without committing a type I error, we used a random intercept model in the LMM analysis to handle the heterogeneity in clusters of the data. [98] This approach avoids a

complete case analysis, in which single cases are deleted entirely if one or more of the measurements are missing. Of course, using this latter approach can threaten the needed sample size and statistical power, leading to a higher risk of having biased results. [98] LMM handles missing values better than analysis of variance (ANOVA) with repeated measurements, for example, which requires wholesale deletion of a case if any data in those cases are missing. [98] An often-used method for replacing missing values is to use the samples' grand mean value to perform single imputations, or to use the predictive distribution of each case having missing data. [100, 134] However, this method can lead to a flawed estimation of the variance and a biased result. [100] We, therefore, used MI with 20 iterations [100] to obtain a pooled estimate for missing data.

The total missing values (N=282) ranged from 12 to 18% from T1 to T4 assessment times. The missing value analysis suggested (produced by an SPSS 25.0 routine), together with our clinical evaluation, that we could validly conduct MI under the missing at random (MAR) assumption. For sensitivity purposes, we compared the MI results with those obtained by the single imputation method, in which we replaced the missing values of a given case with the samples' mean distribution value. This approach yielded similar *P* values but with larger standard errors, supporting our choice to use MI for our analyses.

The reliability of the scales (HADS and EQ-5D-3L) we used was acceptable, as determined statistically. The internal consistency of the HADS questionnaires was accessed by Cronbach's alpha and was good. [83] The Cronbach's alpha for HADS-A was 0.8 and that for HADS-D was 0.79. We chose to not use Cronbach's alpha to assess the internal consistency of the EQ-5D questionnaires, because of its inability to measure the scales' quality. [135] However, we measured the correlation between the EQ-5D-3L VAS scores and EQ-5D-3L index value (VAS UK set) scores, which yielded significant Pearson correlations ranging from 0.58 to 0.64. These values indicated a high degree of consistency between the

two health status scores within the questionnaire. A scale that measures PROMs might be less reliable, because it lacks the ability to fully capture extreme responses (i.e., those in the upper or lower part of the scale), yielding a ceiling effect. [136]

A ceiling effect was present in the EQ-5D-3L measurement of the UK index value score (VAS or TTO based); with this scale, a health state of 1 is the best imaginable health state. The ceiling effect was present three months after surgery (Table 4). Interestingly, EQ-VAS scores demonstrated no ceiling effect at any assessment during follow-up. One reason for these differences is that this scale might measure different qualities of the perceived health state (e.g., EQ-VAS might measure perceptions of 'overall health'), as suggested previously. [137] Therefore, EQ-5D VAS and EQ-5D index value cannot be compared as equal entities. The presence of a ceiling effect indicates that the index value scores of the scale (VAS based) scale might be less sensitive in capturing the full extent of any positive effect of our intervention longitudinally. Similar findings between EQ-5D VAS and EQ-5D index scores and ceiling effect have been observed previously. [138] The cut-off range for the highest scores was set to 97-100 for the EQ-5D VAS score and 1.000 for the index value score; a ceiling effect is present if more than 15% of the responses fall within that cut-off range. [139]

Table 4. Distribution of highest scores of the intervention group for the EQ-5D-3L VAS and EQ-5D-3L index values over time.

Elapsed time after surgery	EQ-5D-3L VAS	EQ-5D-3L UK index value (VAS based)	EQ-5D-3L UK index value (TTO based)
	Scale score from 0 to 100 mm	Scale score from 0.073 to 1	Scale score from 0.594 to 1
1 month (%)	2.1	9.5	9.2
3 months (%)	7.2	40.8	41.1
6 months (%)	8.2	40.5	40.5
1 year (%)	11.7	44.3	44.8

The ancillary analysis of the proportion of unavoidable and avoidable readmissions in the trial was conducted independently by three clinical researchers in the project group.

They were blinded to the individual case group allocation in the trial. Such analysis was subjective, even if only the diagnosis codes were used for assessing the proportions of unavoidable readmissions. Since the assessment was subjective, we chose not to discuss the cases that the evaluators disagreed and failed to reach consensus on. This kind of procedure can represent the real-world context in which local physicians and routine practices may or may not treat similar cases differently on readmission. The proportion of unavoidable readmissions in the AVRre trial reached at least 75% and might even be higher for this SAVR population. Such an analysis might have approached validity more closely if we had used external evaluators. However, we rationalised that this analysis needed to be carried out by those having knowledge of local healthcare systems, which we feel justified the use of internal evaluators.

Doing a CPH regression analysis was the most appropriate choice for our assessment of predictors of 30-DACR after SAVR, especially because of time. The first step was to assess the data population assumptions for conducting a CPH analysis; these were adequately met. The assumptions are as follows: independency of observations, no multicollinearity (low variance inflation factor), no interaction effects, and a constant hazard ratio (HR) across time for the individuals (the proportional hazards assumption). The latter can also be evaluated statistically with SPSS 25.0. The next step was to complete a univariate analysis of chosen covariates of interest. Covariates with a P value < 0.2 were selected for multivariate modeling of possible independent predictors of 30-DACR after SAVR. The chosen covariates were based on a clinical and theoretical assessment of the available data for the prediction model. We could have also used multivariate logistic regression analysis and odds ratios to identify the predictors of 30-DACR after SAVR instead of CPH with HRs, as both methods are measures that evaluate relative risk. [96] Because of proportionality in the CPH modelling, we could not determine whether the association with the covariates was real; rather, it was the

best approximation. [96] The final model in our CPH regression analysis did not overfit the number of predictors — i.e., describing the random error in the data rather than the relationships among variables — and the validation was accurate for the intermediate steps to reach the final model.

Paper IV

One strength of the present thesis work was the nature of the data collected and analysed. The amount and sources of data (i.e., from several different sources) enabled us to compare the various results of different aspects of the AVRre trial, which strengthened the validity of the process evaluation. Results from the focus group were corroborated the results from the field note analysis, even though they were different kinds of data; this nature of the data strengthened the validity of the interpretations of the analyses and findings presented in paper IV. The design of the study and parallel evaluation during the trial prevented confirmation bias. While the main trial was being conducted, the PC assessed in parallel probable interpretations of the field notes and observations during multiple conversations and discussions with the research group.

The parallel interplay between data collection, measurements, and theory-based interim analyses strengthened the trustworthiness of the final qualitative results, and these procedures were in line with recommended qualitative analysis approaches. [87, 88, 105] Although different tools are available for assessing the quality of the qualitative methodology used here, consensus is lacking on which tools are the most appropriate for each situation. [140] We mainly followed the recommendations of Malterud when we conducted the qualitative analyses for paper IV. [141] It is important to note that the PC was aware of the possibility for confirmation bias. This kind of bias can be introduced inadvertently early in a study and can potentially impact the remainder of the study. The presence of confirmation

bias can be challenging to rule out completely due to the subjective nature of qualitative analyses. However, from the beginning in the planning phase, we took precautions to prevent confirmation bias by being aware of it and taking appropriate action.

The risk of confirmation bias was mitigated, in part, by the high participant response rate. All participants who used the hotline phone service answered the questions about their experiences with the hotline, and 84% of the entire study population (same rate in both intervention and control groups) answered questions about their hospital discharge experiences. The high response rate increased our confidence in the validity of the AVRre trial.

For PREM questions related to the discharge, Cronbach's alpha was 0.74, which was an acceptable value for the scale's reliability. Since the results were similar to the findings of the Norwegian national survey of hospital discharge experiences, this increased the generalisability of our present results. While 58% of the hotline callers rated their use of the service to be highly satisfactory, another 20% of the hotline callers chose not to rate the service, instead answering the question as 'not applicable'. It is possible that some in this latter group of hotline callers might be participants who were dissatisfied with the service but declined to rate it negatively, because they viewed the intervention overall to be positive and have potential for being useful during the early rehabilitation phase. We found, however, no systematic associations between the participants and 'not applicable' responses. No other available data suggested reasons that might explain why 20% of the hotline callers chose not to rate the service. Another possible reason for declining to rate the service is simple negligence.

The PREM questions we used in the AVRre trial were based on similar questions used in a Norwegian national survey on patient experiences about hospital stay and discharge; this survey content was tested and found to be valid for its purpose. [142] The Norwegian

national survey was administered one month after hospital discharge in the sample population. By contrast, our questionnaire on discharge experiences was administered three months after surgery, a delay that potentially could have increased the risk of recall bias. [143] In our study, however, the mean hospital LOS was 10 days, with some participants being hospitalised for 2 or 3 weeks before being discharged. Because of these hospital stays, we are confident that the time when the questionnaire was administered (i.e., three months after surgery) was appropriate and was subject to minimal recall bias. Thus, we are confident that the results were valid.

The AVRre research group and PC continuously discussed intervention-related processes (case-related processes) among themselves and with the hotline staff. In addition, the hotline staff participated in educational sessions and consultations. This resulted in a robust setup for the follow-up of the intervention and greatly strengthened the evaluations. The PC essentially conducted active field work using a semi-structured approach, in which both deductive and inductive approaches were applied for evaluation purposes. This mixed-methods prospective approach enabled us to increase opportunities to observe, record, and interpret the outcomes, thereby ensuring valid results. However, applying methods that are highly structured and less inductive, as in this case, can inadvertently cause one to overlook potentially meaningful data, which can threaten the validity of the results. [87] To avoid this possibility, we took steps to ensure that the mixed-methods approach was balanced throughout the intervention, and we actively sought to prevent confirmation and performance bias, and the Hawthorne effect in the university hospital context. The novelty of the study design and outcome measurements were carefully handled and appropriately managed to yield valid and trustworthy results, which strengthened the interpretations of results presented in paper IV.

Study intervention

Paper III

In the AVRre intervention, we employed a novel approach for following up newly discharged SAVR patients by telephone. The telephone service in this intervention comprised two elements — standard TFU plus a 24/7 hotline — for a higher quality of follow-up in the early rehabilitation phase after hospital discharge. TFUs conducted in a timely fashion with educational and practical advice for using the TFU and for managing and monitoring symptoms have reduced readmissions in other clinical settings. [42] Making available the 24/7 hotline service during the initial month after hospital discharge after SAVR, in addition to the standard discharge care (TFU), empowered intervention participants to obtain advice whenever they wanted and thereby increased their level of self-care management. We hypothesised that this kind of follow-up system would reduce the 30-DACR rate after SAVR. However, our data indicated otherwise.

There are several possible reasons why an intervention effect was not observed. Firstly, the planned, and ultimately administered, TFU dose might have been too little (two calls). Perhaps we should have done more TFUs. More follow-up calls might have positively pushed participants at higher risk for readmission to solve simple health concerns. However, most readmissions took place within 14 days (83%), and nearly half of them occurred within one week of hospital discharge. This suggests that more calls should have been made within the first 14 days after discharge.

Secondly, as discussed in paper III, the increased attention to their health condition and symptoms after discharge could have contributed to the increased readmissions in the intervention group. We also noticed a quite high rate of readmissions among control group participants that was not significantly different from that of intervention group. A UK study found a paradoxical result: there were significantly more readmissions among intervention

participants who received home visits from a pharmacist who discussed with them several side effects of the patients' medication. [144] The authors suggested that more attention towards their problems, through the intervention, might have caused the paradox result.

We also performed ancillary analyses on the proportion of unavoidable readmissions and found that slightly more readmissions were deemed unavoidable for the intervention group (81% vs. 69%) compared to the control group. Even though the experienced clinicians in the research group assessed whether a readmission was unavoidable and even though they were blinded to group allocation, their decision to label a readmission as 'unavoidable' is subjective and not completely without bias. This could have contributed, in part, to these large unavoidable readmission percentages. With this in mind, these readmission numbers should be interpreted as rather crude numbers. We recognise that some of the readmissions the experts disagreed on could also have been labelled as 'unavoidable'. This means that the published statistics might represent a conservative estimate of unavoidable readmissions of patients after SAVR.

As alluded to above, the judgements of the clinicians are subjective, being moulded by their experience, professional views, and knowledge about the healthcare system. They are also sensitive to local differences in handling complications. For example, in one hospital a clinician may treat arrhythmias through out-patient consultations, whereas, in another hospital, a clinician would treat arrhythmias by hospitalising the patients. Nonetheless, we still conclude that a clear majority of the readmissions were unavoidable and that this likely affected the intervention's ability to reduce the total number of readmissions. Likely by chance there were slightly more readmissions in the intervention group than in the control group; this made it harder to detect an intervention effect. As we stated in paper III, this trial lacked sufficient statistical power to detect between-group differences in avoidable readmissions. We had no prior information about this possibility. So, this knowledge can now

be used to form new hypotheses and for power calculations in future studies aiming to reduce the 30-DACR rate after SAVR.

Paper IV

The nurses who staffed the intervention hotline had various skill levels of communication ability and styles of communication, and past experiences. This diversity of abilities increases the risk of performance bias when delivering the intervention. The risk of performance bias was reduced by close and continuous follow-up during the trial that included educational sessions, case discussions, and consultations provided together with the hotline staff.

Also, a learning effect for presenting trial-specific information to callers was likely present, and especially since the intervention had a longitudinal design. However, this was handled prospectively with planned follow-up during the trial. Thus, the response from the intervention participants, specifically, information in the questionnaire and the field notes, confirmed that the uptake of the intervention was good, which strengthens the study. Bolstering our evidence that a learning effect was minimised is seen in responses of participants in the control group. Several participants in the control group used the opportunity in the 3-month and the 1-year follow-up questionnaires to express their views (in written narratives) on the discharge care they received during the trial. Their narratives underlined the findings that we independently observed that there was a gap in the care continuum for the SAVR study population. They believed that the intervention was a reasonable offer to minimise the effects of a care gap, especially for the early rehabilitation phase after hospital discharge.

The intervention was delivered from the university hospital by experienced ICU nurses with several robust tools designed for the AVRre trial (the 24/7 hotline manual, the monitoring PC, experienced physicians on duty, own experiences with cardiac-surgery

patients) at their disposal when replying to participants' post-discharge calls. This collection of tools essentially provided the intervention-group participants with a fast and direct route to secure health advice. Approximately one of ten (9%) participants were transferred directly to a rehabilitation facility before transferring to home, and for these participants could have 'short circuited' their participation to some extent; this situation was one reason given by participants for why they chose not to use the hotline. However, the rehabilitation facilities did not have a physician present on duty at any time, and the nurse staffing at the rehabilitation facility is limited during evenings and nights. Having said this, some participants did call the hotline from the rehabilitation centre, as we instructed them to do freely, if needed.

Mostly, participants received the intervention in their specific home context, which could have differentially influenced their degree of adherence to the intervention and might have contributed to a lower external validity in the AVRre trial. We sent an alert SMS prior to doing the actual TFU calls, which allowed the participants to prepare for the call at a chosen time. This procedure might have facilitated the outcomes of the TFU for the participants and contributed to their high expressed satisfaction with the intervention. The field notes revealed that several participants had prepared questions that they were eager to ask when the PC called. The latter suggests, together with the present evidence from the study, that our intervention was truly clinically important for the participants, an important point that strengthens the rationale for conducting the intervention within our context.

External validity of the study

External validity in clinical research and experimental design refers to what degree the results from a clinical trial or experiment are generalizable to a given patient population in different contexts. [116] Are the results credible proxies for SAVR patients in the real world,

representing the true relationship between observations in a study and what the situation is in the real world for SAVR patients? Can we draw credible inferences from the results of the intervention to the observed effects? [116] These are important questions to answer in order to assess the external validity of a study, in our case, the AVRre trial.

We cannot rule out a possible patient selection bias in the AVRre trial, even though we accepted our study cohort as being representative of the elective SAVR patient population in Norway. Given our clinical experience and history at the university hospital, it seems reasonable that the study population was indeed representative. Having met this assumption, we felt justified in using comparative statistical testing.

Several factors increased the external validity in our study. Firstly, the study sampled a broad swath of the SAVR patient population in a university hospital, which is responsible for half of the national population. Secondly, the intervention was not too difficult for the participants to comply with. Thirdly, the gender distribution of participants in the study reflected the real-world gender composition of patients undergoing elective SAVR treatment. However, our statement of external validity should still be cautiously considered, knowing that even with an RCT design, it is challenging to draw true scientific inferences from the results of a study population sampled from the whole population. [145] This is especially the case if the theoretical basis of doing a study is not integrated into the study so that the 'why it worked' can be reached. In addition, it can be difficult with RCTs to determine statistical significance of the effectiveness of a "treatment". [145] As we discussed in paper III, the AVRre trial was a single-centre study in a specific context, which can decrease the generalizability of the study. To be able to make a reasonable critical appraisal of the external validity of a study, it is necessary to accurately describe the methodology used.

Several points strengthen the validity of the trial: the results from paper IV show that the intervention was conducted with high fidelity, the intervention dose that was delivered

occurred as planned, and participants were highly satisfied with the intervention. The high satisfaction of both components in the intervention also increase the external validity. However, as discussed, the lack of absolute adherence to the 'treatment' may be a threat to robust validity. One reason suggests that a slightly higher proportion of unavoidable readmissions occurred in the intervention arm, and another might have been that the usual care was not delivered as uniformly as we had assumed.

Usual care in the AVRre trial was conducted by several nurses and physicians, at different times and locations. We found that local discharge managements differentially affected the readmission proportions, which might also have impacted other factors related to increasing or decreasing readmission within 30 days after discharge. The AVRre trial might have underestimated the complexity of the usual care and might have been able to investigate the local influences more in-depth if we had planned for this aspect and obtained data accordingly. We cannot rule out the possibility of the Hawthorne effect due to the staff being aware of the interventions' intentions. However, we did not register any substantial changes during the intervention period. There was a raised awareness, though, towards the information process in the discharge at the university hospital, but no real changes were implemented that might have altered the discharge care. The participating hospitals are adhering to the ESC guidelines for the treatment of the heart valve patients. [5]

Transferability is an often used term in qualitative research, which can be regarded as the equivalent term of generalizability in quantitative research. [141] The appraisal of whether the sample was adequate and sufficiently varied can account for a higher or lower transferability of the qualitative results. We suggest that the qualitative findings are transferable from the study population to the SAVR population in general, given the same study context. Obviously however, the qualitative findings cannot be used to support generalizability based upon statistical evaluations, as often as it is with the use of quantitative

methodology. The research design, including a process evaluation guided by the MRC model, was able to generate knowledge and information to elaborate on the assessment of the external validity. This resulted in an increased external validity of the trial.

8.2 General discussion

8.2.1 Discussion of prospective protocol for the AVRre trial

We adhered to SPIRIT [70] and CONSORT guidelines, [78] specifications designed to prevent poor and/or irregular reporting of the methodology and the outcomes of clinical trials. A promising initiative to link and consolidate the outcome reporting of these two guidelines is the Planned Endpoints in Clinical Trials (InsPECT), [146] recommendations developed in accordance with the Enhancing the Quality and Transparency of Health Research Quality EQUATOR Network. [147] The reporting from clinical trials have improved over the years, but there are still quality differences between high- and low-impact factor journals, where lower-impact factor journals are more likely to publish trials that are less non-transparent and with a higher risk of bias. [146]

A review published in 2017 determined that 58% of the reported trials had unclear reporting of allocation concealment, which is a crucial aspect for being able to critically appraise an RCT. [148] Moreover, the so-called replication, or reproducibility, crisis in science [149] advocates more initiatives to achieve a higher standard of the reported outcomes. [150] Our protocol of the AVRre trial, together with the reporting in paper III and IV, produced a clear ‘picture’ of what we did, allowing for replication of at least the core elements of the intervention and facilitating a pragmatic replication of the intervention in another real-world context. That being said, we could have improved aspects of trial transparency (e.g., the descriptions of the contexts in which the intervention was conducted); doing this would have facilitated the replicability of the trial protocol reported in paper I.

Furthermore, we might have provided an even more detailed description of the theoretical background in paper I to meet the scientific critique of the problem basis of less commonly reporting of clinical trials. [150] However, the ambiguous and inconsistently reported evidence of readmission reduction effects of prior interventions justified doing the AVRre trial. Moreover, the novel and promising design of the effects of symptom monitoring and educational inputs after the discharge justified doing the trial. The participants' satisfaction and the documented burden within a month after surgery (e.g., symptoms of depression increased from baseline compared to one month after surgery) provided strong evidence of the scientific value of testing the intervention, as the protocol reported in paper I.

8.2.2 Discussion of overall incidence of 30-DACRs after SAVR and TAVR

We found a high and similar 30-DACR rate after SAVR and TAVR reported in paper II (17% and 16%, respectively). The 30-DACR rates were higher than the pooled results reported for the USA and Scandinavia (Denmark and Norway), and lower in other countries. Many of the large studies from the USA were retrospective, multicentre, and registry studies (mostly sourced from administrative databases). Registries might capture more readmissions if they can track patients across hospitals and ensure a high degree of data completeness compared to single-centre studies, which might register readmissions only to their own hospital. On the other hand, registry studies might have a larger error due to for example, incorrect coding, [151] which can lead to an overestimated incidence of the 30-DACR rate.

We found that when working with the readmission rate after SAVR in Norway, coding practice of hospitals and method of extracting the 30-DACRs are flawed for purposes of scientific analyses. In paper II, we concluded that studies using data from registries should begin to provide more transparent statements on how they arrive at their readmission statistics, a deficiency that has been highlighted before by van Walraven and Austin. [152]

However, using a prospective follow-up design and complete and clear definition of the 30-DACR rate, the AVRre trial revealed an overall rate of 22.3%.

Another issue surrounding discharge practices was reported in paper IV. Therein, we reported that local hospitals in Norway have different discharge management systems, which might lead to significant differences in calculated readmission rates. A recent study of Heggestad revealed that shorter hospital LOSs lead to an increase in readmissions. [29] Curtailing the LOS for financial reasons should not, for ethical reasons, be advocated if harm to patients is the result. However, in Norway, which does not penalise hospitals for readmissions and where hospitals' finances are partially driven by achieving diagnosis-related group points, a readmission will economically be positive if the patients are not hindering other patients' stays.

The United States healthcare system, in which economic penalties are enforced when the 30-DACR rate is above an expected risk-adjusted level, is probably less likely to be interested in funding more research on readmission outcomes. The recent literature suggests that 30-DACR rates are trending downward in surgical populations in the USA, with an accelerating decline after introducing the HRRP. [153] The findings of declining readmissions have also sparked a debate among researchers whether the HRRP has had the unintended consequence of producing higher mortality rates among heart failure patients. [154, 155] These considerations demonstrate the methodological challenges facing efforts to improve readmission outcomes and the use of 30-DACR rates as a quality indicator. Further, they may also speak to why there are differences among countries on readmission outcomes. It is known that readmission rates differ among countries, and it has been proposed that a search should begin for answers on how healthcare may be mismanaged, specifically, the hospital LOS or the aftercare by the GPs, for example. [156] Recent literature also reveals an interesting finding in which high-ranking hospitals in the USA have lower mortality rates and

higher patient satisfaction in cardiovascular care compared to lower-ranked hospitals. [157] However, the readmission rates are similar between the high- and low-ranked hospitals [157], which speaks to the critique of and challenges of using the readmission outcome as a quality indicator. [23, 158]

The accurate follow-up of the design and outcomes we analysed in the systematic review and meta-analysis yielded overall readmission-rate estimates, which are likely the best present evidence (within the confidence intervals) of the overall global 30-DACR rates after SAVR and TAVR. In paper III, we continued with a prospective RCT to test an intervention to reduce the 30-DACR rate after SAVR. The intervention failed to significantly reduce readmissions. However, at least it had no negative effect on mortality or other adverse side effects. Anecdotally, we observed some cases in which the telephone support intervention may have been lifesaving. [unpublished observations]

8.2.3 Discussion of 30-DACRs after SAVR and patient-reported outcomes

The meta-analysis of relevant studies reported in the literature demonstrated that the pooled 30-DACR rate for the SAVR population is 17%. In paper III, the AVRre trial yielded an overall 30-DACR rate of 22.3%. The few studies validly reporting the 30-DACR rate after SAVR demonstrate a similarly high rate, typically above 20%, in Scandinavia (Norway and Denmark), for example. [20, 21, 159] There are regional differences, as the meta-analysis demonstrated. Different economic drivers and healthcare systems across countries [156] might account for some of the difference between the AVRre trial's overall 30-DACR rate and the rates found in the meta-analysis.

An interesting finding in a meta-analysis reported in 2014 found that more recent interventions intended to reduce readmissions were significantly less effective than interventions conducted before 2002. [44] Recent research on TFU outcomes (led by nurses)

published after 2010 suggests that TFU does not reduce readmissions. [160] Our finding is in line with these results. However, it is important to discuss why later-tested interventions appear to be less effective in reducing readmissions. Leppin suggested some possible reasons, for example, a shift towards more technology-driven interventions in later years and less direct human contact. We suggest another possible reason: Increased quality of data collection across hospitals in capturing all relevant readmissions and improved data registration in recent years might lead to non-significant results. Thus, we speculate that in later attempts to reduce readmissions, more readmissions occurred in both usual-care and intervention cases, diluting any effect of the intervention.

Although we did not observe a significant reduction in 30-DACR rates as a result of the intervention, we did observe a significant reduction of anxiety symptoms up to one month after surgery and discharge in the intervention group (P= 0.031). Was this statistically significant difference clinically meaningful as well? Really, it needs to be considered from the patients' perspective.

The minimal clinical important difference (MCID) can be achieved by using distribution-based or anchor-based methods. [161] The partial Eta-squared score (a distribution-based score) in the General Linear Model (GLM) univariate (ANCOVA) analysis can be viewed as a correlation between an effect and the dependent variable. Eta and partial Eta describe the association related to the sample, while the Omega squared score estimates the association with the population and might be a stronger measure of the effect size. [162] However, distribution-based methods do not account for patients' perspectives of a meaningful difference. [161] The statistically significant difference we observed between the groups made clinical sense when analysing the findings of the survey on the hotline and the qualitative analyses as reported in paper IV. These findings must be preconditioned in order to state a clinically meaningful difference regarding symptoms of anxiety experienced by the

two groups. The conclusion would have been stronger if inclusion of the participants' perspectives had been designed to function as an anchor-based method together with the distribution-based methods, as recommended. [161] Thus, a cautious interpretation is required, and as we have reported in paper III, due to a small effect size (partial-eta square = 0.019) in the ANCOVA analysis.

Ancillary analyses provided several interesting findings. We found an interesting trend in which the youngest participants, those < 50 years old, had twice as many readmissions compared to patients > 50 years old. Recent research shows that patients < 65 years with concomitant chronic conditions might have more readmissions. [118] This finding warrants more research and clinical attention to be focussed across the age span of SAVR patients in order to determine the optimisation of the hospital discharge and follow-up needs to be titrated according to age. [118] It has already been demonstrated that advancing age increases the incidence of 30-DACR after SAVR. [163]. However, this is different from the findings of the AVRre trial, but is likely related to the observation that much of the research in this area of hospital readmissions has been conducted largely on older populations (> 65 years), at least in the USA. [118] This is appropriate in terms of research methodology and for statistical purposes when comparing different groups within the older segment of the patient population. However, we might not capture important information that is clinically important for the younger segment of the patient population to improve efforts to increase care quality.

Ancillary analyses provided other interesting findings. Firstly, we found that most of the readmissions in the intervention group could be considered unavoidable in the study context. More readmissions were cardiac related, which underlines the higher proportion of unavoidable readmissions we observed. This outcome hurt the likelihood of the telephone intervention reducing readmissions. Secondly, there was slightly more unavoidable

readmissions in the intervention versus the control group; however, this difference was not statistically significant.

We found in the study reported in paper II that there was little evidence about which risk factors specifically lead to 30-DACR after SAVAR. Therefore, we conducted a CPH analysis in that study. We found that patients who had symptoms of anxiety before surgery and/or were undergoing pleural drainage in hospital before discharge independently predicted 30-DACR after SAVR. This new and important finding implies that clinicians and researchers should attempt to improve the discharge and follow-up care of SAVR patients. This new knowledge can be used to individually tailor the discharge to take into account these factors, with the goal of preventing more readmissions.

Preoperative risk assessment, including testing for symptoms of anxiety, can be easily conducted. For patients undergoing pleural drainage, it might be determined that before discharge, they should be scheduled to get an outpatient consultation within one week of being discharged to home. With systematic cooperation of local hospitals, this simple change could reduce the 30-DACR rate. However, cost-benefit analyses need to be carefully undertaken to assess whether this will be cost-effective compared with the actual current discharge care procedures, since CPH modeling is a simplification of the real-world where the strengths of the statistical associations must be critically appraised.

Also, we found that the intervention reduced symptoms of anxiety within a month after the discharge, and combined with the knowledge that preoperative symptoms of anxiety predicts readmissions, indicates that the patients' anxiety state might be an important factor to address for healthcare professionals involved in SAVR discharge and follow-up care.

The monitoring and managing of participants' symptoms after hospital discharge during the AVRre trial prompted us to request two acute referrals to the university hospital because of a life-threatening cardiac tamponade. Both participants were < 60 years old, had

been treated with a mechanical valve, and were being treated with warfarin. A 1986 study conducted in Sweden found that surgical valve replacement (likely mechanical valves at that time) in combination with warfarin treatment was a common factor among patients diagnosed with late tamponade (median elapsed time to occurrence, day 8 after surgery; mean age, 53 years old, and total incidence of 1%). [164] A more recent study found that median day for occurrence of cardiac tamponade was day 17 after surgery. [165] In this study, the mean age was 58.5 years old; the total incidence was 4.3%. Furthermore, having a mechanical valve was an independent risk factor for tamponade. When the tamponade condition is drained in a timely fashion, the problem is solved and very few recurrences occur after the first drainage is completed. [165]

Solem and colleagues demonstrated that nausea and impaired well-being are early symptoms after SAVR, [164] and You and colleagues reported that the symptoms *'can be easily missed'* after discharge. [165] We found a total incidence rate of tamponade of 2.8% (8/282) within 30-days after discharge in the AVRre trial, whereas the incidence of in-hospital tamponade we observed was 4.3% (12/282). Moreover, the mean age for the population experiencing tamponade (N=8) within 30 days after discharge was 54 years, and 7 were males (88%). The symptom descriptions contained more common symptoms like reduced feelings of well-being, generalised chest pain, increasing dyspnoea, and coughing. In line with You et al., [165] specific knowledge and experience is required to interpret this condition as a possible pericardial effusion or tamponade.

The occurrence of more general symptoms in early stages of this complication after discharge requires experienced cardiovascular personnel to detect and diagnose the condition. The work of the experienced nurses staffing the hotline in the AVRre trial confirmed that they had the necessary requirements we needed for reliably delivering the hotline service. Two participants were acutely referred and readmitted with tamponade, requiring immediate

invasive drainage, which demonstrated a potential lifesaving benefit of the intervention. To illustrate one important challenge in the early rehabilitation phase, one patient in the AVRre trial was disallowed admission to a local hospital the evening before he was transferred to the university hospital for acute treatment of his tamponade. No fatalities resulted from the occurrence of the tamponade in the study population, which suggests that the present healthcare system works in favour of the patients experiencing tamponade. However, these patients need a final invasive solution sooner rather than later, which suggests that putting in place a bridge (hospital-home) intervention might be useful to potentially save lives. We noticed that men expressed more symptoms of dyspnoea in hotline calls after discharge compared to women

The theoretical foundation of how the intervention might reduce readmissions after SAVR needs further thought. Managing and monitoring symptoms after the discharge combined with educational input produce fewer readmissions after hospital discharge. [42] However, the expected self-preserving actions of the participants (hotline calls) are also based on the notion that patients will change their behaviour appropriately according to health advice given. In the AVRre trial, we could have emphasised important elements related to behavioural changes more in addition to the medical information provided by the intervention. For example, what strategies related to the discharge are best suited for promoting behavioural changes leading to enhanced self-efficacy and that prevent adverse medical events after SAVR? We might have underestimated the value of making clearer the connections between the components yielding increased self-efficacy and how this might unfold in the specific intervention we offered the SAVR patients. Medical personnel are trained to effectively provide information that promotes knowledge about healthy patient behaviour. However, they are often not trained in the use of specific strategies that might increase the likelihood that patients adhere optimally to a certain treatment or rehabilitation they are offered. [166]

We found that participants' symptoms of depression increased between baseline assessment (before surgery) and one month after surgery (supplemental material in paper III). The perceived health state, as measured by the EQ-5D UK index value score, decreased over the same assessment period (Figs. 6 and 7; supplemental material in paper III). This indicates that, for many of the participants in this study, they struggle to maintain a reasonable level of self-efficacy to achieve good health in the early rehabilitation period. This result has been reported earlier in other studies also. [167]

Fig. 6. Average participant total EQ-5D-3L index value score (UK VAS-based) over time in the AVRre trial

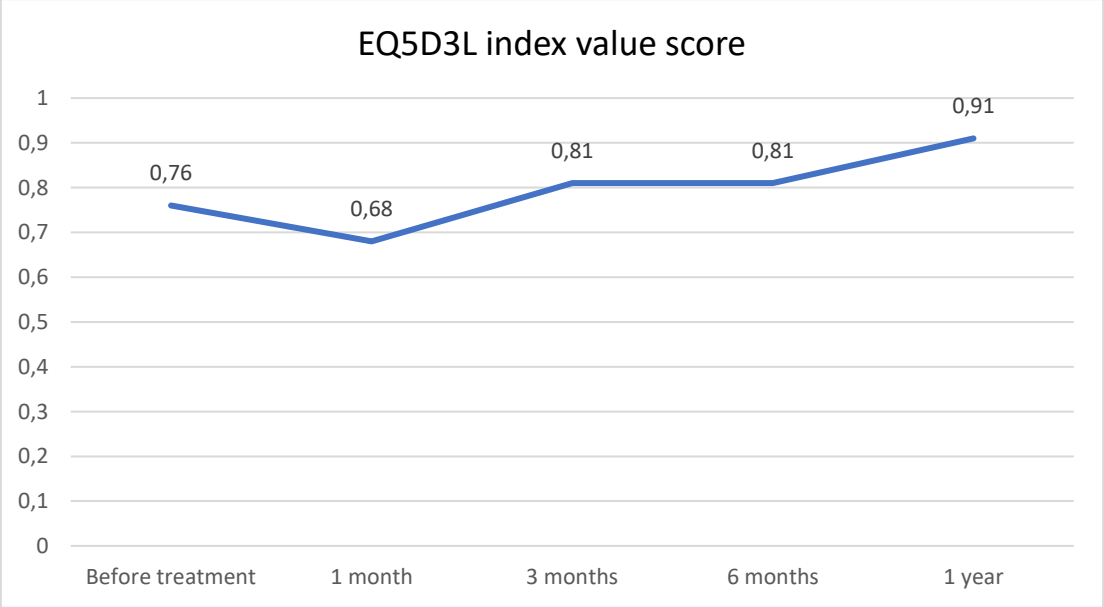
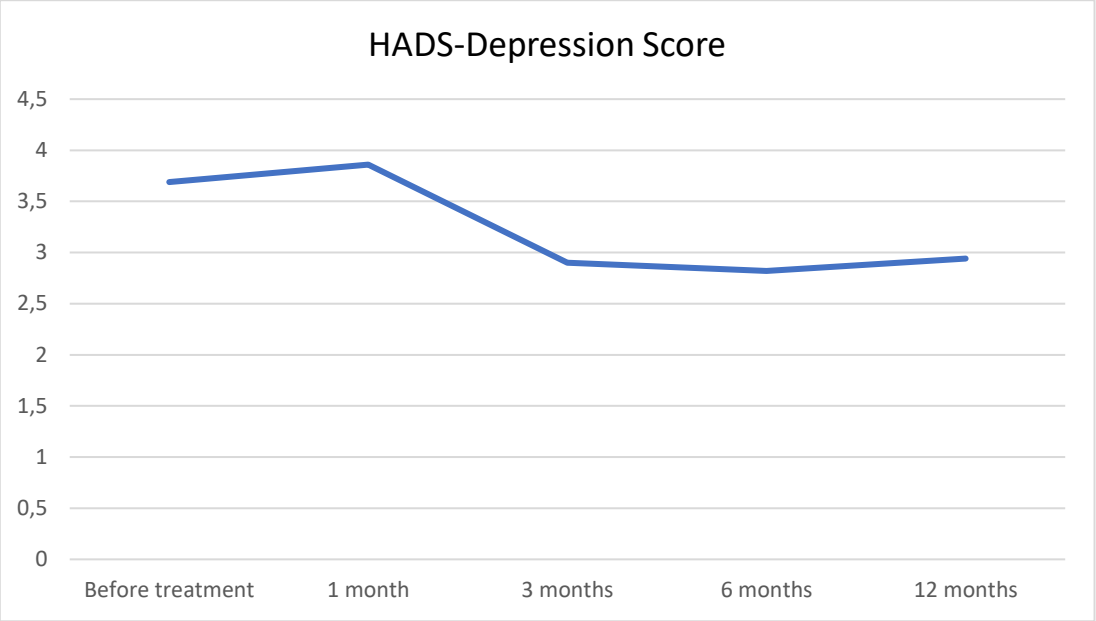


Fig. 7. Average participant total HADS score for depression over time in the AVRre trial



The findings presented in Figures 6 and 7 also suggest that our expectations regarding the outcomes were not adequately realised for several participants in the early rehabilitation phase; the qualitative findings reported in paper IV also support this idea. This result suggests that enhancing patient self-efficacy is both an important goal and a challenge for the SAVR population. Moreover, it suggests that a greater emphasis should be placed on using relevant strategies to encourage behavioural change. This, in turn, might increase adherence to the given treatment and aid reaching a health condition that could prevent new adverse medical events. On the other hand, participants of a RCT need to have a certain level of engagement and understanding in order to change their behaviour and reach healthier decisions. [168] This might have been difficult for many participants in the AVRre trial if they rehabilitated more slowly than they expected in the early phases after surgery.

Self-monitoring of behaviour, risk communication, and use of social support might be effective strategies for promoting behavioural changes by health care personnel. [169] Important determinants to produce behavioural change that promote health and prevent disease are, for example, increasing *knowledge* about health risks and benefits of healthy behaviour, possessing perceived capacity for *self-efficacy*, and *outcome expectations*. [170]

The latter determinant is based on Banduras' work on social cognitive theory, [171, 172] also commonly known as the theory of self-efficacy and its role in changing behaviour, as discussed by Sheer. [171, 172] The telephone intervention in the AVRre trial had elements designed to improve the participants' knowledge, their understanding of symptoms and risk assessments, and support their outcome expectations according to their individual needs. However, the trial might have benefitted more from a design that enhanced these core elements even more. Hence, we specifically emphasised how the elements could increase the participants' level of self-efficacy to positively influence the outcomes in the AVRre trial.

We targeted several aspects of what constitutes the participants' self-efficacy in order to support their healthy choices, which presumably would aid in reducing adverse medical events, including lowering readmissions. Participants received TFUs on two occasions within the first 14 days after discharge. These calls systematically related knowledge to the participants about the importance of engaging in physical activity to lower the probability of experiencing adverse events. Increased physical activity is associated with lower number of readmissions [173], and even reduces the risk of mortality after valve surgery. [174] To realise this aim, the hotline staff participated in an educational session with a specialist physiotherapist during the main trial that covered how to do this. They were also well prepared to always be aware of giving this important advice to intervention-group participants during the early rehabilitation phase after hospital discharge. Increased physical activity after cardiac surgery might also be associated with lower amounts of pleural effusion. [175]

We found that pleural drainage before discharge was a risk factor for readmission. The relationship between physical activity and incidence of pleural effusion in SAVR patients needs more attention from clinicians and needs to be studied more by researchers, especially in terms of optimising discharge and avoiding new medical adverse events. Moreover,

engaging in and adhering to CR should be emphasised more. [176] Our study participants, as reported in paper IV, stated that they often dismissed opportunities to take advantage of CR, because of excessive travel distances to a CR centre; this reason has also been reported to be a barrier to participating in CR. [177] We believe that the AVRre trial intervention provided the participants with sufficient knowledge about the importance of engaging in physical activity. However, we might have underestimated the extent to which we should have emphasised other components necessary for changing behaviour in both the short- and long-term: the risk part of imparted knowledge, the individuals' specific capacity for self-efficacy, and promotion of realistic expectations regarding outcomes.

According to Bandura, one's own beliefs in one's capacity to behave appropriately in prospective situations is crucial for producing outcomes of targeted behaviour [178]: *'Perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments.'* [178] Promoting healthy behaviour requires appropriate communication, an aspect of post-discharge care that the ESC guidelines on secondary prevention and lifestyle modification highly recommend. [179] Actually producing more behavioural changes is still challenging to do (e.g. getting cardiac patients to adhere to CR). [179] One reason might be that healthcare interventions might not fully or effectively use knowledge from the social cognitive theory espoused by Bandura. It is surprising, given there is a range of instruments (disease-specific and generic tools) to measure self-efficacy in a healthcare context. [172] It might be useful in the future to develop a risk-assessment tool specifically for cardiac patients undergoing an invasive valve procedure, which would measure their self-efficacy before a treatment. Having results from such a tool for SAVR patients in early CR might allow researchers to further explore ways to modify behaviour to promote a healthier lifestyle that would avoid readmissions.

8.2.4 Discussion of the process evaluation of the implementation and impact of the intervention

The AVRre trial participants who actively used the hotline support service rated it as satisfactory and much needed, because they felt safe and secure, as reported in the 3-month questionnaire. The results from the qualitatively analysed interview data supported these findings. Although the 30-DACR rate was not reduced in the intervention group, it was likely not because the intervention, as it was carried out, was lacking in fidelity. Monitoring and managing symptoms and giving educational advice to the participants was hypothesised to be the core intervention elements that might lead to fewer readmissions. The high proportion of unavoidable readmissions is part of the explanation (paper III) as to why the 30-DACR rate was not reduced as a result of the telephone support intervention. Ten intervention participants were readmitted due to medical complications, and two of them had acute cardiac tamponade; the latter of which was invasively treated with a favourable outcome. The qualitative findings confirmed that the monitoring and managing assistance for symptoms offered through the intervention was valuable for the participants. The patients appreciated the educational advice given, and the hotline staff evaluated it as being useful, mostly because of the participants' reactions. Analysis of the participating staff's field notes also showed that the participants were satisfied with the 'link' to the university hospital. Therefore, the study's theoretical foundation was justified clinically, as measured by how the participants reacted to the intervention. However, the theoretical basis that would support patient behavioural change according to evidence-based healthcare advice aimed at avoiding readmissions after SAVR needs more investigation to tailor new interventions aimed to reduce readmissions.

An important finding in paper IV was that participants reported experiencing a gap in the care continuum from hospital to discharge to follow-up care. A perceived gap in the care continuum can be caused by several factors. [27] Lack of information as perceived by the

patients and between healthcare institutions and professionals is reported to be a common cause for gaps in the care continuum, often resulting in readmissions. [27] Physicians and nurses inform patients according to their individual needs; Norwegian national legislation mandates this approach. [180] Healthcare professionals target mostly the knowledge part of the delivery of post-discharge patient information. [169] Moreover, the SAVR patients are transferred only when considered physically stable and evaluated to be able to care for themselves after discharge.

Why, then, did the SAVR participants of the AVRre trial still experience a gap in the care continuum? The PROM data in paper III demonstrated that the first month after surgery was demanding for the participants (Figs. 6, 7). However, the intervention participants reported that the telephone support system 'bridged' the care continuum, because they received and had access to trusted healthcare advice. The intervention was conducted at the hospital where the surgery was done, which helped the participants to feel safe and secure, an outcome highly appreciated by them. This outcome might point towards another reason for the participants' perception of a gap in the care continuum after discharge, such as low socio-economic status and health literacy. [181, 182] Being independent and self-caring at home after surgery and discharge might come too early for some patients, as their statements suggested that they needed that kind of support.

In addition to supporting and securing the participants' physical condition after surgery, a more systematic approach was needed to enhance their socio-psychological support too; assistance in reorienting themselves after discharge also seemed warranted. Patients are vulnerable during transition of care. [183] We found that symptoms of anxiety before treatment predicted readmissions, as reported in paper III. Being aware of this possibility can help identify patients with anxiety, who can be followed up after discharge, hopefully to prevent readmissions.

Eight of 10 participants in the intervention group reported feeling more safe and secure because of the hotline availability, an observation that was corroborated by the contents of the field notes. This underlines the interventions' effect on reducing symptoms of anxiety and can be understood in terms of the qualitative finding that participants experienced a gap in the care continuum. The national surveys in Norway of patients reporting less satisfaction with the discharge supports the participants' statements, and challenges the idea that the transition of care between hospital to home and primary care is known. [43] The main objective of the Norwegian Healthcare Coordination Reform initiated in 2012 was to construct a more efficient healthcare system for patients moving from hospital to primary care, all the while without compromising the quality of care. [184] Our finding in which the participants indicated that the transition of care related to the discharge and follow-up needed to improve suggests that the objective of the coordination reform has not yet reached SAVR patients. There is no evidence to suggest that the coordination reform has led to more readmissions. [185] A recent governmental initiative in Norway designed to enhance patient satisfaction with hospital discharge (named 'Safe discharge') became part of the national Patient Safety Program in 2017. [186]

Half of the AVRre intervention participants did not use the hotline service for various reasons, as reported in paper IV. However, these participants stated that they recognised the value of the intervention for more vulnerable patients and also said that they appreciated the availability of it if they needed it themselves. The non-users underlined the potential positive effects of the intervention in the early rehabilitation phase. Therefore, we cannot conclude that the TFU somehow limited the effect of the hotline could have had on the outcomes. The hotline was perceived as an attractive and necessary service if the non-callers should need any advice on their health condition in addition to the two scheduled TFU phone calls.

The TFU phone calls on day 2 and 9 after discharge were administered as planned and were greatly appreciated by the participants. The PC experienced a learning curve as the TFU service unfolded, a phenomenon that could have introduced performance bias, in which the later participants received more tailored information than earlier ones. However, the PC was prepared for the possibility of a learning effect issue and had expert physicians for consulting purposes standby. Moreover, all intervention participants received at least two follow-up calls. In retrospect, the hotline staff said they wished they had more education and training in the pre-trial phase. This feeling of the staff not being at the asymptote of the learning curve before the main trial began might have impaired the fidelity of the delivery of the intervention early on in the main trial. However, the results were convincing enough for us to conclude that the intervention was delivered with high fidelity. The robust follow-up of staff during the main trial was one important factor contributing to the high fidelity.

More research attention needs to be focussed on how contextual factors and the diversity and complexity of local hospitals discharge practices affect the readmission rates following SAVR. We found significant statistical differences among comparable local hospitals and the 30-DACR rates. Unfortunately, we did not have the relevant data to analyse this difference. This lack could be seen as a limit of the study. What local factors impact the 30-DACR rate is a future question to be investigated, one that requires a different design for data collection. Our results from the AVRre trial, however, can provide a foundation for developing hypotheses. Clues derived from the evaluation suggest possibilities.

We observed challenges related to patient transfer to local hospitals, the admission process at local hospitals, competence levels dealing with the SAVR patients' condition at local hospitals, and local hospital discharge management practices. One recent study demonstrated that top-ranked hospitals in the US do not have fewer readmissions but still have lower mortality rates and more satisfied patients compared to lower-ranked hospitals.

[157] When considering the SAVR treatment and the risk of readmission within 30-days after the discharge, it might be reasonable to expect a higher number of readmissions related to ensuring patient safety. Previous studies have demonstrated that cardiac valve surgery yields higher readmission rates than CABG, for example. [187] In the AVRre trial, monitoring and managing of symptoms during the intervention resulted in admissions of participants to local hospitals due to complications requiring hospital care as presented. The university hospital where the intervention was performed is the largest in Norway. However, most readmissions in the AVRre trial were locally initiated. This aspect is a limitation, as we cannot evaluate more deeply the local contexts' influences. With our mixed-methods design, we were able to at least broaden the understanding of the intervention's outcomes and reduce problems associated with a potential 'black box' evaluation. [64] Moreover, the World Health Organization states that because of the diversity of results in attempts to reduce adverse events during transition of care across different settings, it is very important to thoroughly describe the intervention implementation so the healthcare providers can understand what is most effective for improving the quality of care. [183]

To optimise the discharge and follow-up after SAVR in the AVRre trial, we found it appropriate and useful to use a mixed-methods approach. Thus, we integrated user experiences into the overall trial evaluation right from the start with the development of the hotline manual to the process evaluation of the completed trial. Having easy access to a direct phone line to secure health advice from trained project staff who were attuned to their specific health condition and individual needs, together with their self-management behaviour, was an important sign of the 'good' healthcare according to the participants' perspectives in this study. Thus, we were able to understand more of the complexity concerning the discharge and follow-up of SAVR patients, other than just understanding the effectiveness of the tested intervention on primary outcomes.

In a real-world context, where science unfolds quickly and stakeholders strive to understand the core elements of a specific topic, one cannot fully describe all aspects of a RCT in one or two papers. However, those studies can accumulate certain pieces of evidence that can contribute to theory and then bring the clinical problem towards a relevant solution to apply in the clinic. This notion is in line with traditions spanning Donabedian ideas [46] to the MRC organising framework [51] and others that attempt to expand our scientific knowledge through the integration of different methodologies to profile human subjects and their behaviour in the healthcare context and the society where behaviour takes place. A valid scientific description can be accurate at the time of its birth; however, it evolves over time, pushing the science to change accordingly. With the new technology revolution of today and political disputes about science, it seems *à propos* to mention these issues, as it affects the context where research is taking place. We can appeal to the science communities to increase their flexibility of scientific thinking and to adapt to the evolution of an issue in order to overcome contextual challenges. This shift in thinking will, I believe, preserve knowledge as the most valuable human asset, one that can be used to guide development when applied wisely.

8.3 Ethical considerations

The AVRre trial was approved by the Regionale komiteer for medisinsk og helsefaglig forskningsetikk sør-øst (REK) (Norwegian Regional Committees for Medical Research Ethics; REK South-East B; approval no. 2013/2031B) that oversees human subjects' research at the University of Oslo. The OUH Data Protection Officer approved the focus group interviews with the nurses staffing the hotline. To ensure anonymity, all digital data were stored on a secure server within the OUH system, and all patients paper documents were stored in a closed safe at the Center for Patient-centered Heart and Lung Research,

Department of Cardiothoracic Surgery, Division of Cardiovascular and Pulmonary Diseases, OUH, Ullevål, Oslo, Norway. Finally, all participants gave their written informed consent to participate in the AVRre trial, per the Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects. [67]

We observed that some participants were not content with being assigned to the control group during randomisation. The PC spent time reasoning with these patients, informing them about the important role that control-group participants play in clinical trials and how valuable their answers are in follow-up questionnaires. Still, some control-group participants stated that the 24/7 hotline should obviously be provided to all patients. We paid attention to these concerns by taking additional time to make sure that they understood information about the trial, and most importantly, the control-group concept, that comparing treated and untreated groups is necessary in order to determine the true effects of an intervention. We seriously considered their explicit disappointment of being randomly allocated to the control group by discussing their concerns with them, emphasising the critical contribution of control patients, that without them, we cannot determine whether the intervention is indeed effective or useful.

For researchers conducting the trial, the concerns of control-group patients were an important topic worthy of consideration and notice. It led to discussions to determine, or at least to arrive at some hypotheses as to, what their concerns might mean. Also, we discussed ways on how to inform control patients in cases when we perceive that all of the participants greatly desire to receive the intervention. The researchers discussed *a priori* whether the control group should be offered a lighter version of the intervention, an offer we know, in retrospect, would have likely led control-group participants to be more satisfied after discharge. However, due to the risk of introducing additional 'noise' into the analysis of the intervention effects, we chose not to do so and not to construct a 'black box problem' larger

than necessary for understanding the experimental outcomes. However, the control participants did receive the standard care at the time, which the ethical committee took into consideration when assessing our study protocol for approval.

In 2000, Emanuel and colleagues proposed an ethical framework for conducting ethical clinical research using an RCT design. [188] Value is one of seven requirements in the proposed framework, which highlights the necessary aims of producing potentially positive effects for the participants and society, and the responsibility of publishing both negative and positive results. [188] The participants in the AVRre trial evaluated the intervention as a service that promotes a feeling of safety and trust in patients, and as a satisfactory early rehabilitation intervention after SAVR surgery.

8.4 Surplus data and future research considerations

The AVRre trial collected ancillary data that was not analysed as part of this thesis work due to time and economic constraints. Firstly, we did not conduct a cost-utility analysis. This will be part of future research, which will address the main research question: Is the intervention in the AVRre trial more cost-effective than usual care? Secondly, an investigation of the one-year readmissions data from the AVRre trial will be conducted. This planned study will address the following research questions: What is the one-year incidence of readmissions after SAVR? What causes readmissions in the first year after SAVR? What predicts one-year readmissions after SAVR? What is the incidence of planned and unplanned out-patient consultations in the first year after SAVR? Are there any differences between the trial groups with regard to the incidence of readmissions in the first year after surgery? Thirdly, we will analyse the content of research interviews with physicians, patients, and caregivers regarding their experiences with the transition of care after SAVR. The following possible research questions will be addressed: How do stakeholders perceive the discharge and care after

SAVR? What do the qualitative perspectives of trial participants reveal in terms of what can improve the discharge and follow-up of SAVR patients? Fourthly, I wish to continue the work with the NPR to gain valid thirty-day readmission data on SAVR, which also would be useful for TAVR populations with a similar care trajectory after the discharge. Fifthly, a Master's thesis in Nursing Science was completed in 2019 at the University of Oslo. The Master's thesis analysed data obtained from the registration forms used to log information about the hotline calls of the AVRre trial; the goal was to identify the various themes participants talked about in the calls. [189] This topic will be a similar objective for a planned peer-reviewed paper reporting on what concerns participants had during the intervention period.

9 Conclusions

The thesis work discussed here, and reported in the four published papers, provides new and important knowledge for optimising the discharge and follow-up care of SAVR patients. Firstly, this work found a high proportion of 30-DACRs following SAVR surgery, providing clinicians and researchers vital knowledge on to what extent readmissions burden not only patients and by extension their informal caregivers, but also the healthcare professionals charged with their care and the healthcare system, in general. These findings, therefore, serve as an impetus to improve healthcare related to discharge and readmissions after SAVR. Secondly, the AVRre trial found that the telephone intervention reduced patient symptoms of anxiety within the first 30 days after surgery but failed to reduce the 30-DACR rate. Attempts to reduce symptoms of anxiety is warranted, because less anxiety is associated with a lower risk of mortality and other adverse events. [190] Thirdly, the trial also found a high proportion of unavoidable readmissions associated with the SAVR treatment. This finding is important and provides clues to tailoring new interventions to improve discharge and follow-up of these patients. Fourthly, patient participants experienced the intervention as being useful, bolstering their trust in the intervention and giving them an overall sense of security. The 24/7 hotline also increased their satisfaction with the discharge process and follow-up after SAVR.

Taken together, these findings demonstrate that the novel scientific mixed-methods approach employed in the AVRre trial was useful, as was having users participate from the trial design stages to formal evaluation of this clinical trial, for gaining the knowledge needed to optimise the discharge and follow-up of SAVR patients.

Conclusions as they relate to the aims of the AVRre study[2]:

I.

[2] Roman numerals identify the published papers.

- **A protocol paper was published in a timely manner to ensure transparency in the reporting of the AVRre trial outcomes.**

II.

- **The overall worldwide incidence of 30-DACRs after SAVR is relatively high at 17% (95% CI: 16-18%), which is similar to the incidence after TAVR, which was 16% (95% CI: 15-18%).**
- Multi-centre studies yielded statistically significantly higher 30-DACR rates than single-centre studies after SAVR and TAVR.
- There is a lack of prospective studies on 30-DACR rate after SAVR.
- There is lack of evidence on independent risk factors for 30-DACR after SAVR, and there is lack of transparent reporting on the validation of readmission data used in clinical research.

III.

- **The intervention did not significantly reduce the 30-DACR rate after SAVR nor did it, in general, improve patient-reported outcomes, except for symptoms of anxiety (which did significantly decrease up to 30 days after surgery).**
- Total incidence of unplanned 30-DACR after SAVR was 22.3%, and of these, most (83%) occurred within 14 days after SAVR, providing an impetus to tailor future readmission-reduction interventions to target the first 30 days after hospital discharge.
- Independent risk factors for readmission within 30 days after SAVR are the presence of symptoms of anxiety before surgery and pleural drainage before hospital discharge.
- Unavoidable readmissions after SAVR were estimated to be as high as 75%.

IV.

- **Participating SAVR patients were generally satisfied with the intervention and perceived it as valuable in bolstering their trust in their care and providing them with a sense of security. These findings underline the value of an additional process evaluation that involves investigating the implementation of and patient reactions to an intervention in a clinical trial.**
- Some lack of hotline staff preparedness might have been a barrier to the fidelity of carrying out the intervention; however, robust support given to the staff during the main trial enabled the trial to be performed safely.
- Context influences the 30-DACR rate after SAVR, and local hospitals should focus more attention on determining and analysing the causes of the significant differences in discharge practices among hospitals in order to identify factors that might be targets for reducing the overall readmission rate after SAVR.

10 Future perspectives

A large amount of data was collected in the AVRre trial through its mixed-methods design. As described in section 8.4, a Master's thesis based on data from the AVRre study has been completed, and more research data will be finalised in order to further analyse other discharge-related outcomes for this SAVR patient population. However, due to similarities with the present TAVR population and for future benchmarking of various treatment choices available today, physicians may choose to apply the less invasive TAVR technique to other at-risk AS patient populations. In that case, the kinds of analyses done in this thesis work will have to be replicated in these other kinds of TAVR populations.

Knowledge gained from this PhD dissertation work will guide future research on discharge-related outcomes of those AVR patients. To fully capture an intervention's effect on a patient population, more clinical research will have to be predicated upon obtaining intervention-user knowledge and perspectives, and these must be sampled and considered at all phases of healthcare intervention development and conduct. Furthermore, user knowledge and perspectives should be included in a timely fashion, as should careful ethical considerations, scientific assessment of clinical relevance, and validation of results in terms of the user's perspectives. Systematic implementation of these aspects will help future researchers and clinicians better grasp the validity and efficacy of a clinical trial, enabling them to reach appropriate and more robust conclusions on the effectiveness of the healthcare intervention they are assessing. For future interventions aimed at optimising patient discharge and follow-up after surgical aortic valve replacement, faithfully implementing such a research programme will likely reduce hospital readmissions, improve patient-reported outcomes, and improve healthcare in general.

11 References

1. Vahanian A, Alfieri O, Andreotti F, Antunes MJ, Barón-Esquivias G, Baumgartner H, et al. Guidelines on the management of valvular heart disease (version 2012). European Heart Journal. 2012;33(19):2451-96. https://doi.org/10.1093/eurheartj/ehs109

2. Vahanian A, Iung B, Himbert D, Nataf P. Changing demographics of valvular heart disease and impact on surgical and transcatheter valve therapies. The International Journal of Cardiovascular Imaging. 2011;27(8):1115-22. https://doi.org/10.1007/s10554-011-9804-7

3. Qualcommlife. Hospital Readmissions in Europe. Using remote monitoring to contain costs. The Burill Report.; 2013. [Available from: https://qualcommlife.com/wp-content/themes/qualcommlife/assets/white-papers/documents/hospital-readmissions-in-europe.pdf

4. Falk V, Baumgartner H, Bax JJ, De Bonis M, Hamm C, Holm PJ, et al. 2017 ESC/EACTS Guidelines for the management of valvular heart disease. European Journal of Cardio-thoracic Surgery. 2017;52(4):616-64. https://doi.org/10.1093/ejcts/ezx324

5. Baumgartner H, Falk V, Bax JJ, De Bonis M, Hamm C, Holm PJ, et al. 2017 ESC/EACTS Guidelines for the management of valvular heart disease. European Heart Journal. 2017;38(36):2739-91. https://doi.org/10.1093/eurheartj/ehx391

6. Norwegian Haeart Registry. Norwegian Cardiac Surgery Registry Report 2016. National Service Institution for Medical Quality Registry; 2016.

7. Leon MB, Smith CR, Mack M, Miller DC, Moses JW, Svensson LG, et al. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. The New England Journal of Medicine. 2010;363(17):1597-607. https://doi.org/10.1056/NEJMoa1008232

8. Fujita B, Ensminger S, Bauer T, Mollmann H, Beckmann A, Bekeredjian R, et al. Trends in practice and outcomes from 2011 to 2015 for surgical aortic valve replacement: an

update from the German Aortic Valve Registry on 42 776 patients. European Journal of Cardio-thoracic surgery. 2017. https://doi.org/10.1093/ejcts/ezx408

9. Benstoem C, Moza A, Autschbach R, Stoppe C, Goetzenich A. Evaluating outcomes used in cardiothoracic surgery interventional research: a systematic review of reviews to develop a core outcome set. PloS One. 2015;10(4):e0122204-e. https://doi.org/10.1371/journal.pone.0122204

10. Goldfarb M, Drudi L, Almohammadi M, Langlois Y, Noiseux N, Perrault L, et al. Outcome Reporting in Cardiac Surgery Trials: Systematic Review and Critical Appraisal. Journal of the American Heart Association. 2015;4(8):e002204. https://doi.org/10.1161/jaha.115.002204

11. Holmes C, Briffa N. Patient-Reported Outcome Measures (PROMS) in patients undergoing heart valve surgery: why should we measure them and which instruments should we use? Open Heart. 2016;3(1):e000315. https://doi.org/10.1136/openhrt-2015-000315

12. Hannan EL, Samadashvili Z, Jordan D, Sundt TM, 3rd, Stamato NJ, Lahey SJ, et al. Thirty-Day Readmissions After Transcatheter Aortic Valve Implantation Versus Surgical Aortic Valve Replacement in Patients with Severe Aortic Stenosis in New York State. Circulation Cardiovascular Interventions. 2015;8(8):e002744. https://doi.org/10.1161/circinterventions.115.002744

13. Iribarne A, Chang H, Alexander JH, Gillinov AM, Moquete E, Puskas JD, et al. Readmissions After Cardiac Surgery: Experience of the National Institutes of Health/Canadian Institutes of Health Research Cardiothoracic Surgical Trials Network. The Annals of Thoracic Surgery. 2014;98(4):1274-80. https://doi.org/http://doi.org/10.1016/j.athoracsur.2014.06.059

14. Cribier AMD, Eltchaninoff HMD, Bash AP, Borenstein NMD, Tron CMD, Bauer FMD, et al. Percutaneous Transcatheter Implantation of an Aortic Valve Prosthesis for Calcific Aortic Stenosis: First Human Case Description. Circulation. 2002;106(24):3006-8.

15. Popma J, Michael D, Yakubov S, Mubashir M, Heiser J, Merhi W, et al. Transcatheter Aortic-Valve Replacement with a Self-Expanding Valve in Low-Risk Patients. The New England Journal of Medicine. 2019:1706-15. https://doi.org/10.1056/NEJMoa1816885

16. Howard C, Jullian L, Joshi M, Noshirwani A, Bashir M, Harky A. TAVI and the future of aortic valve replacement. Journal of Cardiac Surgery. 2019. https://doi.org/10.1111/jocs.14226

17. Nombela-Franco L, del Trigo M, Morrison-Polo G, Veiga G, Jimenez-Quevedo P, Abdul-Jawad Altisent O, et al. Incidence, Causes, and Predictors of Early (</=30 Days) and Late Unplanned Hospital Readmissions After Transcatheter Aortic Valve Replacement. JACC: Cardiovascular Interventions. 2015;8(13):1748-57. https://doi.org/10.1016/j.jcin.2015.07.022

18. Lindman A, Kristoffersen, DT., Hassani, S., Tomic, O., Helgeland, J. 30-day readmissions among the elderly 2011–2013. Results for hospitals and municipalities.: Norwegian Research Centre for Health Services; May 2015.

19. Barreto-Filho JA, Wang Y, Dodson JA, Desai MM, Sugeng L, Geirsson A, et al. Trends in aortic valve replacement for elderly patients in the United States, 1999-2011. JAMA. 2013;310(19):2078-85. https://doi.org/10.1001/jama.2013.282437

20. Sibilitz KL, Berg SK, Thygesen LC, Hansen TB, Kober L, Hassager C, et al. High readmission rate after heart valve surgery: A nationwide cohort study. International Journal of Cardiology. 2015;189:96-104. https://doi.org/10.1016/j.ijcard.2015.04.078

21. Auensen A, Hussain AI, Bendz B, Aaberge L, Falk RS, Walle-Hansen MM, et al. Morbidity outcomes after surgical aortic valve replacement. Open Heart. 2017;4(1):e000588. https://doi.org/10.1136/openhrt-2017-000588

22. Brown ML, Schaff HV, Lahr BD, Mullany CJ, Sundt TM, Dearani JA, et al. Aortic valve replacement in patients aged 50 to 70 years: Improved outcome with mechanical versus biologic prostheses. The Journal of Thoracic and Cardiovascular Surgery. 2008;135(4):878-84. https://doi.org/http://dx.doi.org/10.1016/j.jtcvs.2007.10.065

23. Fischer C, Lingsma HF, Marang-van de Mheen PJ, Kringos DS, Klazinga NS, Steyerberg EW. Is the Readmission Rate a Valid Quality Indicator? A Review of the Evidence. PloS One. 2014;9(11):e112282. https://doi.org/10.1371/journal.pone.0112282

24. Lie I, Danielsen SO, Tonnessen T, Solheim S, Leegaard M, Sandvik L, et al. Determining the impact of 24/7 phone support on hospital readmissions after aortic valve replacement surgery (the AVRre study): study protocol for a randomised controlled trial. Trials. 2017;18(1):246. https://doi.org/10.1186/s13063-017-1971-y

25. Centers for Medicare & Medicaid Services Services; CMS.gov. Hospital Readmissions Reduction Program (HRRP); 2019 [updated 01/16/2019. [Available from: https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html

26. Joynt Maddox KE. Readmissions have declined, and mortality has not increased: The importance of evaluating unintended consequences. JAMA. 2017;318(3):243-4. https://doi.org/10.1001/jama.2017.8705

27. Braet A. Unplanned readmissions in Belgian hospitals -importance of quality on in-hospital care and quality of care transition. [Doctoral thesis.]. KU Leuven, Belgium; 2016.

28. OECD (2017). Hospital discharges. In Health at a Glance 2017: OECD Indicators. Paris; 2017. https://doi.org/10.1787/health_glance-2017-en

29. Heggestad T. Do hospital length of stay and staffing ratio affect elderly patients' risk of readmission? A nation-wide study of Norwegian hospitals. Health Services Research. 2002;37(3):647-65.

30. van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. Journal of Clinical Epidemiology. 2012;65(2):126-31. https://doi.org/10.1016/j.jclinepi.2011.08.002

31. Winkley Shroyer AL, Bakaeen F, Shahian DM, Carr BM, Prager RL, Jacobs JP, et al. The Society of Thoracic Surgeons Adult Cardiac Surgery Database: The Driving Force for Improvement in Cardiac Surgery. Seminars in Thoracic and Cardiovascular Surgery. 2015;27(2):144-51. https://doi.org/10.1053/j.semtcvs.2015.07.007

32. The Norwegian Directorate of Health. Norwegian Patient Registry; 2019 [Available at: https://helsedirektoratet.no/english/norwegian-patient-registry

33. Braet A, Weltens C, Sermeus W, Vleugels A. Risk factors for unplanned hospital re-admissions: a secondary data analysis of hospital discharge summaries. Journal of Evaluation in Clinical Practice. 2015;21(4):560-6. https://doi.org/10.1111/jep.12320

34. Tsai TC, Joynt KE, Orav EJ, Gawande AA, Jha AK. Variation in surgical-readmission rates and quality of hospital care. The New England Journal of Medicine. 2013;369(12):1134-42. https://doi.org/10.1056/NEJMsa1303118

35. Fischer C, Anema HA, Klazinga NS. The validity of indicators for assessing quality of care: a review of the European literature on hospital readmission rate. European Journal of Public Health. 2012;22(4):484-91. https://doi.org/10.1093/eurpub/ckr165

36. Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The incidence and severity of adverse events affecting patients after discharge from the hospital. Annals of Internal Medicine. 2003;138(3):161-7.

37. van Walraven C, Bennett C, Jennings A, Austin PC, Forster AJ. Proportion of hospital readmissions deemed avoidable: a systematic review. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2011;183(7):E391-402. https://doi.org/10.1503/cmaj.101860

38. Horwitz L, Partovian, C., Lin, Z., Herrin, J., Grady, J., Conover, M., Montague, J., Dillaway, C., Bartczak, K., Ross, J., Bernheim, S., Drye, E., Krumholz, HM. Hospital-Wide (All-Condition) 30-Day Risk-Standardized Readmission Measure DRAFT Measure Methodology Report. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHSC/CORE); 2011 August 10, 2011.

39. Waring J MF, Bishop S, et al. Chapter 2, Hospital discharge and patient safety: reviews of the literature. Southampton (UK): NIHR Journals Library; 2014 [Available from: https://www.ncbi.nlm.nih.gov/books/NBK259995/

40. Galvin EC, Wills T, Coffey A. Readiness for hospital discharge: A concept analysis. Journal of Advanced Nursing. 2017;73(11):2547-57. https://doi.org/10.1111/jan.13324

41. World Health Organization (WHO). Transitions of Care: Technical Series on Safer Primary Care.; 2016. [Available at: https://apps.who.int/iris/bitstream/handle/10665/252272/9789241511599-eng.pdf;jsessionid=6A7C5DE361A9661C6AF703F3CC23AB86?sequence=1

42. Burke RE, Guo R, Prochazka AV, Misky GJ. Identifying keys to success in reducing readmissions using the ideal transitions in care framework. BMC Health Services Research. 2014;14:423. https://doi.org/10.1186/1472-6963-14-423

43. Holmboe O, Bjertnaes OA. Inpatients' experiences with somatic hospitals in 2014. National results. PasOpp-report 2-2015.: Oslo: National Research Center for Health Care, 2015.

44. Leppin AL, Gionfriddo MR, Kessler M, Brito JP, Mair FS, Gallacher K, et al. Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials. JAMA Internal Medicine. 2014;174(7):1095-107. https://doi.org/10.1001/jamainternmed.2014.1608

45. Hansen LO, Young RS, Hinami K, Leung A, Williams MV. Interventions to reduce 30-day rehospitalization: a systematic review. Annals of Internal Medicine. 2011;155(8):520-8. https://doi.org/10.7326/0003-4819-155-8-201110180-00008

46. Donabedian A. Evaluating the quality of medical care. The Milbank Memorial Fund quarterly. 1966;44(3):Suppl:166-206.

47. Speight J, Barendse SM. FDA guidance on patient reported outcomes. BMJ (Clinical research ed). 2010;340:c2921. https://doi.org/10.1136/bmj.c2921

48. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. JAMA. 2013;309(8):814-22. https://doi.org/10.1001/jama.2013.879

49. Manary MP, Boulding W, Staelin R, Glickman SW. The patient experience and health outcomes. The New England Journal of Medicine. 2013;368(3):201-3. https://doi.org/10.1056/NEJMp1211775

50. Buhse S, Muhlhauser, I. Development of complex interventions. Outcome modeling. In: Richards D, Hallberg, IR., editors. Complex interventions in health An overview of research methods. 1th ed. New York, US.: Routledge.; 2015. p. 96-102.

51. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ (Clinical research ed). 2015;350:h1258. https://doi.org/10.1136/bmj.h1258

52. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ (Clinical research ed). 2008;337:a1655. https://doi.org/10.1136/bmj.a1655

53. Richards D. The complex intervention framework. In: Richards D, Hallberg IR., editors. Complex interventions in health. New York: Routledge; 2015. p. 1-16.

54. Bleijenberg N, de Man-van Ginkel JM, Trappenburg JCA, Ettema RGA, Sino CG, Heim N, et al. Increasing value and reducing waste by optimizing the development of complex interventions: Enriching the development phase of the Medical Research Council (MRC) Framework. International Journal of Nursing Studies. 2018;79:86-93. https://doi.org/https://doi.org/10.1016/j.ijnurstu.2017.12.001

55. Ioannidis JPA, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. The Lancet. 2014;383(9912):166-75. https://doi.org/https://doi.org/10.1016/S0140-6736(13)62227-8

56. Leeman J, Chang Y, Voils CI, Crandell JL, Sandelowski M. A mixed-methods approach to synthesizing evidence on mediators of intervention effects. Western Journal of Nursing Research. 2011;33(7):870-900. https://doi.org/10.1177/0193945911402365

57. Nelson G, Macnaughton E, Goering P. What qualitative research can contribute to a randomized controlled trial of a complex community intervention. Contemporary Clinical Trials. 2015;45:377-84. https://doi.org/https://doi.org/10.1016/j.cct.2015.10.007

58. Rothman K. Epidemiology : an introduction. 2nd. ed. Oxford: Oxford Univeristy Press; 2012.

59. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. Trials. 2017;18(1):122-. https://doi.org/10.1186/s13063-017-1870-2

60. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. Journal of the Royal Society of Medicine. 2011;104(12):510-20. https://doi.org/10.1258/jrsm.2011.110180

61. New England Journal of Medicine Catalyst. What is patient-centered care? 2017 [Available from: https://catalyst.nejm.org/what-is-patient-centered-care/

62. Rosvik AH, Movik E, Nylenna M. [Do patients participate in the development of clinical guidelines in Norway?]. Journal of the Norwegian Medical Association. 2010;130(22):2236-8. https://doi.org/10.4045/tidsskr.09.1511

63. Kernick D. Wanted--new methodologies for health service research. Is complexity theory the answer? Family Practice. 2006;23(3):385-90. https://doi.org/10.1093/fampra/cml011

64. Scriven M. The Fine Line Between Evaluation and Explanation. Research on Social Work Practice. 1999;9(4):521-4. https://doi.org/10.1177/104973159900900407

65. Sermeus W. Modelling process and outcomes in complex interventions. In: Richards D, Hallberg IR., editors. Complex interventions in health. 1st ed. New York: Routledge; 2015. p. 111-20.

66. Sharp D, Palmore T, Grady C. The ethics of empowering patients as partners in healthcare-associated infection prevention. Infection control and hospital epidemiology. 2014;35(3):307-9. https://doi.org/10.1086/675288

67. World Medical Association. Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects 2013 [Available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

68. Farup PG, Skar V. Collaboration by use of the Internet yields data of high quality and detects non-uniform management of patients with Helicobacter pylori infection. Scandinavian Journal of Gastroenterology. 2002;37(12):1466-70.

69. Russo N, Compostella L, Tarantini G, Setzu T, Napodano M, Bottio T, et al. Cardiac rehabilitation after transcatheter versus surgical prosthetic valve implantation for aortic stenosis in the elderly. European Journal of Preventive Cardiology. 2014;21(11):1341-8. https://doi.org/10.1177/2047487313494029

70. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gotzsche PC, Krleza-Jeric K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Annals of Internal Medicine. 2013;158(3):200-7. https://doi.org/10.7326/0003-4819-158-3-201302050-00583

71. Kvale S, Brinkmann, S. InterView: Learning the Craft of Qualitative Research Interviewing. 3.ed.2015 Oslo: Gyldendal Akademisk; 2009.

72. Wheeldon J, Faubert J. Framing Experience: Concept Maps, Mind Maps, and Data Collection in Qualitative Research. International Journal of Qualitative Methods. 2009;8(3):68-83. https://doi.org/10.1177/160940690900800307

73. Farrand P, Hussain F, Hennessy E. The efficacy of the 'mind map' study technique. Medical education [Internet]. Wiley; 2002;36(5):426-31. https://doi.org/10.1046/j.1365-2923.2002.01205.x

74. The Norwegian Medical Association. Norwegian index for medical acute support. 3 ed. Stavanger, Norway: Laerdal Medical A/S; 2009.

75. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Journal of Clinical Epidemiology. 2009;62(10):1006-12. https://doi.org/10.1016/j.jclinepi.2009.06.005

76. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Medical Informatics and Decision Making. 2007;7(1):16. https://doi.org/10.1186/1472-6947-7-16

77. GA Wells BS, D O'Connell, J Peterson, V Welch, M Losos, P Tugwell. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Accessed Novemer 10, 2017.: Ottawa Hospital Research Institute; 2014 [Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

78. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ (Clinical research ed). 2010;340:c869. https://doi.org/10.1136/bmj.c869

79. Sangha O, Stucki G, Liang MH, Fossel AH, Katz JN. The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. Arthritis and Rheumatism. 2003;49(2):156-63. https://doi.org/10.1002/art.10993

80. MAPI Research Trust. Hospital Anxiety and Depression Scale (HADS) ePROVIDE: Mapi Research Trust; 2015 [Available from: https://eprovide.mapi-trust.org/instruments/hospital-anxiety-and-depression-scale

81. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy (Amsterdam, Netherlands). 1990;16(3):199-208.

82. Xie F, Pickard AS, Krabbe PFM, Revicki D, Viney R, Devlin N, et al. A Checklist for Reporting Valuation Studies of Multi-Attribute Utility-Based Instruments (CREATE). Pharmacoeconomics. 2015;33(8):867-77. https://doi.org/10.1007/s40273-015-0292-9

83. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. Journal of Psychosomatic Research. 2002;52(2):69-77.

84. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatrica Scandinavica. 1983;67(6):361-70.

85. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology. 2007;60(1):34-42. https://doi.org/10.1016/j.jclinepi.2006.03.012

86. Pinnock H, Barwick M, Carpenter CR, Eldridge S, Grandes G, Griffiths CJ, et al. Standards for Reporting Implementation Studies (StaRI) Statement. BMJ (Clinical research ed). 2017;356:i6795. https://doi.org/10.1136/bmj.i6795

87. Maxwell JA. Qualitative Research Design. An interactive approach.: SAGE Publications Inc.; 2013.

88. Malterud K. Qualitative methods in medical research. An introduction. Oslo: Universitetsforlaget.; 2013.

89. Kreuger RA, Casey, M.A. Focus groups: a practical guide for applied research. 5th ed. 2014. SAGE Publications Inc.

90. Engzelius J. Readmission at Bærum hospital. A coordination pilot project. Oslo, Norway: Health East South.; 2012.

91. DerSimonian R, Laird N. Meta-analysis in clinical trials revisited. Contemporary Clinical Trials. 2015;45(Pt A):139-45. https://doi.org/10.1016/j.cct.2015.09.002

92. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. Research Synthesis Methods. 2010;1(2):112-25. https://doi.org/10.1002/jrsm.11

93. Stata Corp. 2015. Stata Statistical Software: Release 14.: College Station, TX: StataCorp LP.

94. MedCalc. MedCalc Software: MedCalc Software bv, Ostend, Belgium; 2016 [Available from: https://www.medcalc.org.

95. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34.

96. Aalen O, Riksheim, M. Survival analysis. In: Veierød MB, Lydersen S, Laake P, editors. Medical Statistics: in clinical and epidemiological research. 1st ed. Oslo: Gyldendal Akademiske; 2012. p. 167-98.

97. Vickers AJ, Altman DG. Statistics notes: Analysing controlled trials with baseline and follow up measurements. BMJ (Clinical research ed). 2001;323(7321):1123-4. https://doi.org/10.1136/bmj.323.7321.1123

98. Thoresen M, Gjessing HK. Mixed models. In: Veierød MB, Lydersen S, Laake P, editors. Medical Statistics: in clinical and epidemiological research. 1st ed. Oslo: Gyldendal Akademiske; 2012. p. 231-58.

99. Bjørnstad J, Lydersen S. Missing data. In: Veierød MB, Lydersen S, Laake P, editors. Medical Statistics: in clinical and epidemiological research. 1st ed. Oslo: Gyldendal Akademiske; 2012. p. 428-61.

100. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ (Clinical research ed). 2009;338:b2393. https://doi.org/10.1136/bmj.b2393

101. Harell Jr. FE. Regression Modelling Strategies. New York: Springer Verlag; 2001.

102. van Walraven C, Jennings A, Taljaard M, Dhalla I, English S, Mulpuru S, et al. Incidence of potentially avoidable urgent readmissions and their relation to all-cause urgent

readmissions. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2011;183(14):E1067-72. https://doi.org/10.1503/cmaj.110400

103. NVivo. NVivo 10. London, UK.: QRS International.; 2014.

104. NVivo. NVivo 11 Pro. London, UK.: QRS International; 2018.

105. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scandinavian Journal of Public Health. 2012;40(8):795-805. https://doi.org/10.1177/1403494812465030

106. Basu AP, Pearse JE, Rapley T. Publishing protocols for trials of complex interventions before trial completion - potential pitfalls, solutions and the need for public debate. Trials. 2017;18(1):5. https://doi.org/10.1186/s13063-016-1757-7

107. Piaggio G, Elbourne D, Schulz KF, Villar J, Pinol APY, Gülmezoglu AM, et al. The reporting of methods for reducing and detecting bias: an example from the WHO Misoprostol Third Stage of Labour equivalence randomised controlled trial. BMC Medical Research Methodology. 2003;3:19-. https://doi.org/10.1186/1471-2288-3-19

108. Vardakas KZ, Tsopanakis G, Poulopoulou A, Falagas ME. An analysis of factors contributing to PubMed's growth. Journal of Informetrics. 2015;9(3):592-617. https://doi.org/https://doi.org/10.1016/j.joi.2015.06.001

109. Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Systematic Reviews. 2017;6(1):245. https://doi.org/10.1186/s13643-017-0644-y

110. Eden J LL, Berg A, et al. Finding What Works in Health Care: Standards for Systematic Reviews Washington (DC, US): Institute of Medicine (US) Committee on Standards for Systematic Reviews of Comparative Effectiveness Research; 2011 [Available from: https://www.ncbi.nlm.nih.gov/books/NBK209517/

111. Borenstein M, Hedges, LV., Higgins, JPT., Rothstein, HR. Introduction to meta-analysis. 1th ed. UK: Wiley; 2009.

112. Ekmekci PE. An increasing problem in publication ethics: Publication bias and editors' role in avoiding it. Medicine, Health Care, and Philosophy. 2017;20(2):171-8. https://doi.org/10.1007/s11019-017-9767-0

113. Paez A. Grey literature: An important resource in systematic reviews. Journal of Evidence-based Medicine. 2017. https://doi.org/10.1111/jebm.12265

114. Dudas V, Bookwalter T, Kerr KM, Pantilat SZ. The impact of follow-up telephone calls to patients after hospitalization. The American Journal of Medicine. 2001;111(9, Supplement 2):26-30. https://doi.org/https://doi.org/10.1016/S0002-9343(01)00966-4

115. Koehler BE, Richter KM, Youngblood L, Cohen BA, Prengler ID, Cheng D, et al. Reduction of 30-day postdischarge hospital readmission or emergency department (ED) visit rates in high-risk elderly medical patients through delivery of a targeted care bundle. Journal of Hospital Medicine. 2009;4(4):211-8. https://doi.org/10.1002/jhm.427

116. Polit D, Beck, CT. Nursing research. Generating and assessing evidence for nursing pracice. 9th. ed. Philadelphia, US.: Wolters Kluwer/Lippincott Williams & Wilkins; 2012.

117. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ (Clinical research ed). 2011;343:d5928. https://doi.org/10.1136/bmj.d5928

118. Berry JG, Gay JC, Joynt Maddox K, Coleman EA, Bucholz EM, O'Neill MR, et al. Age trends in 30 day hospital readmissions: US national retrospective analysis. BMJ (Clinical research ed). 2018;360:k497. https://doi.org/10.1136/bmj.k497

119. Kaptchuk TJ. Effect of interpretive bias on research evidence. BMJ (Clinical research ed). 2003;326(7404):1453-5. https://doi.org/10.1136/bmj.326.7404.1453

120. Meyer G, Kopke S. Introduction to section 1: Developing complex interventions. In: Richards D, Hallberg, IR., editors. Complex interventions in health An overview of research methods. 1th ed. New York, US.: Routledge.; 2015. p.53-56

121. Rao G, Lopez-Jimenez F, Boyd J, D'Amico F, Durant NH, Hlatky MA, et al. Methodological Standards for Meta-Analyses and Qualitative Systematic Reviews of Cardiac Prevention and Treatment Studies: A Scientific Statement From the American Heart Association. Circulation. 2017. https://doi.org/10.1161/cir.0000000000000523

122. Higgins J, Green, S (Editors). Identifying and measuring heterogeneity Cochrane Org.: Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0; 2011 [Available from: https://handbook-5-1.cochrane.org/

123. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ (Clinical research ed). 2003;327(7414):557-60. https://doi.org/10.1136/bmj.327.7414.557

124. Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, Norman G, et al. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. Journal of Clinical Epidemiology. 2013;66(7):736-42; quiz 42.e1-5. https://doi.org/10.1016/j.jclinepi.2013.02.004

125. Mercuri M, Gafni A. The evolution of GRADE (part 3): A framework built on science or faith? Journal of evaluation in clinical practice. 2018;24(5):1223-31. https://doi.org/10.1111/jep.13016

126. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of Chronic Diseases. 1987;40(5):373-83. https://doi.org/10.1016/0021-9681(87)90171-8

127. Hall WH, Ramachandran R, Narayan S, Jani AB, Vijayakumar S. An electronic application for rapidly calculating Charlson comorbidity score. BMC Cancer. 2004;4(1):94. https://doi.org/10.1186/1471-2407-4-94

128. Laor A, Tal S, Guller V, Zbar AP, Mavor E. The Charlson Comorbidity Index (CCI) as a Mortality Predictor after Surgery in Elderly Patients. The American Surgeon. 2016;82(1):22-7.

129. Quan H, Parsons GA, Ghali WA. Validity of information on comorbidity derived rom ICD-9-CCM administrative data. Medical Care. 2002;40(8):675-85. https://doi.org/10.1097/01.mlr.0000020927.46398.5d

130. Bjørnnes A. Pain experiences of men and women following cardiac surgery. [Doctoral thesis.]. University of Oslo, Norway: Oslo; 2016.

131. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Intention-to-treat versus per-protocol analysis. Perspect Clin Res. 2016;7(3):144-6. https://doi.org/10.4103/2229-3485.184823

132. McCoy CE. Understanding the Intention-to-treat Principle in Randomized Controlled Trials. West Journal of Emergency Medicine. 2017;18(6):1075-8. https://doi.org/10.5811/westjem.2017.8.35985

133. Hernán MA, Hernández-Díaz S. Beyond the intention-to-treat in comparative effectiveness research. Clin Trials. 2012;9(1):48-55. https://doi.org/10.1177/1740774511420743

134. Thoresen M. Longitudinal analysis. In: Veierød MB, Lydersen S, Laake P, editors. Medical statistics: in clinical and epidemiological research. 1st ed. Oslo: Gyldendal Akademiske; 2012. p. 259-87.

135. Konerding U. What does Cronbach's alpha tell us about the EQ-5D? A methodological commentary to "Psychometric properties of the EuroQol Five-Dimensional

Questionnaire (EQ-5D-3L) in caregivers of autistic children". Quality of Life Research. 2013;22(10):2939-40. https://doi.org/10.1007/s11136-013-0430-9

136. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Medical Research Methodology. 2010;10:22. https://doi.org/10.1186/1471-2288-10-22

137. Feng Y, Parkin D, Devlin NJ. Assessing the performance of the EQ-VAS in the NHS PROMs programme. Quality of Life Research. 2014;23(3):977-89. https://doi.org/10.1007/s11136-013-0537-z

138. McCaffrey N, Kaambwa B, Currow DC, Ratcliffe J. Health-related quality of life measured using the EQ-5D-5L: South Australian population norms. Health and Quality of Life Outcomes. 2016;14(1):133-. https://doi.org/10.1186/s12955-016-0537-0

139. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Quality of Life Research. 1995;4(4):293-307.

140. Munthe-Kaas HM, Glenton C, Booth A, Noyes J, Lewin S. Systematic mapping of existing tools to appraise methodological strengths and limitations of qualitative research: first stage in the development of the CAMELOT tool. BMC Medical Research Methodology. 2019;19(1):113. https://doi.org/10.1186/s12874-019-0728-6

141. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet (London, England). 2001;358(9280):483-8. https://doi.org/10.1016/s0140-6736(01)05627-6

142. Iversen HH, Bjertnaes OA, Skudal KE. Patient evaluation of hospital outcomes: an analysis of open-ended comments from extreme clusters in a national survey. BMJ Open. 2014;4(5):e004848. https://doi.org/10.1136/bmjopen-2014-004848

143. Daddato AE, Dollar B, Lum HD, Burke RE, Boxer RS. Identifying Patient Readmissions: Are Our Data Sources Misleading? Journal of the American Medical Directors Association. 2019;20(8):1042-4. https://doi.org/10.1016/j.jamda.2019.04.028

144. Holland R, Lenaghan E, Harvey I, Smith R, Shepstone L, Lipp A, et al. Does home based medication review keep older people out of hospital? The HOMER randomised controlled trial. BMJ (Clinical research ed). 2005;330(7486):293. https://doi.org/10.1136/bmj.38338.674583.AE

145. Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. Social Science & Medicine. 2018;210:2-21. https://doi.org/10.1016/j.socscimed.2017.12.005

146. Butcher NJ, Monsour A, Mew EJ, Szatmari P, Pierro A, Kelly LE, et al. Improving outcome reporting in clinical trial reports and protocols: study protocol for the Instrument for reporting Planned Endpoints in Clinical Trials (InsPECT). Trials. 2019;20(1):161. https://doi.org/10.1186/s13063-019-3248-0

147. Groves T. Enhancing the quality and transparency of health research. BMJ (Clinical research ed). 2008;337:a718. https://doi.org/10.1136/bmj.a718

148. Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, et al. Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study. BMJ (Clinical research ed). 2017;357:j2490. https://doi.org/10.1136/bmj.j2490

149. Schooler JW. Metascience could rescue the 'replication crisis'. Nature. 2014;515(7525):9. https://doi.org/10.1038/515009a

150. Ioannidis JPA. Why Most Clinical Research Is Not Useful. PLOS Medicine. 2016;13(6):e1002049. https://doi.org/10.1371/journal.pmed.1002049

151. van Walraven C, Bennett C, Forster AJ. Administrative database research infrequently used validated diagnostic or procedural codes. Journal of clinical epidemiology. 2011;64(10):1054-9. https://doi.org/10.1016/j.jclinepi.2011.01.001

152. van Walraven C, Jennings A, Forster AJ. A meta-analysis of hospital 30-day avoidable readmission rates. Journal of Evaluation in Clinical Practice. 2012;18(6):1211-8. https://doi.org/10.1111/j.1365-2753.2011.01773.x

153. Mehtsun WT, Papanicolas I, Zheng J, Orav EJ, Lillemoe KD, Jha AK. National Trends in Readmission Following Inpatient Surgery in the Hospital Readmissions Reduction Program Era. Annals of Surgery. 2018;267(4):599-605. https://doi.org/10.1097/sla.0000000000002350

154. Gupta A, Allen LA, Bhatt DL, Cox M, DeVore AD, Heidenreich PA, et al. Association of the Hospital Readmissions Reduction Program Implementation With Readmission and Mortality Outcomes in Heart Failure. JAMA Cardiol. 2018;3(1):44-53. https://doi.org/10.1001/jamacardio.2017.4265

155. Khera R, Dharmarajan K, Krumholz HM. Rising Mortality in Patients With Heart Failure in the United States: Facts Versus Fiction. JACC Heart Fail. 2018;6(7):610-2. https://doi.org/10.1016/j.jchf.2018.02.011

156. Westert GP, Lagoe RJ, Keskimaki I, Leyland A, Murphy M. An international study of hospital readmissions and related utilization in Europe and the USA. Health Policy (Amsterdam, Netherlands). 2002;61(3):269-78.

157. Wang DE, Wadhera RK, Bhatt DL. Association of Rankings With Cardiovascular Outcomes at Top-Ranked Hospitals vs Nonranked Hospitals in the United States. JAMA Cardiol. 2018. https://doi.org/10.1001/jamacardio.2018.3951

158. Shih T, Dimick JB. Reliability of readmission rates as a hospital quality measure in cardiac surgery. The Annals of Thoracic Surgery. 2014;97(4):1214-8. https://doi.org/10.1016/j.athoracsur.2013.11.048

159. Eide LS, Ranhoff AH, Fridlund B, Haaverstad R, Hufthammer KO, Kuiper KK, et al. Readmissions and mortality in delirious versus non-delirious octogenarian patients after aortic valve therapy: a prospective cohort study. BMJ Open. 2016;6(10):e012683. https://doi.org/10.1136/bmjopen-2016-012683

160. Woods CE, Jones R, O'Shea E, Grist E, Wiggers J, Usher K. Nurse-led post-discharge telephone follow-up calls: A mixed study systematic review. Journal of Clinical Nursing. 2019. https://doi.org/10.1111/jocn.14951

161. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. Mayo Clinic Proceedings. 2002;77(4):371-83. https://doi.org/10.4065/77.4.371

162. Pierce CA, Block RA, Aguinis H. Cautionary Note on Reporting Eta-Squared Values from Multifactor ANOVA Designs. Educational and Psychological Measurement. 2004;64(6):916-24. https://doi.org/10.1177/0013164404264848

163. Kuo K, Shah P, Hiebert B, Love K, Menkis AH, Manji RA, et al. Predictors of survival, functional survival, and hospital readmission in octogenarians after surgical aortic valve replacement. The Journal of Thoracic and Cardiovascular Surgery. 2017. https://doi.org/10.1016/j.jtcvs.2017.05.047

164. Solem JO, Kugelberg J, Stahl E, Olin C. Late cardiac tamponade following open-heart surgery. Diagnosis and treatment. Scandinavian Journal of Thoracic and Cardiovascular Surgery. 1986;20(2):129-31.

165. You SC, Shim CY, Hong GR, Kim D, Cho IJ, Lee S, et al. Incidence, Predictors, and Clinical Outcomes of Postoperative Cardiac Tamponade in Patients Undergoing Heart Valve Surgery. PloS One. 2016;11(11):e0165754. https://doi.org/10.1371/journal.pone.0165754

166. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychology. 2004;23(5):443-51. https://doi.org/10.1037/0278-6133.23.5.443

167. Savage LS, Grap MJ. Telephone monitoring after early discharge for cardiac surgery patients. American Journal of Critical Care. 1999;8(3):154-9.

168. Adams RJ. Improving health outcomes with better patient understanding and education. Risk Manag Healthc Policy. 2010;3:61-72. https://doi.org/10.2147/RMHP.S7500

169. van Achterberg T, Huisman-de Waal GGJ, Ketelaar NABM, Oostendorp RA, Jacobs JE, Wollersheim HCH. How to promote healthy behaviours in patients? An overview of evidence for behaviour change techniques. Health Promot Int. 2011;26(2):148-62. https://doi.org/10.1093/heapro/daq050

170. Bandura A. Health promotion by social cognitive means. Health Education & Behavior. 2004;31(2):143-64. https://doi.org/10.1177/1090198104263660

171. Bandura A. The Explanatory and Predictive Scope of Self-Efficacy Theory. Journal of Social and Clinical Psychology. 1986;4(3):359-73. https://doi.org/10.1521/jscp.1986.4.3.359

172. Sheer VC. A meta-synthesis of health-related self-efficacy instrumentation: problems and suggestions. Journal of Nursing Measurement. 2014;22(1):77-93.

173. Takahashi T, Kumamaru M, Jenkins S, Saitoh M, Morisawa T, Matsuda H. In-patient step count predicts re-hospitalization after cardiac surgery. Journal of Cardiology. 2015;66(4):286-91. https://doi.org/10.1016/j.jjcc.2015.01.006.

174. Lund K, Sibilitz KL, Berg SK, Thygesen LC, Taylor RS, Zwisler AD. Physical activity increases survival after heart valve surgery. Heart. 2016;102(17):1388-95. https://doi.org/10.1136/heartjnl-2015-308827

175. Valenza-Demet G, Valenza MC, Cabrera-Martos I, Torres-Sanchez I, Revelles-Moyano F. The effects of a physiotherapy programme on patients with a pleural effusion: a randomized controlled trial. Clinical Rehabilitation. 2014;28(11):1087-95. https://doi.org/10.1177/0269215514530579

176. Olsen SJ, Schirmer H, Bonaa KH, Hanssen TA. Cardiac rehabilitation after percutaneous coronary intervention: Results from a nationwide survey. European Journal of Cardiovascular Nursing. 2018;17(3):273-9. https://doi.org/10.1177/1474515117737766

177. De Vos C, Li X, Van Vlaenderen I, Saka O, Dendale P, Eyssen M, et al. Participating or not in a cardiac rehabilitation programme: factors influencing a patient's decision. European Journal of Preventive Cardiology. 2013;20(2):341-8. https://doi.org/10.1177/2047487312437057

178. Bandura A. Self-efficacy. The exercise of control. USA: WH Freeman and company; 1997. 525 p.

179. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). European Heart Journal. 2016;37(29):2315-81. https://doi.org/10.1093/eurheartj/ehw106

180. Norwegian Ministry of Health and Care/NMHCS. The Patient and User Rights Act. Law-1999-07-02-63. Oslo, Norway (1999).

181. Choudhry AJ, Younis M, Ray-Zack MD, Glasgow AE, Haddad NN, Habermann EB, et al. Enhanced readability of discharge summaries decreases provider telephone calls and patient readmissions in the posthospital setting. Surgery. 2018. https://doi.org/https://doi.org/10.1016/j.surg.2018.10.014

182. Shebehe J, Hansson A. High hospital readmission rates for patients aged >/=65 years associated with low socioeconomic status in a Swedish region: a cross-sectional study in primary care. Scandinavian Journal of Primary Health Care. 2018:1-8. https://doi.org/10.1080/02813432.2018.1499584

183. World Health Organization (WHO). Transitions of Care: Technical Series on Safer Primary Care Geneva: 2016 [Licence: CC BY-NC-SA 3.0 IGO]. Available from: https://apps.who.int/iris/bitstream/handle/10665/252272/9789241511599-eng.pdf;jsessionid=6A7C5DE361A9661C6AF703F3CC23AB86?sequence=1

184. Norwegian Ministry of Health and Care Services/NMHCS. Report no.47 to the Storting (2008-2009): The Coordination Reform, Proper treatment -at the right place and right time. Oslo, Norway 2009.

185. Ambugo EA, Hagen TP. Effects of introducing a fee for “inpatient overstays” on the rate of death and readmissions across municipalities in Norway. Social Science & Medicine. 2019. https://doi.org/https://doi.org/10.1016/j.socscimed.2019.04.006

186. Directorate of Health. Safe discharge. Safe discharge with the patient as an equal party. Oslo 2019 [Available from: https://www.pasientsikkerhetsprogrammet.no/om-oss/innsatsområder/trygg-utskrivning

187. D'Agostino RS, Jacobs JP, Badhwar V, Paone G, Rankin JS, Han JM, et al. The Society of Thoracic Surgeons Adult Cardiac Surgery Database: 2016 Update on Outcomes and Quality. The Annals of Thoracic Surgery. 2016;101(1):24-32. https://doi.org/10.1016/j.athoracsur.2015.11.032

188. Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? JAMA. 2000;283(20):2701-11.

189. Andersen,VB. An offer of telephone support to aortic valve surgery patients. Who uses the offer and which symptoms do the patients experience after hospital discharge? [Master thesis]. University of Oslo. Oslo, Norway; 2019.

190. Williams JB, Alexander KP, Morin JF, Langlois Y, Noiseux N, Perrault LP, et al. Preoperative anxiety as a predictor of mortality and major morbidity in patients aged >70 years undergoing cardiac surgery. The American Journal of Cardiology. 2013;111(1):137-42. https://doi.org/10.1016/j.amjcard.2012.08.060