

**Comparison of two multidisciplinary rehabilitation programmes in patients with chronic low back pain
- A randomised controlled trial**

PhD dissertation

Anne Mette Schmidt

Health
Aarhus University

2020

Supervisors

Thomas Maribo, PT, PhD, Associate professor
Department of Public Health, Aarhus University, Denmark
DEFACTUM, Central Denmark Region, Denmark

Trine Bay Laurberg, MD, PhD, Associated professor
Department of Rheumatology, Aarhus University Hospital, Denmark
Sano, Denmark

Berit Schiøttz-Christensen, MD, PhD, Professor
Spine Centre of Southern Denmark, University Hospital Lillebælt, Denmark
Institute for Regional Health Research, University of Southern Denmark, Denmark

Assessment Committee

Associate professor Dorte Rytter (Chairman)
Department of Public Health, Aarhus University, Denmark

Professor Sarah Dean
College House, University of Exeter, St Luke's Campus, United Kingdom

Clinical professor Janus Laust Thomsen
Department of Clinical Medicine, Center for General Practice, Aalborg University, Denmark

Acknowledgements

Since October 2015, I have been employed at Sano, Denmark to conduct this research project in close collaboration with DEFACTUM, Central Denmark Region and Department of Public Health, Aarhus University. This PhD dissertation comprises the scientific work during this period. Numerous people deserve a mention for their contribution to this research project or for their support and engagement along my way.

I am truly grateful to my three supervisors. Thomas Maribo, you were the reason I started this academic journey. It has been a pleasure being raised under your wing. I have learned a lot from our many conversations, discussions and reflections where we left no stone unturned. I am grateful for the well-balanced way you have challenged, supported and encouraged me in my efforts. In addition, I have enjoyed our informal talks and laughs during lunch breaks, and when attending courses and congresses, not least when travelling to congresses in South Africa. Thank you for hosting me at DEFACTUM, I have indeed profited from being a part of the academic community. I owe Trine Bay Laurberg thanks for referring the participants, and for the efforts you made to make this research project alive in the clinical setting. But most of all, I owe you huge thanks for your respectful way of supervising me, and for insisting that “livet er en morgengave”. Berit Schjøttz-Christensen, you have provided this research project with your clinical experience and your useful helicopter perspective – thank you for that; it has been really valuable.

A special thank you to the Danish Rheumatism Association who gave me the opportunity to conduct this research project. This also applies more specifically to the management at Sano; Karin Jul Pedersen, Annie Abildtrup, and especially Kirstine Bak Gulstad. Thank you for believing in me from the very beginning, for providing the clinical setting and the participants, and for doing some magic whenever needed. This research project could not have been performed without the outmost willing participants, the exceptionally helpful administrative staff, especially Bente Munch Nilsson, and the dedicated providers at Sano who pulled together – thank you for your efforts. A special thank you to Camilla Vejlgård for your never-ending support, and to Gorm Høi Jensen, Johan Riisgaard Laursen, and Brian Clausen for constructive and entertaining collaboration.

My colleagues at DEFACTUM, I am grateful for the friendly and inspiring research environment you have created – I will miss you. A special thank you to Inger Hornbech for your never-ending technical and practical support. In particular, I am very grateful for my trusty wingmen, and our many café latte breaks. Charlotte Ibsen, you welcomed me the very first day at DEFACTUM, and since then you have been a (nearly) daily lifeline. Thank you for our many conversations about everything and nothing, and for your good company all over the world – it is wonderful to know

you. Line Thorndal Moll, you have made a strong active academic and personal contribution to my process. It has made an indescribable difference to me that I met you. Finally, Charlotte Maria Jensen and Cecilie Nørby Lyhne, thank you for growing with me in “Spirekassen” – I enjoyed to share office and life with you for a while.

Along this research project, I have collaborated with some external persons who need a special thank you. Lianne Wood for inviting me to and planning my research exchange at Keele University. Nadine Foster for hosting me at Keele University. I am humbled and honored that you spent your valuable time on me, and my research project - thank you for the many enlightening discussions. Helle Terkildsen Maindal for your contribution to my first paper - it was a pleasure to collaborate with you. You created the overview I needed by introducing me to the world of complex interventions. Jens Søndergaard Jensen, for your ability to explain difficult statistical knowledge in a patient and understandable way. I also owe a special thank you to Dorte Rytter, Sarah Dean, and Janus Laust Thomsen for the time and effort you spent on the thorough evaluation of my work.

Looking back, I owe my gratitude to a couple of persons. Oluf Schiøtz for being the mentor that provoked, motivated and supported me to conduct my first research project. Hanne Vebert Olesen and Vibsen Bregnballe you were there too – thanks for still being in the background. Martin Vesterby, for many years you have provided priceless support when I most needed it – I will not forget that.

A big thank you to close friends, siblings and the rest of my family for showing interest in me and my research project; but most of all for being there. My deepest thanks go to my parents for always supporting and encouraging me in any way at any time – I am grateful for all of you. My two four-legged friends, Nugga and Coco, thank you for being good company at my home office and for forcing me out in the fresh air. Last but not least, Henrik, my “livsfordobler”, I cannot thank you enough for your unending positivity through thick and thin. Together with our two boys, Nicolai and Casper, you are second to none!

Anne Mette Schmidt, June 2020

Financial support

Sano

Aarhus University

The Danish Rheumatism Association

Familien Hede Nielsens Fond

Christian og Otilia Brorsons Rejselegat

Lundbeckfonden

Augustinus Fonden

List of scientific papers

The PhD dissertation is based on the following papers:

Study 1

Schmidt, A.M., Maindal H.T., Laurberg, T.B., Schiøttz-Christensen, B., Ibsen, C., Gulstad, K.B., Maribo T.

The Sano study: justification and detailed description of a multidisciplinary biopsychosocial rehabilitation programme in patients with chronic low back pain.

Clinical Rehabilitation 2018 Nov;32(11):1431-1439. doi: 0.1177/0269215518780953.

Study 2

Schmidt, A.M., Schiøttz-Christensen, B., Foster, N.E., Laurberg, T.B., Maribo T.

The effect of an integrated multidisciplinary rehabilitation programme alternating inpatient interventions with home-based activities for patients with chronic low back pain: a randomized controlled trial.

Clinical Rehabilitation 2020 doi: 10.1177/0269215519897968

Study 3

Schmidt, A.M., Laurberg, T.B., Moll, L.T., Schiøttz-Christensen, B., Maribo T.

The effect of an integrated multidisciplinary rehabilitation programme for patients with chronic low back pain: long-term follow up of a randomised controlled trial.

Submittet to Clinical Rehabilitation March 2020.

The studies are found in Appendices 1-3.

List of abbreviations

CERT: Consensus on Exercise Reporting Template

CI: Confidence Interval

CLBP: Chronic Low Back Pain

CONSORT: Consolidated Standards of Reporting Trials

CReDECI 2: Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline

EQ-5D 5L: EuroQol-5 Domain 5-level

GRIPP-2: Guidance for Reporting Involvement of Patients and the Public 2

HRQoL: Health Related Quality of Life

ICD-10: 10th revision of the International Statistical Classification of Diseases and Related Health Problems

ICF: International Classification of Functioning, Disability and Health

ITT: Intention-To-Treat

LBP: Low Back Pain

MDI: Major Depression Inventory

MRC: Medical Research Council

NRS: Numeric Rating Scale

ODI: Oswestry Disability Index

PPI: Patient and Public Involvement

PRECIS-2: Pragmatic Explanatory Continuum Indicator Summary (second version)

PSEQ: Pain Self-Efficacy Questionnaire

RCT: Randomised Controlled Trial

SD: Standard Deviation

TIDieR: Template for Intervention Description and Replication

WHO: World Health Organization

Overview of figures and tables

Figures

Figure 1. The Medical Research Council's four-stage model with the 12 related elements.

Figure 2. Flow-chart illustrating the updated literature search and assessment in February 2016.

Figure 3. Logic model of the integrated programme.

Figure 4. The three key functions and subdivisions of process evaluation addressed (column 1), process research questions (column 2) and methods used for the process evaluation (column 3).

Figure 5. The PRECIS-2 wheel (Study 2).

Figure 6. Comparison of the integrated programme and the existing programme.

Figure 7. The interrelationship between the 10 key components.

Figure 8. Flow chart of patients in Studies 2 and 3.

Figure 9. Mean change in back-specific disability measured by the Oswestry Disability Index during the 12-month follow-up period, including the four identical measurement time points.

Figure 10. Mean changes in secondary outcomes during the 12-month follow-up period, including the four identical measurement time points.

Figure 11. Mean changes over the 12-month follow-up period, including all nine measurement time points (t0-t8).

Tables

Table 1. Overview of methods used in the research process according to the four stages and the 12 related elements described in the Medical Research Council's guidance.

Table 2. Overview of the four RCTs comparing two or more rehabilitation programmes having disability as the primary outcome.

Table 3. Key components and the related 38 clinical activities in the rehabilitation programmes.

Table 4. Key components, subdivision and corresponding explanation of a process evaluation.

Table 5. Overview of design, applied reporting guidelines, outcome domains and outcome measures, follow up and primary statistical analyses.

Table 6. Items and item categories of the TIDieR checklist along with a brief item description.

Table 7. Overview of domains, rationale for design choices and the corresponding score given by applying PRECIS-2 to this RCT.

Table 8. Differences and similarities between the integrated programme and the existing programme.

Table 9. The activity sheet developed to describe the 38 clinical activities; using aqua gymnastics as an example.

Table 10. Characteristics of the population at baseline. Columns 1 and 2 show patients randomised (n=165), and columns 3 and 4 show patients providing 6-month follow-up data (n=139).

Table 11. Summary of 6-month follow-up data on primary and secondary outcomes. Between-group and within-group change; complete case analysis.

Table 12. Summary of 12-month follow-up data on primary and secondary outcomes. Between-group and within-group change; intention-to-treat analysis.

Appendices

Appendix 1

Study 1

Supplementary file: Appendix 8

Appendix 2

Study 2

Supplementary file 1: Appendix 5

Supplementary file 2: Figure 4 in this dissertation

Appendix 3

Study 3

Supplementary file 1: Appendix 7

Appendix 4

Permission of the UK Medical Research Council to reproduce Figure 1.

Appendix 5

A completed GRIPP2 short form

Appendix 6

Statistical analysis plan – Study 2

Appendix 7

Statistical analysis plan – Study 3

Appendix 8

Activity sheets describing each of the 38 clinical activities

Appendix 9

Graphical depiction comparing the integrated programme and the existing programme

Appendix 10

Examples process research question 1 and 2

Table of contents

Summary	11
Dansk resumé	13
1. Background	15
1.1 The burden of low back pain	15
1.2 The biopsychosocial approach	15
1.3 Management of patients with CLBP	16
1.4 Multidisciplinary rehabilitation - a complex intervention.....	17
2. The integrated programme	19
2.1 The MRC's guidance applied to the research process	20
2.1.1 Development	23
2.1.2 Feasibility/piloting	33
2.1.3 Evaluation	33
2.1.4 Implementation	38
2.2 Patient and public involvement.....	39
3. Aims	40
Aim 1 (Study 1).....	40
Aim 2 (Study 2).....	40
Aim 3 (Study 3).....	40
4. Methods	41
4.1 Study 1	42
4.2 Studies 2 and 3	43
4.2.1 Choice of design and its explanation.....	43
4.2.2 Setting and participants	45
4.2.3 Randomisation and blinding	46
4.2.4 Interventions	46
4.2.5 Baseline characteristics, outcome domains, outcome measures, and other variables	49
4.2.6 Data collection.....	51
4.2.7 Statistical analyses.....	52
4.2.8 Ethical approval and trial registration	53
5. Results	54
5.1 Study 1	54
5.1.1 Item 1. Brief name	54
5.1.2 Item 2. Why.....	54
5.1.3 Item 3. What (materials)	54
5.1.4 Item 4. What (procedures)	54
5.1.5 Item 5. Who provided	55
5.1.6 Item 6. How	55
5.1.7 Item 7. Where.....	56
5.1.8 Item 8. When and how much	56

5.1.9 Item 9. Tailoring	56
5.1.10 Item 10. Modifications	56
5.1.11 Item 11. How well (planned)	56
5.1.12 Item 12. How well (actual)	56
5.2 Patient flow and baseline characteristics in Studies 2 and 3	57
5.3 Results Study 2	60
5.4 Results Study 3	62
5.5 Process evaluation	66
5.5.1 Process research question 1	66
5.5.2 Process research question 2	66
5.5.3 Process research question 3	66
5.5.4 Process research question 4	67
5.5.5 Process research question 5	68
5.5.6 Process research question 6	68
6. Discussion	69
6.1 Reflections on the integrated programme including Study 1.....	69
6.1.1 Development stage.....	69
6.1.2 Feasibility/piloting	70
6.1.3 Evaluation stage.....	71
6.2 PPI.....	73
6.3 Main results from the RCT compared with the literature	73
6.4 Main findings from the process evaluation and their relationship to the main results from the RCT	75
6.4.1 Process research question 1.....	75
6.4.2 Process research question 2.....	76
6.4.3 Process research question 3.....	76
6.4.4 Process research question 4.....	76
6.4.5 Process research question 5.....	76
6.4.6 Process research question 6.....	77
6.5 Strengths and limitations in Studies 2 and 3	77
6.6 External validity	79
7. Conclusion.....	81
8. Implications	82
8.1 For clinical practice.....	82
8.2 For research.....	82
8.3 For public health	83
9. References.....	84
10. Appendices	90

Summary

Background

Low back pain (LBP) is a substantial burden worldwide constituting the most frequent cause of years lived with disability and having an enormous impact on individuals affected, their relatives and society in general. Reasons for the development and persistence of chronic LBP (CLBP) and related disability include a complex interaction between multiple biopsychosocial factors. Thus, a biopsychosocial approach is recommended when guiding management of patients with CLBP, and therefore, recent evidence-based clinical practice guidelines in the CLBP area recommend the biopsychosocial model to inform clinical practice. Those evidence-based clinical practice guidelines recommend multidisciplinary rehabilitation as one of the possible second-line treatment options for patients with CLBP. However, the optimal dose, content or delivery of such multidisciplinary rehabilitation programmes remains unknown. Against this backdrop, an integrated programme was developed, feasibility-tested and evaluated. Thus, the aims of this dissertation were: 1) to justify and describe the integrated programme, 2) to compare the effectiveness of the integrated programme with an existing programme at the 6-month follow up, and 3) to compare the effectiveness of the integrated programme with an existing programme at the 12-month follow up.

Methods

In Study 1, the Template for Intervention Description and Replication (TIDieR) checklist was used as a structural framework for the description of the integrated rehabilitation programme. As part of the description, the Medical Research Council's guidance, 'Developing and evaluating complex interventions', was used as a framework to justify the integrated rehabilitation programme. The process was underpinned by patient and public involvement.

In Studies 2 and 3, a single-centre, pragmatic, two-arm parallel, randomised controlled trial (RCT) was conducted in a Danish rheumatology inpatient rehabilitation centre. Adults with CLBP for more than 12 months were randomly allocated, using computer-generated randomisation (1:1 ratio) to the integrated programme or to the existing programme. The integrated programme comprised a pre-admission day, 2 weeks at home preparing for the next inpatient stay, 2 weeks as an inpatient followed by home-based activities plus two 2-day inpatient booster sessions, and a 6-month follow up visit, whereas the existing programme comprised a 4-week inpatient stay, and a 6-month follow up visit. Patient-reported outcomes were collected at baseline, 6-month follow up (Study 2) and 12-month follow up (Study 3). The primary outcome was disability measured by the Oswestry Disability Index (ODI). Secondary outcomes included pain intensity (Numerical Rating Scale), pain self-efficacy (Pain Self-Efficacy Questionnaire), health-related quality of life (EQ-5D),

and depression (Major Depression Inventory). In Study 2, a complete case analysis was performed using linear regression, and in Study 3, analysis was by intention-to-treat, using linear mixed models.

A process evaluation using both qualitative and quantitative methods was nested into the RCT.

Results

In Study 1, the integrated programme was justified and described in detail. The inpatient part of the programme consisted of 38 clinical activities, some of them delivered more than once. The 38 clinical activities were described in an activity sheet developed for this purpose, combining five items from the TIDieR checklist.

In Studies 2 and 3, between February 2016 and August 2018, 303 patients were assessed for eligibility. Of them, 165 patients (mean age: 50 years (SD 13) with mean ODI score of 42 (SD 11)) were randomised, 82 patients to the integrated rehabilitation programme and 83 patients to the existing programme. Baseline characteristics were comparable between programmes.

In Study 2, 139 patients (70 from the integrated programme and 69 from the existing programme) provided 6-month follow-up data. The between-group difference in the ODI score was -0.28 (95% CI: -4.02 ; 3.45), which was neither statistically nor clinically significant. No significant differences were found in the secondary outcomes.

In Study 3, the between-group difference in the ODI score was -0.53 (95% CI: -4.08 ; 3.02) at the 12-month follow up, being neither statistically nor clinically significant. No significant differences were found in the secondary outcomes.

Conclusion

The integrated programme for patients with CLBP was justified and described. The intervention description was used for structuring and standardising the content and delivery of the integrated programme in the RCT. The integrated programme did not lead to improved back-specific disability or other outcomes for patients with CLBP when compared with the existing programme at 6-month and 12-month follow up. The null effect may be due to the identical comprehensive nature of the two rehabilitation programmes being compared, the continuous organisational challenges attenuating any difference between the two rehabilitation programmes or a number of other reasons.

Dansk resumé

Baggrund

Lænderygsmarter er en meget udbredt helbredstilstand på verdensplan. Samtidig er det den hyppigste årsag til leveår med påvirket funktionsevne, og det har en væsentlig indflydelse på personen, der lider af lænderygsmarterne, deres pårørende og samfundet generelt. Et komplekst sammenspil mellem biopsykosociale faktorer anses som årsag til udvikling og kronificering af lænderygsmarter (kroniske lænderygsmarter) og den afledte nedsatte funktionsevne. Af samme grund bygger de seneste evidensbaserede kliniske retningslinjer inden for dette område også på en anbefaling af en biopsykosocial tilgang i klinisk praksis. I de evidensbaserede kliniske retningslinjer anbefales tværfaglig rehabilitering som en af de mulige anden valgs behandlinger til patienter med kroniske lænderygsmarter, der ikke responderer på førstevalgs behandling. Den optimale dosis, det optimale indhold og den optimale opbygning af et sådant tværfagligt rehabiliteringsprogram er fortsat ukendt. På baggrund heraf blev et integreret program udviklet, testet og evalueret, hvorfor formålet med denne afhandling var: 1) at underbygge og beskrive det integrerede program, 2) at undersøge effekten af det integrerede program sammenlignet med et eksisterende program efter 6 måneder, og 3) at undersøge effekten af det integrerede program sammenlignet med et eksisterende program efter 12 måneder.

Metode

I Studie 1 blev tjeklisten the Template for Intervention Description and Replication (TIDieR) brugt til at strukturere beskrivelsen af det integrerede program. Som en del af beskrivelsen blev også the Medical Research Council's guidance om udvikling og evaluering af komplekse interventioner brugt som rammeværk i forhold til at understøtte rationalet for det integrerede program. Processen blev underbygget ved at involvere patienter og øvrige samarbejdspartnere.

Studie 2 og 3 var et pragmatisk randomiseret kontrolleret studie (RCT) med to parallelle grupper udført på et dansk reumatologisk rehabiliteringscenter. Voksne med kroniske lænderygsmarter i mere end 12 måneder blev ved hjælp af computergenereret randomisering tilfældigt fordelt (1:1 ratio) til det integrerede program eller det eksisterende program. Det integrerede program bestod af en forundersøgelsesdag, 2 uger hjemme med forberedelse til det næste forløb, 2 ugers indlæggelse efterfulgt af aktiviteter i hjemmet plus 2 gange booster sessioner af to dages varighed. Programmet blev afsluttet med opfølgning efter 6 måneder. Det eksisterende program bestod af fire ugers indlæggelse og blev afsluttet med opfølgning efter 6 måneder. Patientrapporterede målemetoder blev anvendt til at indsamle data ved baseline, samt 6 måneder (Studie 2) og 12 måneder (Studie 3) efter interventionsstart. Det primære outcome var funktionsevne målt med Oswestry Disability Index (ODI). Sekundære outcomes var smerte intensitet (Numerisk Rang

Skala), smerte self-efficacy (Pain Self-Efficacy Questionnaire), helbredsrelateret livskvalitet (EQ-5D) og depression (Major Depression Inventory). I Studie 2 blev en complete-case analyse udført ved hjælp af lineær regression. I studie 3 blev en intention-to-treat analyse udført ved hjælp af lineær mixed models.

En procesevaluering der brugte både kvalitative og kvantitative metoder blev gennemført sideløbende med RCT'et.

Resultater

Rationalet for det integrerede program blev underbygget, og programmet blev beskrevet i detaljer (Studie 1). Programmet bestod af 38 forskellige kliniske aktiviteter, hvoraf nogle af dem blev leveret mere end en gang. De 38 kliniske aktiviteter blev beskrevet i et aktivitetsskema udviklet til samme formål og bestod af fem punkter fra TIDieR tjeklisten.

Studierne 2 og 3 blev udført i perioden fra februar 2016 til august 2018. 303 patienter blev undersøgt med henblik på om de opfyldte in- og eksklusionskriterierne. Af dem blev 165 patienter randomiseret (mean alder var 50 år (SD 13) og mean ODI score var 42 (SD 11)), hvoraf 82 patienter blev tildelt det integrerede program og 83 patienter det eksisterende program. Demografiske og kliniske karakteristika var sammenlignelige mellem de to programmer ved baseline.

I Studie 2 bidrog 139 patienter (70 fra det integrerede program og 69 fra det eksisterende program) med data. Forskellen mellem grupperne i ODI score var $-0,28$ (95% CI: $-4,02$; $3,45$) efter 6 måneder, hvilket hverken var statistisk eller klinisk signifikant. Der blev ikke fundet signifikant forskel mellem grupperne på nogle af de sekundære outcomes.

I Studie 3 var forskellen mellem grupperne i ODI score $-0,53$ (95% CI: $-4,08$; $3,02$) efter 12 måneder, hvilket hverken var statistisk eller klinisk signifikant. Der blev ikke fundet signifikant forskel mellem grupperne på nogle af de sekundære outcomes.

Konklusion

Et integreret program til patienter med kroniske lænderygsmarter blev underbygget og beskrevet. Beskrivelsen af interventionen blev brugt til at strukturere og standardisere indholdet og opbygningen af det integrerede program i RCT'et. Sammenlignet med deltagelse i det eksisterende program medførte deltagelse i det integrerede program ikke forbedret funktionsevne eller andre outcomes hos patienter med kroniske lænderygsmarter. Dette gjaldt både sammenligningen foretaget ved 6 og 12 måneder. Flere årsager kan forklare den manglende forskel, herunder at de to programmer var for ens og at de organisatoriske udfordringer medførte at forskellen mellem de to programmer blev udvisket.

1. Background

1.1 The burden of low back pain

Low back pain (LBP) is a highly prevalent health condition affecting more than 540 million people worldwide (1). Moreover, it constitutes the most frequent cause of years lived with disability (1). LBP has substantial impacts on individuals, families, communities and health care systems, as well as its societal costs due to work absenteeism and its associated effects on the social support system (2, 3).

LBP is defined by the location of pain on the posterior aspect of the body below the costal margin and above the inferior gluteal folds, and it can be accompanied by pain in one or both legs (sciatica) (2-4). The prevalence and burden of LBP increases with age (2, 5), and LBP is more common in women than in men (2). The majority of patients have non-specific LBP, meaning that no underlying pathology or cause can be identified, and often concurrent pain is present elsewhere in the body (2, 3). LBP with a pain duration persisting for more than three months is often defined as chronic LBP (CLBP) (2-4). A minority of an estimated 10-15% of patients with LBP develops CLBP (2, 3, 6), and they comprise the group of patients with LBP bearing the greatest proportion of the disease burden (3, 6-8).

Reasons for the development and persistence of LBP and related disability include a complex interaction between biophysical factors, genetic factors, psychological factors, social factors, lifestyle factors, and comorbidities (2-4, 9, 10). The contribution of these different factors varies from individual to individual (8), and over time (6). Besides pain and disability, patients with LBP often experience psychological consequences with signs of anxiety and depression, and reduced quality of life (2, 3, 7). Furthermore, negative consequences can be experienced for social life including leisure, and work life (2, 3, 7).

In light of the biopsychosocial influence on the development and persistence of LBP and related disability, some elaboration on the biopsychosocial approach, that allows for and incorporates all the biophysical, psychological, and social factors, is needed.

1.2 The biopsychosocial approach

In 1977, George Engel described the biopsychosocial model (11), and 10 years later, Gordon Waddell suggested the model as a new theoretical framework for the understanding and management of LBP (12). This contributed to a fundamental change in the existing approach to LBP (12, 13), and nowadays, the biopsychosocial approach is still widely recognised as the dominant one in the understanding and management of LBP and related disability (2, 3, 6, 8, 13-15).

Years later, in 2001, the biopsychosocial model formed the basis of the International Classification of Functioning, Disability and Health (ICF), published by the World Health Organisation (WHO) (16). ICF is the international standard to describe and measure health and disability; it focuses on functioning and disability, and how these are influenced by contextual factors (16), and it seeks to capture how people with a specific health condition function in their daily life (17).

More recently, leading proponents, for example the alliance behind the third edition of the White Book on Physical and Rehabilitation Medicine in Europe, and four editorials in the journal *Clinical Rehabilitation*, have endorsed the biopsychosocial model as a rational approach to rehabilitation (15, 17-20).

In the field of LBP, the biopsychosocial approach is recommended when guiding management of patients in view of the key contributors to LBP and related disability (2, 3, 6, 14, 21). Therefore, recent evidence-based clinical practice guidelines in the CLBP field recommend the biopsychosocial approach to inform clinical practice (22-24).

1.3 Management of patients with CLBP

The aforementioned recent evidence-based clinical practice guidelines endorse education and self-care (including advice to remain active and patient education), and non-pharmacological therapy (including exercise therapy and cognitive behavioural therapy) as first-line treatment options for patients with CLBP (14, 22, 23). Multidisciplinary rehabilitation is recommended as one of the possible second-line treatment options for patients who experience disability due to CLBP and who are not responding to first-line treatments (3, 14, 22-24).

The recommendations about multidisciplinary treatment in recent evidence-based clinical practice guidelines (14, 22-24) are based on research including a Cochrane systematic review and meta-analysis aimed at assessing the long term effects of multidisciplinary rehabilitation for patients with CLBP (7). In total, 41 RCTs including 6858 participants comparing multidisciplinary rehabilitation with usual care, physical treatment, surgery and waiting list, respectively, were included (7). In total, 16 RCTs provided moderate quality evidence that multidisciplinary rehabilitation was more effective than usual care in terms of decreasing pain and disability (7). Furthermore, 19 RCTs provided low quality evidence that multidisciplinary rehabilitation was more effective compared with physical treatments in decreasing pain and disability (7). The effect lasted longer than 1 year regardless of the comparator being either usual care or physical treatments (7). Additionally, 12 RCTs comparing two or more multidisciplinary rehabilitation programmes were identified but not included in a comparative effectiveness analysis, as this did not inform the main research question of the review (7). Thus, the optimal dose, content or delivery of a multidisciplinary rehabilitation programme remains unknown (7).

1.4 Multidisciplinary rehabilitation - a complex intervention

Several formal definitions of the concept rehabilitation exist (25, 26). According to the World Report on Disability, produced jointly by WHO and the World Bank, rehabilitation is defined as “*a set of measures that assist individuals who experience, or are likely to experience, disability to achieve and maintain optimal functioning in interaction with their environments*” (27).

Rehabilitation is a problem-solving process (17, 28, 29) aimed at optimising functioning and minimising disability of people with a given health condition (17, 20, 27, 30) enabling them to continue with their lives (6). The process of rehabilitation is described as dynamic and iterative, encompassing the following four actions: 1) assessment, 2) goal-setting, 3) intervention, and 4) evaluation (15, 17, 19, 27, 29, 31).

When used in this dissertation, multidisciplinary rehabilitation is defined as detailed in the aforementioned Cochrane review (7). The premise is that multidisciplinary rehabilitation is characterised in accordance with the biopsychosocial approach, the intervention involves a physical component and a psychological and/or social/work targeted component, it is delivered by more than one profession, the intensity and approach varies, and it can be provided in inpatient- or outpatient settings (7).

In general, health care interventions cover a wide range of complexity and there is no clear cut-off between simple and complex interventions (32). However, when imagining health care interventions as a hierarchy of complexity, management of patients with disability is near the top. Rehabilitation interventions are thought of as complex interventions (15, 19, 28, 33) because they:

- Involve several providers from different professions within different teams, organisations, and sectors (17, 19, 27, 31, 33, 34)
- Usually consist of several or many primary interventions aimed at different parts of the biopsychosocial approach (15, 17, 27-29, 33)
- Use a person-centered approach by tailoring the rehabilitation intervention to the needs, goals and preferences of the patient (17, 27, 28, 33, 35)
- Pay attention to and acknowledge the context (15, 17, 28, 33, 35)

Several of the abovementioned aspects of rehabilitation are similar with dimensions defining complex interventions (32).

Summing up, the burden of LBP is indisputable, and it is well established that the development and persistence of LBP and associated disability is attributed to a complexity of biopsychosocial factors. Multidisciplinary rehabilitation is a complex intervention targeting the wide range of modifiable factors known to contribute to CLBP and it is usually based on the widely accepted biopsychosocial approach; it is recommended as a second-line treatment for patients with CLBP. However, the optimal dose, content or delivery of such a multidisciplinary rehabilitation programme is unknown. Against this backdrop, a new multidisciplinary rehabilitation programme, named *the integrated programme* was developed, feasibility-tested and evaluated in a randomised controlled trial (RCT).

2. The integrated programme

The development, feasibility-testing and evaluation of the integrated programme was based on the Medical Research Council's (MRC) guidance on developing and evaluating complex interventions (32, 36). The MRC's guidance was used as the underlying methodological framework as it offered a systematic four-stage model to structure the research process. Figure 1 illustrates the model comprising four stages, and 12 related elements (32). The process is seen as dynamic and iterative rather than linear or cyclical (32).

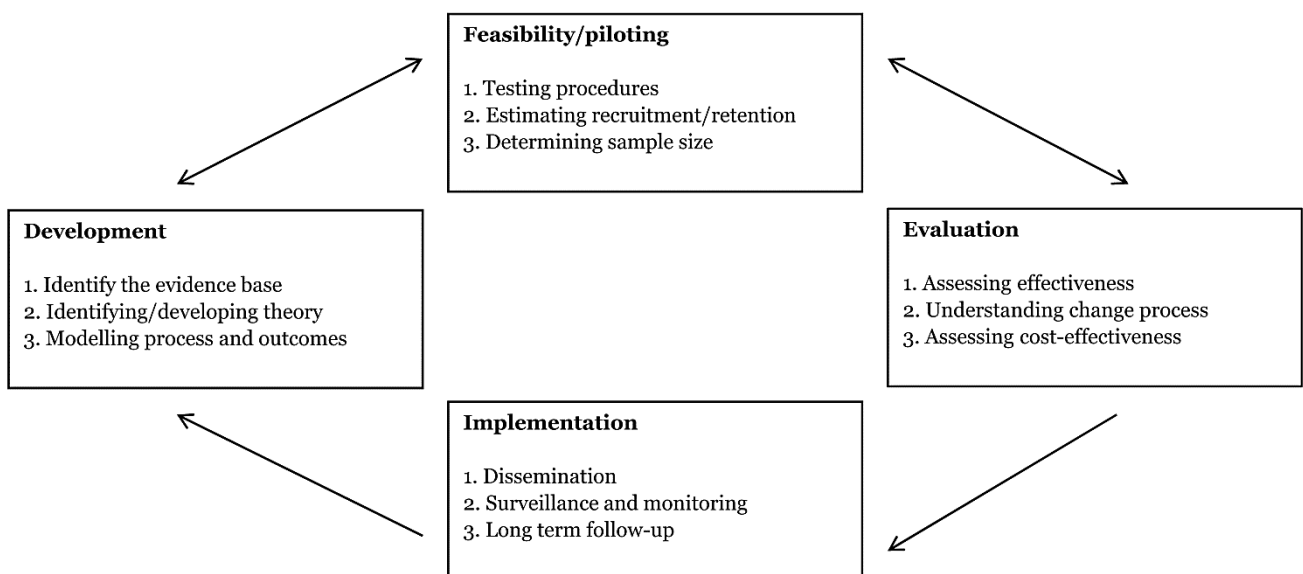


Figure 1. The Medical Research Council's four-stage model with the 12 related elements (32). Reproduced with permission of the UK Medical Research Council (Appendix 4).

In the first part of this chapter, the MRC's guidance (32, 36) in combination with inspiration from Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline (CReDECI 2) (37) was used to structure and ensure comprehensive reporting. In the last part of this chapter, the Patient and Public Involvement (PPI) process was reported with inspiration from Guidance for Reporting Involvement of Patients and the Public 2 short form (GRIPP2) (38).

2.1 The MRC's guidance applied to the research process

To improve the transparency and create an overview, the methods used at each of the four stages and the 12 related elements following the MRC's guidance (32, 36) are outlined in Table 1. The results of this research process are presented subsequently and follow the same structure.

Table 1. Overview of methods used in the research process according to the four stages and the 12 related elements described in the Medical Research Council’s guidance (32, 36).

Stage	Elements	Methods
Development	Identifying the evidence base	A systematic approach was used to identify a Cochrane systematic review and meta-analysis (7). An updated literature search in CENTRAL, Medline, Embase, PsycINFO and CINAHL was performed in February 2016. Clearly ineligible trials were excluded based on title and abstract, and the remaining trials were reviewed in full text. The updated literature search and assessment were identical to criteria used in the Cochrane review (7).
	Identifying/developing theory	When conducting an RCT within the field of CLBP and rehabilitation, the biopsychosocial approach was considered essential. The Chronic Care Model was identified from discussions with other researchers.
	Modelling processes and outcomes	A six-step modelling scenario consisted of the following steps (39): Step 1 Installing a project team and formulating the key objectives: meetings. Step 2 Getting consensus on the components: meetings with patients and stakeholders (includes providers, administrative and management staff). Step 3 Clustering of clinical activities into key components and building a process flow: meetings with providers and management staff. Step 4 Organising the process and allocating resources: meetings with administrative and management staff. Step 5 Describing key interventions: existing descriptions were updated by providers and the project team. Step 6 Translation into a set of process and outcome indicators: literature, focus groups with patients using a semi-structured interview guide, and providers voting for their two favourite outcomes. The final decision about the outcomes was taken by the PhD candidate. The outcome measures were primarily chosen on the basis of the literature. Three providers with different backgrounds chose their preferred psychological outcome measure amongst Pain Self-Efficacy Questionnaire, Fear Avoidance Questionnaire and Pain Catastrophising Scale.

Stage	Elements	Methods
Feasibility/ Piloting	Testing procedures	Opinions about whether the integrated programme was appropriate, feasible and acceptable were gained from meetings and focus groups using a semi-structured interview guide with former patients in the existing programme, an internal pilot group (n = 3) (40), and stakeholders. Two patients were involved in the development of the participant information and informed consent form. Two patients tested the database set-up in terms of time consumption and layout.
	Estimating recruitment/ retention	Patient flow, recruitment strategy and duration of inclusion were discussed at meetings with administrative and management staff. Focus groups with patients using a semi-structured interview guide were used to learn about barriers to recruitment, and the proportion of eligible patients willing to participate.
	Determining sample size	Sample size was calculated based on the literature and a feasibility-test including 12 patients completing the Oswestry Disability Index when they started and ended the existing programme (Chapter 4).
Evaluation	Assessing effectiveness	The design was chosen based on meetings involving members of the project team and recommendations from a statistician.
	Understanding change process	No exact recipe on how to perform a process evaluation exists (41-43), but it is recommended to: 1) plan, 2) design and conduct, 3) analyse, and 4) report (41, 42). Background, purpose and methods applied will be presented in Section 2.1.3.2 as more elaboration was needed regarding the process evaluation.
	Assessing cost-effectiveness	This was outside the scope of this dissertation.
Implementation	Dissemination	Publication/submission of three peer-reviewed papers. Presentations at national and international meetings and conferences. Results from the process evaluation and the evaluation of the RCT were fed back to stakeholders at meetings. Patients included in the RCT will be emailed the results in layman language formulated in cooperation with a patient representative.
	Surveillance and monitoring	Design of a database and obtained informed consent allowing for 12-month follow up served the purpose of surveillance, monitoring and long-term follow up with the possibility of prolongation
	Long-term follow-up	

2.1.1 Development

2.1.1.1 Identifying the evidence base

A Cochrane systematic review and meta-analysis on multidisciplinary biopsychosocial rehabilitation for CLBP (7) together with two additional RCTs (44, 45) were identified (Figure 2).

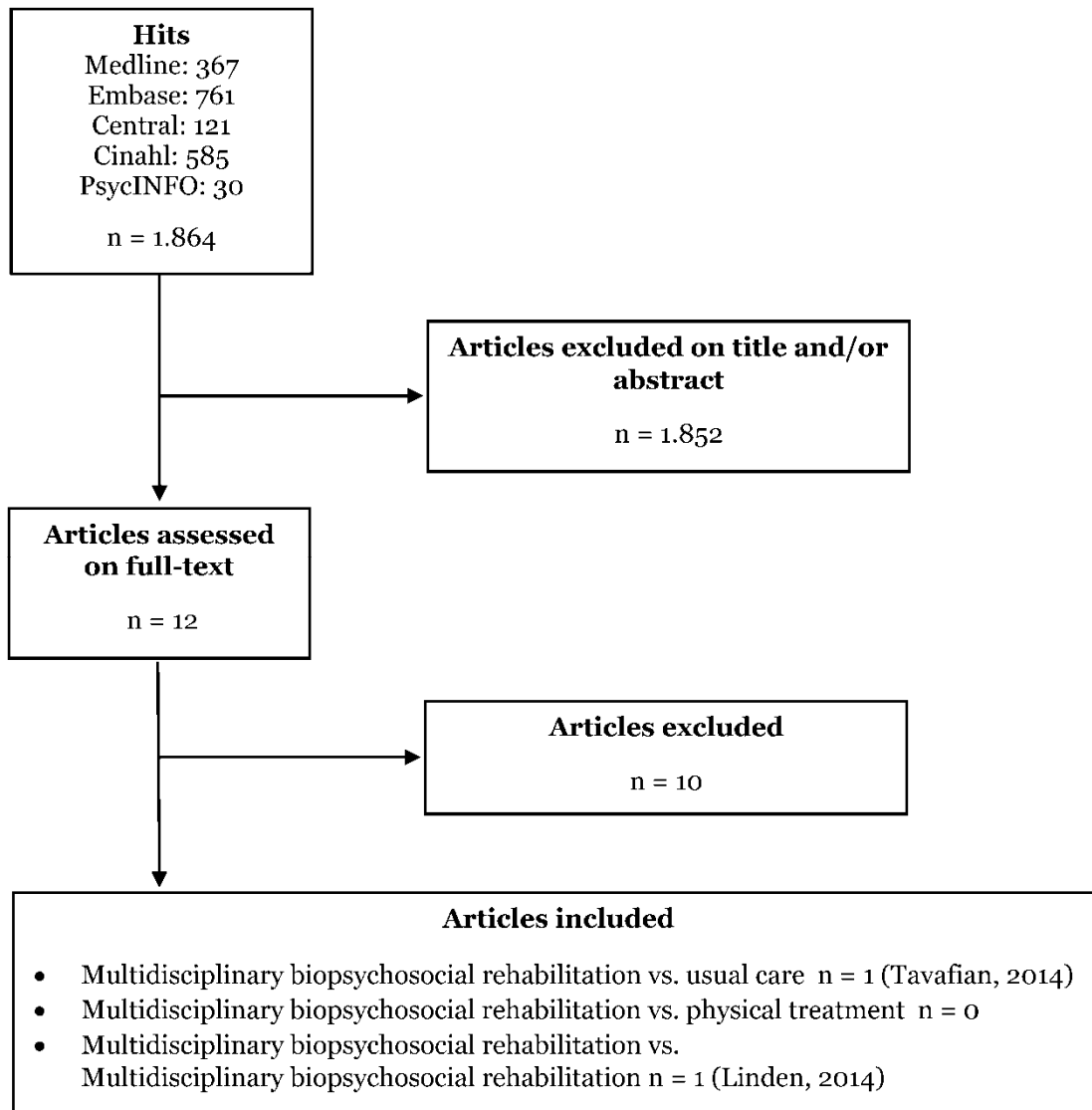


Figure 2. Flow-chart illustrating the updated literature search and assessment in February 2016.

In total, 13 specific RCTs comparing two or more rehabilitation programmes were of notable interest in the development process and formed the evidence base for the integrated programme; 12 RCTs were identified in the Cochrane review (7) and one additional RCT (44) from the updated literature search. In brief, the 13 RCTs were very heterogenous as they compared a variety of inpatient- and outpatient programmes, using diverse clinical activities, different provider compositions, various follow-up lengths, and different outcome domains and outcome measures (7, 44). Four of the RCTs had disability as the primary outcome (46-49) (Table 2).

Of the 13 RCTs, only one RCT assessed the effect of adding booster sessions (seven phone calls) to a 4-week multidisciplinary inpatient rehabilitation programme, aiming to demonstrate if adding the booster sessions had additional benefit in stabilising treatment successes one year after discharge (48). The RCT found a slight, but not statistically significant, advantage of adding the booster sessions compared with a 4-week multidisciplinary inpatient rehabilitation programme without booster sessions (48) (Table 2).

None of the 13 RCTs explicitly claimed to assess integration of knowledge, skills and behaviours from an inpatient stay into the daily life of the patients.

Table 2. Overview of the four RCTs comparing two or more rehabilitation programmes having disability as the primary outcome.

Author [Year] (Ref.)	Number of participants (in groups)	Intervention 1	Intervention 2	Intervention 3	Outcome domains (outcome measurements)	Follow up	Conclusion in terms of disability
Abbasi [2012] (46)	36 (12/12/12)	Standard medical care -> continuation of routine treatment based on ordinary medical care.	P-MPMP -> Conventional patient-oriented multidisciplinary pain management programme. 7x weekly sessions of 2 hours (6/group) + session with doctor + session with physiotherapist Light mobilisation, coping skills training, education regarding anatomy, physiology, medication, exercise session. Outpatient	SA-MPMP -> As P-MPMP + involvement of spouse	Primary: - Disability (Roland and Morris Disability Questionnaire) - Pain severity (Visual Analogue Scale) Secondary: - Fear avoidance (Tampa Scale of Kinesiophobia) - Pain catastrophizing (Pain Catastrophizing Scale) - Psychological distress (Depression, Anxiety and Stress Scale) - Marital adjustment (Marital Adjustment Test)	12 months	No significant between-group differences were found for Roland and Morris Disability Questionnaire scores.

Author [Year] (Ref.)	Number of participants (in groups)	Intervention 1	Intervention 2	Intervention 3	Outcome domains (outcome measurements)	Follow up	Conclusion in terms of disability
Leeuw [2008] (47)	85 (42/43)	<p>Exposure in vivo treatment (EXP).</p> <p>16x 1hr sessions, 2/week.</p> <p>Information about diagnosis, imaging, continued active approach, treatment rationale. Establishment of hierarchy of feared activities, explanation of fear avoidance model, gradual, systematic exposure to feared activities. Behavioral experiments to test consequences of engagement in feared activities.</p>	<p>Graded Activity (GA).</p> <p>26x 1hr sessions, 2/week. Information about diagnosis, imaging, continued active approach, treatment rationale. Identification of functional treatment goals, quota-based gradual increase in performance of functional activities.</p> <p>Two sessions included spouses.</p>		<p>Primary:</p> <ul style="list-style-type: none"> - Disability (Quebec Back Pain Disability Scale) - Main complaints (Patient Specific Complaints) <p>Secondary:</p> <ul style="list-style-type: none"> - Harmfulness of activities (Photograph Series of Daily Activities) - Pain catastrophizing (Pain Catastrophizing Scale) - Daily activity (accelerometer) - Pain intensity (McGill Pain Questionnaire) 	6 months	No significant difference was found between treatment conditions in functional disability.

Author [Year] (Ref.)	Number of participants (in groups)	Intervention 1	Intervention 2	Intervention 3	Outcome domains (outcome measurements)	Follow up	Conclusion in terms of disability
Mangles [2009] (48)	363 (131/113/119)	<p>Traditional Orthopedic Rehabilitation treatment.</p> <p>3 weeks of individual sessions or open groups (inpatient)</p> <p>Mostly physical/orthopedic treatment including active physiotherapy, passive modalities and occupational therapy.</p>	<p>Behavioral-medical Rehabilitation Treatment.</p> <p>4-week inpatient stay in groups of 10-12 patients.</p> <p>Treatment as in the Traditional Orthopedic Rehabilitation treatment + implementation of explicit psychologic treatment elements.</p> <p>- Group session of psychologic pain management (9 sessions of 90 minutes) conducted by behavioral psychotherapist.</p> <p>- Home assignments</p> <p>- Progressive muscle relaxation group</p>	<p>Behavioral-medical Rehabilitation Treatment plus subsequent booster sessions.</p> <p>4-week inpatient stay in groups of 10-12 patients.</p> <p>Treatment as in Behavioral-medical Rehabilitation Treatment + booster sessions consisting of 7 telephone calls of 20 minutes over 12 weeks to reinforce inpatient topics, problem-solving, goal setting, relaxation, coping etc.; ending with home-work. Handouts were given on all topics before beginning this aftercare treatment.</p>	<p>- Disability (Pain Disability Index)</p> <p>- Depression (Beck Depression Inventory)</p> <p>- Pain perception (The Pain Perception Scale)</p> <p>- Health-status (SF-12 Health Survey)</p> <p>- Coping (German Pain Management Questionnaire)</p> <p>- Pain Self-efficacy (Pain Self-Efficacy Questionnaire)</p> <p>- Life satisfaction concerning health (German Life Satisfaction Questionnaire)</p> <p>Not stated which was primary.</p>	12 months	<p>No treatment effect in terms of disability.</p> <p>Only slightly advantages for the subsequent booster sessions were found.</p>

Author [Year] (Ref.)	Number of participants (in groups)	Intervention 1	Intervention 2	Intervention 3	Outcome domains (outcome measurements)	Follow up	Conclusion in terms of disability
			- Opportunity to engage in weekly individual sessions with psychotherapist	Phone calls were conducted by clinical psychologist.			
Smeets [2008] (49)	223 (58/53/61 and waiting list controls=51)	Graded activity with problem solving training (GAP) 3 group sessions + max 17 individual sessions of 30 minutes), problem-solving (10 group sessions x 90 min), modification of dysfunctional beliefs, HEP increasing activity.	Active physical treatment (APT) 10 weeks. 3x/week 1 hour and 45 minutes of aerobic training, strength and endurance.	Combination of GAP + APT APT three/week and GAP once/week Frequency and duration as in APT and GAP.	Primary: - Disability (Roland and Morris Disability Questionnaire) Secondary: - Pain (Visual Analogue Scale and McGill) - Patients main complaints - Self-perceived improvement - Depression (Beck Depression Inventory) - Six physical performance tests	6 months 12 months	The combination treatment is no better treatment option.

2.1.1.2 Identifying/developing theory

No single theory of rehabilitation exists (50), but based on the knowledge presented in Chapter 1, the biopsychosocial approach (15, 16, 18-20) was the dominant theory. In addition, the Chronic Care Model (51-53) was considered appropriate. Together, the two approaches formed the theoretical foundation of the integrated programme.

The biopsychosocial approach has been explained, discussed, and interpreted during the past 40 years. Two interpretations of the biopsychosocial approach were used to underpin the integrated programme, namely the ICF model (16) and four editorials arguing for a new rehabilitation approach (15, 18-20).

The ICF model embraces the biopsychosocial dimensions of disability by viewing functioning and disability, and contextual factors relevant to a given health condition (16). The ICF model offers a framework for applying the biopsychosocial approach into clinical practice (6) especially into rehabilitation where the consequences of disability are managed (6, 15, 17, 27, 30, 54). In the LBP field, it has been suggested that the ICF model can serve as the foundational underpinning of a comprehensive framework to better guide rehabilitation of patients with LBP (55). The ICF model is an important element in any proposed theory of rehabilitation, and a valuable framework from which to begin any programme of rehabilitation, although it cannot stand alone (56).

The four editorials enhanced the following elements as sound, logical and/or evidence-based in rehabilitation (15, 18-20):

- Rehabilitation works in the context of a complex system (18).
- The patient must be actively engaged and practise activities as much as possible in the fundamental process of learning (15, 18). In terms of learning, there are some important principles: 1) the patient must want to learn, 2) the patient must practise, 3) the patient needs feedback on his/her performance, 4) the patient needs to take responsibility for his/her learning, and 5) practice should be in the context where the activity is to be undertaken normally (15). Face-to-face therapy is an important, but not dominant, component in rehabilitation (18). Of great importance is the amount of time the patient spends learning (20).
- A skilled multidisciplinary rehabilitation team is required to effectively manage the complex problems seen in rehabilitation (18-20). The goal is to teach the patient knowledge, skills and self-management, and arrange the context in an optimal way (19, 20).

The abovementioned elements correspond with elements recommended in the rehabilitation of patients with LBP (6).

Moreover, some of the elements have also been described in the Chronic Care Model (51-53). In the Chronic Care Model, evidence of effective changes needed to improve chronic care has been synthesised (51-53). At the center of the Chronic Care Model is a productive interaction between an informed, actively engaged patient and a prepared, proactive practice team united in a collaborative relationship (52, 53). The role of the practice team is to ensure that patients are equipped with the requisite confidence and skills to self-manage their health condition (51-53). This can be supported by using regular follow-up interactions (hereafter termed “booster sessions”) (51-53). Furthermore, due to the time horizon and fluctuating course of chronic health conditions, regular interactions between the patient and the practice team are essential. Booster sessions can take the form of practice-initiated follow-up visits, but methods other than face-to-face contacts can be used (e.g. telephone calls). Regardless of the form of contact, the purpose of the booster sessions is to evaluate response to therapy and self-management competences, and to adjust treatment (51, 53).

Collectively, the chosen theories underpinned the justification of the integrated programme. The integrated programme targeted the biopsychosocial factors driving disability, and further, it acknowledged the influence of the context. With the integrated programme, the delivery of an existing rehabilitation programme was changed by alternating inpatient stays (including booster sessions) and home-based activities. During the inpatient stays, patients were motivated, assisted and challenged by a multidisciplinary team in the process of learning. During the home-based periods, the patients were encouraged to integrate their newly obtained knowledge, skills and behaviours into their daily life. The booster sessions further allowed for adjustment and modification of the rehabilitation plan.

2.1.1.3 Modelling process and outcomes

Step 1

A project team comprising the PhD candidate, supervisors and management staff was established. It was decided to name the new rehabilitation programme “the integrated programme”, since the main feature was to enhance integration of knowledge, skills, and behaviours from inpatient stays into the patients’ daily life.

Step 2

The overall aim of the RCT was to compare two different ways of delivering a rehabilitation programme, thus, it was decided to include exactly the same clinical activities in the integrated programme and the comparator being the existing programme. The clinical activities in the existing programme were originally based on the literature and knowledge about CLBP using the biopsychosocial approach.

During the consensus process on how to deliver the integrated programme, the providers argued for the booster sessions and the prolonged rehabilitation time as it would give the patients the opportunity to integrate knowledge, skills and behaviours into their daily life while still being in contact with the providers. Most patients were positive about the alternation between inpatient stays and home-based activities.

The organisational impact of the integrated programme was assessed, and it was deemed possible to conduct the proposed RCT in the setting under study.

Step 3

The clinical activities were grouped into 10 key components delivered by the providers (Table 3). Together, the 38 clinical activities targeted biophysical, psychological, social, and lifestyle factors.

Table 3. Key components and the related 38 clinical activities in the rehabilitation programmes (57).

Key component	Clinical activity
Clinical assessment	Physical assessment Psychosocial assessment
Motivation and change	Instruction in exercise app Introduction to rehabilitation Exercise theory Introduction to mindfulness Involvement of relatives Motivation and anchoring The next step
Pain knowledge and management	Chronic pain and chronic back pain Experiences with pain Knowledge about pain Knowledge about analgaesic medicine Living with pain
Multidisciplinary intervention	Welcome meeting Multidisciplinary conference Open counselling Midterm evaluation
Exercise and physical activity	Aqua gymnastics Circuit training Intro electric bicycle Exercise capacity training Healthy feet Closing activity
Individual counselling	Individual nurse counselling (focusing on psychologic aspects) Individual physiotherapy counselling Individual occupational therapy counselling Individual dietary counselling Individual rheumatologic counselling
Essential activity	Activity and health Balanced activity Activity and social relations Lifelong activity
Activities of daily living	Sleep House and garden
Nutrition and weight loss	Permanent weight loss strategies Healthy lifestyle
Individual exercise	Individual exercise

Step 4

Tasks and roles of providers and administrative staff were described in internal documents and in a trial flow illustrating the patients' trajectory through the inclusion process.

Step 5

A detailed description of the integrated programme following the Template for Intervention Description and Replication (TIDieR) checklist (58) was undertaken and published (57).

Step 6

The primary and secondary outcomes and outcome measures were chosen (Section 4.2.5).

2.1.2 Feasibility/piloting

2.1.2.1 Testing procedures

The main changes and initiatives were:

- 1) Fine-tuning of administrative procedures (e.g. revision of the welcome letter, nominating persons responsible for phone calls and data collection, and documentation of informed consent form in the electronic health records).
- 2) Overbooking of two to three patients in each group due to patient-initiated postponements.
- 3) Arranging a place to rest on the pre-admission day and at the 6-month follow-up visit, and assigning all groups a specific dinner table.
- 4) Adjusting layout and minor features in the database.

2.1.2.2 Estimating recruitment/retention

The number of eligible patients and the intended willingness to participate were considered large enough to recruit a sufficient number of patients.

2.1.2.3 Determining sample size

Sample size was calculated (Section 4.2.7.1).

2.1.3 Evaluation

2.1.3.1 Assessing effectiveness

A single-centre, pragmatic, two-arm parallel RCT was chosen as an appropriate design. The RCT is further described in Chapter 4.

2.1.3.2 Understanding change processes

As advocated when working with complex interventions, a process evaluation was nested into the RCT (32). A process evaluation is useful when interpreting the results to explain discrepancies between expected and observed outcomes, or when clarifying why and how a successful intervention works (36, 43, 59). In brief, an RCT assesses if an intervention is working or not, while a process evaluation seeks to understand why the intervention is working or not (41, 43, 59).

The process evaluation relied on a framework that consisted of the following three key components:

- 1) understand how a complex intervention is implemented,
- 2) clarify mechanisms of impact, and
- 3) identify interaction with the context

(41, 42, 60).

The three key components, the related subdivision and a corresponding explanation is provided in Table 4.

Table 4. Key components, subdivision and corresponding explanation of a process evaluation. Inspired by and combined from relevant articles (41, 42, 60).

Key components	Subdivision	Explanation
<p><i>Implementation - what is implemented and how?</i> The extent to which the intervention has been implemented and received by the intended participants.</p>	Implementation process.	How delivery is achieved; training, resources etc.
	<p>What is delivered:</p> <ol style="list-style-type: none"> 1) Fidelity 2) Dose 3) Adaptations 4) Reach 	<ol style="list-style-type: none"> 1) The extent to which the intervention was delivered as intended. 2) The quantity of intervention implemented. 3) Adaptations often arise when the intervention is implemented across different contexts. 4) Whether the intended participants come into contact with the intervention, and how.
<p><i>Mechanisms of impact – how does the delivered intervention produce change?</i> When capturing mechanisms of impact, participants' response to, and interaction with, the intervention, as well as unexpected pathways and consequences can be described.</p>	Participant response to and interactions with, the intervention.	The mechanisms through which interventions bring about change.
	Mediators.	
	Unexpected pathways and consequences.	
<p><i>Context - how does context affect implementation and outcomes?</i> When focusing on context, focus can be on how context shapes implementation and influences whether intervention mechanisms work.</p>	Contextual factors that shape theories of how the intervention works.	Anything external to the intervention that may act as a barrier or facilitator to its implementation, or its effect.
	Contextual factors that affect (and may be affected by) implementation, intervention mechanisms and outcomes.	
	Causal mechanisms present with the context which act to sustain the status quo, or potential effects.	

In the following the methods used to 1) plan, 2) design and conduct, 3) analyse, and 4) report (41, 42) the process evaluation will be elaborated.

Firstly, the plan was pre-specified, and all issues known to potentially affect the results were discussed. The purpose of the process evaluation was to shed light on selected aspects of implementation, mechanisms of impact and context related to the integrated programme in order to seek to understand and explain the results of the RCT.

Secondly, to underpin the design and conduct, a logic model was developed and refined until it depicted the integrated programme including justification, planned work and intended results (61). A pictorial representation can be seen in Figure 3. Logic models seeks to illuminate the chosen theories, is a map of the road ahead (61), and may reveal possible interactions (62) making it useful when identifying relevant process research questions (41). Thus, the logic model is a simplification of a complex reality, and is useful across all the stages of the research process.

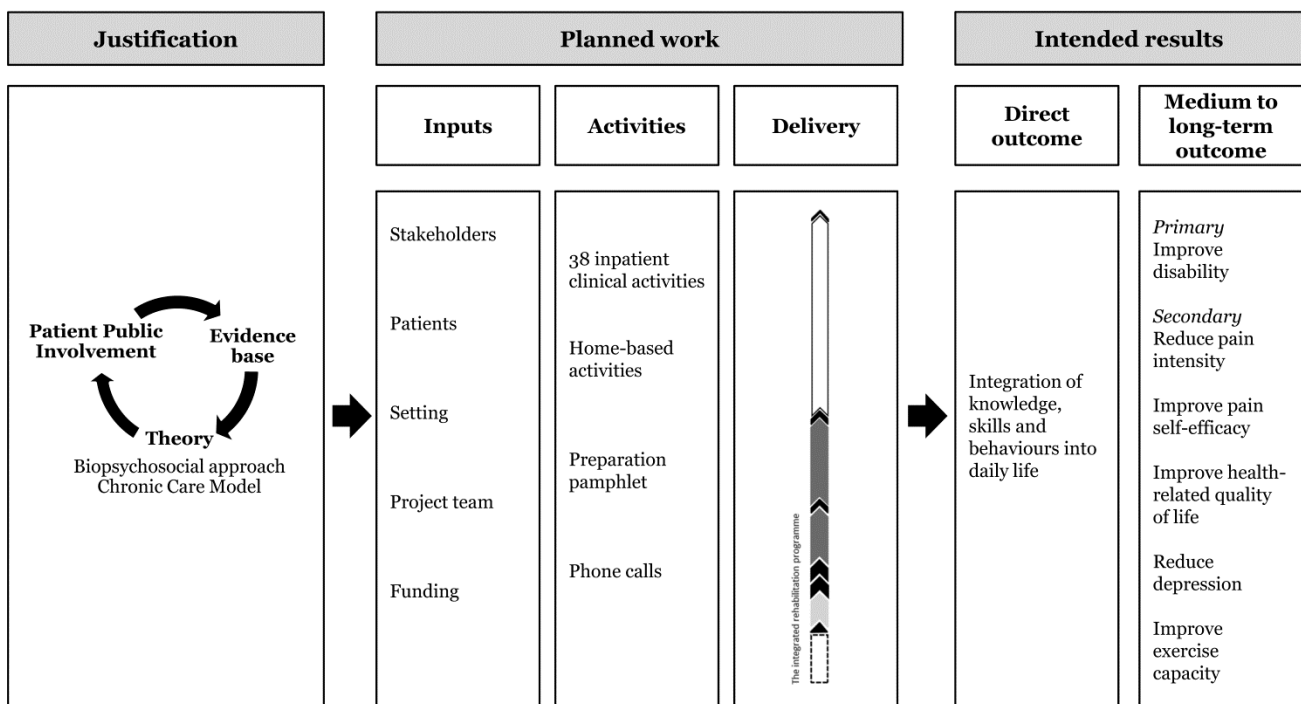


Figure 3. Logic model of the integrated programme.

Process research questions 1-5 were identified prior to the RCT (Figure 4). The process evaluation was designed with flexibility, as complex interventions are inherently unpredictable (41, 42), thus, research question 6 occurred during the evaluation of the RCT (Figure 4). A combination of

qualitative (process research questions 1-4) and quantitative methods (process research questions 5-6) were used.

The PhD candidate analysed the process research questions, except for process research questions 3 and 4, which were analysed in-depth by two Masters students (data not published).

The PhD candidate provided ongoing verbal and written feedback at meetings or via email on process findings to stakeholders (includes providers, administrative and management staff), thereby striving for quality improvement (61).

Thirdly, the analyses of the qualitative process findings (process research questions 1-4) were performed before the statistical analyses. Quantitative data (process research questions 5-6) on adherence and waiting time were used in secondary analyses.

Fourthly and lastly, findings were reported systematically and explicitly in line with recommendations for process evaluations (41-43). The findings from process research questions 1-4 were briefly summarised (Section 5.5), and the findings from process research questions 5 and 6 were reported in Studies 2 and 3 (Sections 5.3 and 5.4).

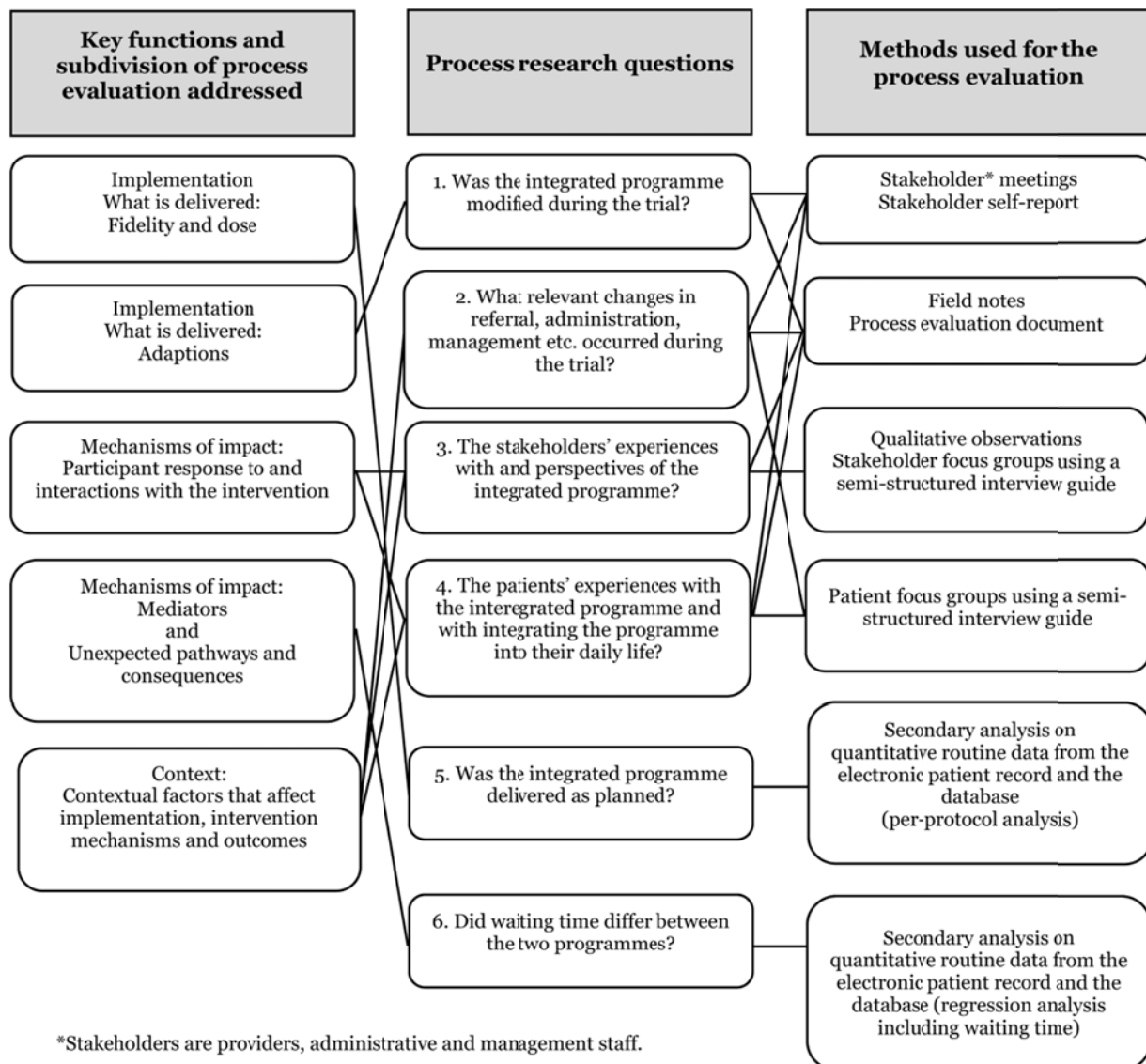


Figure 4. The three key functions and subdivisions of process evaluation addressed (column 1), process research questions (column 2) and methods used for the process evaluation (column 3) (63).

2.1.3.3 Assessing cost-effectiveness

This was outside the scope of this dissertation.

2.1.4 Implementation

Implementation in clinical practice was outside the scope of this dissertation. However, in line with recommendations (32), implementation was considered and discussed with stakeholders early in the research process as there was no need to develop and evaluate a rehabilitation programme that was unlikely to be implemented following the evaluation.

2.2 Patient and public involvement

Patient and Public Involvement (PPI) is defined as research being carried out with or by members of the public rather than to, about, or for them (64). The interest in PPI is growing (38, 65), and in rheumatology, the European League Against Rheumatism has published eight recommendations for the inclusion of patients in research (66). Although sparsely described in the MRC's guidance (32, 36), PPI was prioritised as a part of this research process since preferences and concerns might differ between researchers, patients, and stakeholders (64-66).

In the current research process the purpose of PPI was to consult and collaborate with patients and stakeholders as research partners in the development, feasibility-testing and evaluation of the RCT.

The patients recruited for the PPI process were randomly selected among those who were inpatients at different time points relevant to the research project. Thus, they represented a diversity of patients who could contribute with a variety of different reflections and personal experiences with CLBP and related disability. Furthermore, it minimised the logistical challenge of gathering the same group of patients repeatedly. All patients who participated in the PPI process volunteered. The stakeholders involved had experiences with delivering rehabilitation in the setting under study. Data obtained from meetings and focus groups were stored in meeting summaries, the PhD candidate's field notes and a process evaluation document. PPI in this research project did require time resources, but no direct costs.

A completed GRIPP2 short form reporting PPI in the research process can be found in Appendix 5.

3. Aims

The overall aim of this dissertation was to contribute new knowledge about the delivery of multidisciplinary rehabilitation programmes to patients with CLBP. For this purpose, the integrated programme was developed, feasibility-tested and evaluated.

The three specific aims listed below correspond to the three studies undertaken:

Aim 1 (Study 1)

To justify and describe the integrated programme in detail prior to an RCT.

Aim 2 (Study 2)

To compare the effectiveness of the integrated programme with the existing programme in terms of back-specific disability in patients with chronic low back pain at the 6-month follow up.

It was hypothesised, that the integrated programme would be superior to the existing programme.

Aim 3 (Study 3)

To compare the effectiveness of the integrated programme with the existing programme in terms of back-specific disability in patients with chronic low back pain at 12-month follow up.

It was hypothesised, that the integrated programme would be superior to the existing programme.

4. Methods

Study 1 is a justification and detailed description of the integrated programme. The evaluation of the RCT is reported in Study 2 (6-month follow up) and Study 3 (12-month follow up). Table 5 presents an overview of the three studies included in this dissertation. Following the overview, the methods used are presented in further detail; Study 1 on its own, and Studies 2 and 3 together as they were closely related.

Table 5. Overview of design, applied reporting guidelines, outcome domains and outcome measures, follow up and primary statistical analyses.

Study	Design	Reporting guidelines	Outcome domains (outcome measures)	Follow up	Primary statistical analyses
1	Descriptive study	TIDieR (58) CERT (67)	-	-	-
2	RCT	CONSORT 2010 (68) CONSORT extensions (69-71) GRIPP2 short form (38)	Primary: Disability (ODI) Secondary: Pain intensity (NRS) Pain self-efficacy (PSEQ) Health-related quality of life (EQ-5D 5L) Depression (MDI) Physical activity (Three questions) Exercise capacity (Astrand test)	6-month	Complete case analysis. Multiple linear regression using change score as the dependent variable, rehabilitation programme as the independent variable, and the corresponding baseline score as covariate.
3	RCT	CONSORT 2010 (68) CONSORT extensions (69-71) GRIPP2 short form (38)	Primary: Disability (ODI) Secondary: Pain intensity (NRS) Pain self-efficacy (PSEQ) Health-related quality of life (EQ-5D 5L) Depression (MDI) Physical activity (Three questions) Exercise capacity (Astrand test)	12-month	Intention-to-treat analysis including the four similar measurement time points. Difference in change between the two groups (integrated programme minus existing programme) from baseline to 12-month using a linear mixed model with a random intercept including time, group and the interaction between group and time as the explanatory variables.

TIDieR: Template for Intervention Description and Replication; CERT: Consensus on Exercise Reporting Template; CONSORT: Consolidated Standards of Reporting Trials; GRIPP2 short form: Guidance for Reporting Involvement of Patients and the Public 2 short form; ODI: Oswestry Disability Index; NRS: Numerical Rating Scale; PSEQ: Pain Self-Efficacy Questionnaire; EQ-5D 5L: EuroQol-5 Domain 5-level; MDI: Major Depression Inventory

4.1 Study 1

Structure and detail was provided primarily using the TIDieR checklist (58) of which the 12 items are listed and explained in Table 6.

Table 6. Items and item categories of the TIDieR checklist along with a brief item description (58).

Item	Item category	Brief item description
1	Brief name	Provide the name or a phrase that describes the intervention
2	Why	Describe any rationale, theory, or goal of the elements essential to the intervention
3	What (materials)	Describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or in training of intervention providers. Provide information on where the materials can be accessed (for example, online appendix, URL)
4	What (procedures)	Describe each of the procedures, activities, and/or processes used in the intervention, including any enabling or support activities.
5	Who provided	For each category of intervention provider (for example, psychologist, nursing assistant), describe their expertise, background and any specific training given.
6	How	Describe the modes of delivery (such as face to face or by some other mechanism, such as internet or telephone) of the intervention and whether it was provided individually or in a group.
7	Where	Describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features.
8	When and how much	Describe the number of times the intervention was delivered and over what period of time including the number of sessions, their schedule, and their duration, intensity or dose.
9	Tailoring	If the intervention was planned to be personalised, titrated or adapted, then describe what, why, when, and how.
10	Modifications	If the intervention was modified during the course of the study, describe the changes (what, why, when, and how).
11	How well (planned)	If intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them.
12	How well (actual)	If intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned.

For the purpose of clarity, the 38 different clinical activities included in the integrated programme were each presented in an activity sheet combining “What (procedures)”, “Tailoring”, “Who provided”, “How” and “When (and how much)” from the TIDieR checklist (58). In addition, the Consensus on Exercise Reporting Template (CERT) checklist (67) was used to ensure proper reporting of the exercise interventions, highlighting the importance of including specific information about e.g. frequency, duration, intensity and supervision.

Thus, the two checklists (58, 67) were used to achieve transparent and complete description of the 38 clinical activities. An example of a completed activity sheet is presented in Table 9.

4.2 Studies 2 and 3

4.2.1 Choice of design and its explanation

These two studies occurred in the context of a single-centre, pragmatic, two-arm parallel, RCT (63, 72).

When developing the RCT, we aimed for a high degree of applicability, meaning that the results of the RCT were intended to have the ability to directly inform, and be relatively easy to apply to, clinical practice (73, 74). Aiming to maximise applicability corresponds to conducting a pragmatic trial (73). While a pragmatic trial is designed to determine intervention effects under usual conditions, an explanatory trial is designed to determine intervention effects under ideal circumstances (73, 75). The difference between the two approaches is not dichotomous rather a continuum (73, 74). PRECIS-2 (second version of Pragmatic Explanatory Continuum Indicator Summary) is a tool developed for use when designing the trial; it aims to help the researcher to make the purpose explicit and ensure that design choices and purpose are concordant, and further, to make the researcher aware of the trial's location on the pragmatic-explanatory continuum (73).

In the development of the RCT, we were not aware of PRECIS-2 (73), and hence the tool was used retrospectively in order to reveal the (actual) position of the RCT on the pragmatic-explanatory continuum. The PhD candidate and the main supervisor assessed each domain individually. After that, a discussion followed and consensus was obtained.

The PRECIS-2 assessment was performed following four steps (73):

Step 1. What design approach are you taking? (Section 4.2.1)

Step 2. Consider your trial design choices for each of the nine PRECIS-2 domains (Table 7).

Step 3. Score 1 to 5 for the choices made in step 2 and/or mark on the PRECIS-2 wheel (score 1 is “very explanatory”, score 2 is “rather explanatory”, score 3 is “equally pragmatic and explanatory”, score 4 is “rather pragmatic”, and score 5 is “very pragmatic”) (Table 7 and Figure 5).

Step 4. Review the PRECIS-2 wheel (Figure 5).

Table 7. Overview of domains, rationale for design choices and the corresponding score given by applying PRECIS-2 to this RCT.

	Domain	Rational for design choices	Score
1	Eligibility	The eligibility criteria were very much identical to those in the usual setting – except for exclusion of patients with axial spondyloarthritis and Oswestry Disability Index score <21 (Section 4.2.2).	4
2	Recruitment	Patients referred to the usual setting were recruited without knowledge of the RCT. No extra effort was made to enhance referral to the usual setting.	5
3	Setting	Full accordance between the setting of the RCT and the setting where the results are likely to be applied.	5
4	Organisation	The integrated programme drew upon existing competences among providers and administrative staff.	3
5	Flexibility (delivery)	The integrated programme was partially tailored to ensure a degree of person-centeredness. There was no strict protocol in terms of the content and no restrictions of co-interventions.	5
6	Flexibility (adherence)	Adherence to inpatient days was recorded, but patients were not excluded based on this record and no effort was made to improve adherence to inpatient days. Adherence to home-based activities was not measured.	5
7	Follow-up	Identical follow up in both rehabilitation programmes (6-month and 12-month), but slightly more data collection in the integrated programme.	4
8	Primary outcome	Choice based on the literature and PPI confirming that the primary outcome was meaningful to patients and providers.	5
9	Primary analysis	A complete case analysis only including data from completers (Study 2).	1
		An intention-to-treat analysis using all available data regardless of adherence and/or complete follow-up data (Study 3).	5

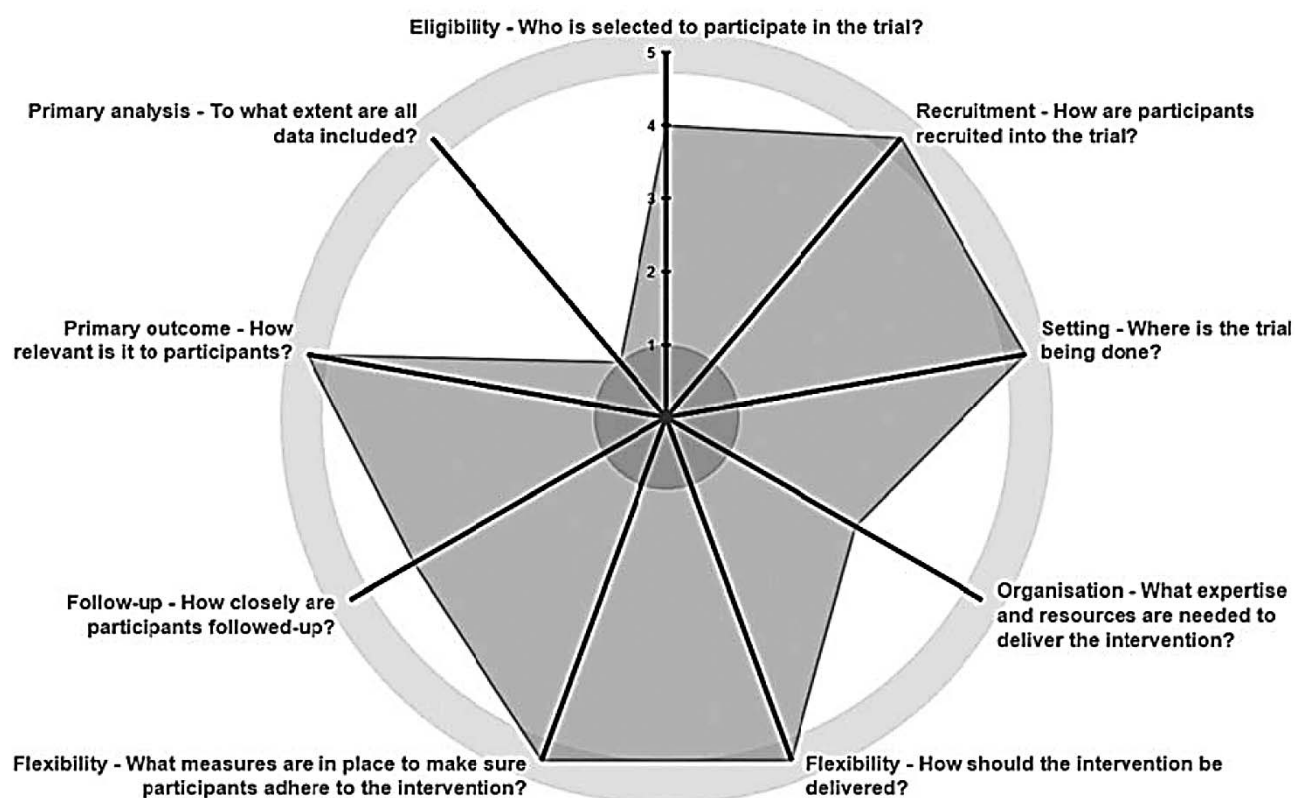


Figure 5. The PRECIS-2 wheel depicting Study 2 (76).

4.2.2 Setting and participants

The setting under study was a multidisciplinary rehabilitation centre, owned and operated by the Danish Rheumatism Association. Patients with rheumatic and musculoskeletal diseases were referred from general practitioners or hospital departments to a tax-funded rehabilitation programme. Allowing for approximately 28 patients to be inpatients from Monday to Friday, a running schedule starting a group of six to eight patients on a rehabilitation programme each week meant that there were constantly four groups of patients at the rehabilitation centre at the same time.

Patients were included in the RCT if they had had CLBP for more than 12 months (+/- sciatica and/or widespread pain) and if they were 18 years or older. The RCT exclusion criteria were: 1) severe systemic diseases (American Society of Anesthesiologists Physical Status classification system > 3 (77)), 2) a diagnosis of axial spondyloarthritis, 3) spinal fracture within the last three months, 4) severe osteoporosis, 5) active cancer, 6) severe psychiatric disease, 7) pregnancy, 8) lack of fluency in Danish, and 9) minimal back-specific disability (Oswestry Disability Index (ODI) score < 21 (78)).

The rheumatologist at the rehabilitation centre identified potentially eligible patients based on the referral request and a list of ICD-10 diagnosis codes. Before inclusion, a research assistant performed eligibility checks by telephoning potentially eligible patients. Participant information and an informed consent form were emailed by the research assistant, and if a signed version was returned, the patient was included. The final eligibility check was performed by the rheumatologist on the first inpatient day.

4.2.3 Randomisation and blinding

A secure electronic database was used to email questionnaires, store data, and generate the random allocation sequence. A computer-generated randomisation with 1:1 allocation in random blocks of six ensuring allocation concealment (79) was performed by the research assistant. Randomisation was stratified on the basis of baseline disability (ODI score over/under 41) (78) to achieve approximate balance in mean disability levels (79). The research assistant informed patients about intervention allocation and appointed dates for the allocated rehabilitation programme. The patients had wait time until the next available rehabilitation programme, as this was usual practice at the rehabilitation centre.

Blinding of patients and providers was not possible due to the nature of the rehabilitation programmes (31, 79), although attempts were made to blind patients to the hypothesis by informing them that the RCT aimed to compare two rehabilitation programmes both which met current recommendations. This was done in order to equalise expectations for the allocated rehabilitation programme (31, 79). The physiotherapists were blinded to the result of the baseline exercise capacity test when performing the 6-month follow-up exercise capacity test. The PhD candidate performed the statistical analyses and was blinded to allocation.

4.2.4 Interventions

The key difference between the two rehabilitation programmes was how they were delivered (Figure 6).

Patients in *the integrated programme* participated in: 1) a pre-admission day, 2) two weeks of home-based activities, 3) a two-week inpatient stay (nine days excluding the weekend), 4) four weeks of home-based activities, 5) an initial two-day inpatient booster session, 6) six weeks of home-based activities, 7) a second two-day inpatient booster session, and 8) a 6-month follow-up visit. In total, the integrated programme comprised 15 inpatient days.

Patients in *the existing programme* participated in: 1) a four-week inpatient stay (20 days excluding the weekends), and 2) a 6-month follow-up visit, resulting in 21 inpatient days. The existing programme has been usual practice for more than 15 years.

Differences and similarities between the two rehabilitation programmes are reported in Table 8.

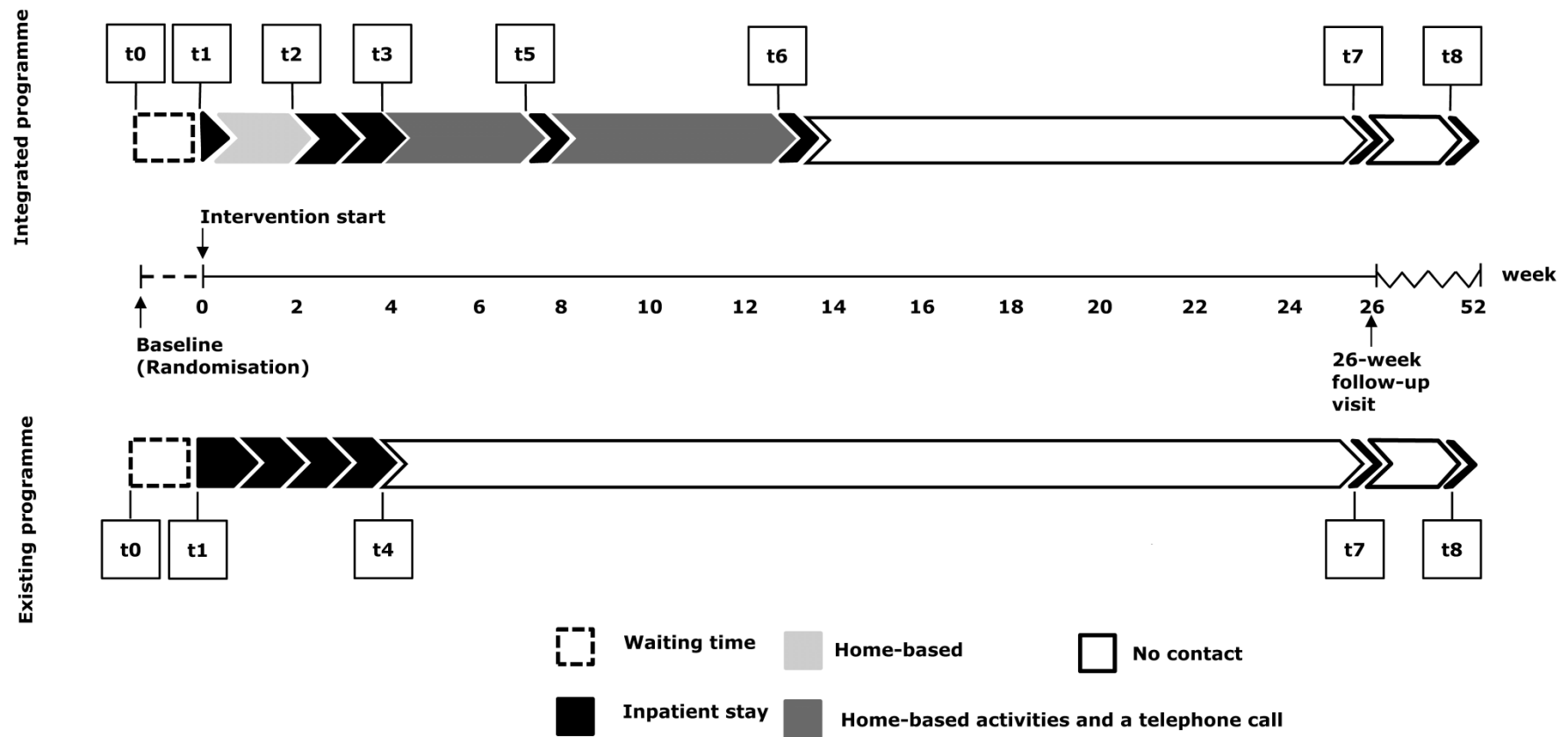


Figure 6. Comparison of the integrated programme and the existing programme (72).

Table 8. Differences and similarities between the integrated programme and the existing programme (72).

Differences		
Characteristics	Description	
	Integrated programme	Existing programme
Delivery of the rehabilitation programme	1) pre-admission day 2) two weeks of home-based activities 3) two-week inpatient stay 4) four weeks of home-based activities 5) initial two-day inpatient booster session 6) six weeks of home-based activities 7) second two-day inpatient booster session 8) 6-month follow-up visit In total 15 inpatient days. In between the inpatient stays, patients were at home (in total, 11 weeks). The integrated programme lasted for 14 weeks. During the integrated programme, continuous focus was on integration of knowledge, skills and behaviours into daily life supported by the home-based activities.	1) 4-week inpatient stay 2) 6-month follow-up visit In total, 21 inpatient days. The existing programme lasted for 4 weeks. At the end of the existing programme focus was on the integration of knowledge, skills and behaviours into daily life.
Elements to support integration of knowledge, skills and behaviours	<ul style="list-style-type: none"> - Pamphlet - A phone call half-way through each home-based period (in total, two phone calls) - Focus on integration during the inpatient stays, but especially in the clinical activities comprising individual counselling. 	-
Development of the rehabilitation programme	A systematic research process following the MRC's guidance on developing and evaluating complex interventions.	Not developed using a systematic research process.
Patient and public involvement	A major part of the development, feasibility-test and evaluation.	-
Theoretical underpinning	The biopsychosocial approach and the Chronic Care Model.	The biopsychosocial approach.

Similarities	
Characteristics	Description
Rehabilitation programmes for patients with a variety of rheumatic health conditions	The rehabilitation programmes encompassed the four actions characterising the process of rehabilitation: 1) assessment, 2) goal setting, 3) intervention, and 4) evaluation (Section 1.4).
Clinical activities	38 clinical activities targeted the biopsychosocial factors driving disability. Some of the clinical activities were delivered more than once.
Providers	The multidisciplinary team consisted of six physiotherapists, three occupational therapists, three nurses (educated as coaches and primarily focusing on the psychologic aspect of the programme), a rheumatologist, and a nutritional counsellor. The majority of the providers were trained in the Motivational Interviewing approach.
Contact hours	Identical in the two rehabilitation programmes (approximately 50 contact hours).
Mode	A combination of group lecture and dialogue, group sessions (supervised and non-supervised), individual counselling and non-supervised individual exercise.
Setting	A rehabilitation centre in Denmark.
Tailoring	The rehabilitation programmes were partially standardised and partially tailored, the latter to ensure a degree of person-centeredness. The tailoring occurred primarily during: 1) the clinical assessment (including goal setting), 2) the multidisciplinary conference, 3) individual counselling, 4) exercise and physical activity, and 5) individual exercise (Table 3).
Additional contact	Permission to contact the providers twice via the exercise app.

4.2.5 Baseline characteristics, outcome domains, outcome measures, and other variables

Baseline characteristics comprised sex, age, marital status, smoking, leg pain, employment status and educational level.

The primary measurement time point was 6-month follow up (63). In addition, 1-year follow-up results were analysed (72) as the long-term effect of such a rehabilitation programme is interesting given that time is thought to be related to successful integration of knowledge, skills, and behaviours (31).

The choice of outcome domains and outcome measures was based on PPI in combination with international recommendations (80-83).

They were:

Primary outcome:

- Disability assessed by the ODI version 2.1a. ODI is a 10-item scale where each item contains six statements describing increasing degrees of disability. Each item is scored on a 0-5 point scale. The total score is doubled and expressed as a percentage; it is suggested rounding the percentage to a whole number. Minimum score is 0 and maximum score is 100; the higher the score the greater the disability (78). The index can be aggregated into five levels where 0-20% indicates minimal disability, 21-40% indicates moderate disability, 41-60% indicates severe disability, 61-80% indicates crippled and 81-100% indicates patients being bed-bound or exaggerating their symptoms (78). ODI has shown good psychometric properties (84).

Secondary outcomes:

- Pain intensity assessed by a Numerical Rating Scale (NRS) (80). It is scored on a 0-10 point scale where 0 defines absence of pain and 10 describes unbearable pain; the higher the score, the greater the pain disability (82). The NRS has shown good psychometric properties (83).
- Pain self-efficacy using the Pain Self-Efficacy Questionnaire (PSEQ). It consists of 10 questions about the patient's confidence in carrying out various normal activities despite pain. There are seven response options ranging from 0 (not at all confident) to 6 (very confident). The total score ranges from 0 to 60 points with higher scores indicating higher perceived pain self-efficacy (85). The psychometric properties of the PSEQ have been found to be sound (86).
- Health-Related Quality of Life (HRQoL) was measured by EQ-5D 5L ©, a generic questionnaire containing five dimensions each with five response options. The final score has a unique 5-digit descriptor corresponding to each dimension. Using a crosswalk value set for Denmark, health status was calculated and ranged between 1 (optimal health) and 0.624 (worst health); a high score indicates better HRQoL (87). EQ-5D 5L has not been tested in a LBP population (82, 83, 88).
- Depression using the Major Depression Inventory (MDI) as a depression rating scale covering the symptoms of depression (89) which contains 10 items scored on a 6 point Likert scale from 0 (at no time) to 5 (all the time); higher scores indicate a higher degree of depression. The individual items measure the proportion of time with present symptoms during the past two weeks (89). The MDI has shown adequate psychometric properties (89).

- Physical activity was identified by asking three questions (90):
 - 1) minutes spent in physical exercise during a typical week (0 minutes, less than 30 minutes, 30-59 minutes, 60-89 minutes, 90-120 minutes and more than 120 minutes).
 - 2) minutes spent in physical activity during a typical week (0 minutes, less than 30 minutes, 30-59 minutes, 60-89 minutes, 90-149 minutes, 150-300 minutes and more than 300 minutes).
 - 3) hours spent sitting in a typical 24-hour period (nearly the whole day, 13-15 hours, 10-12 hours, 7-9 hours, 4-6 hours, 1-3 hours, and never).

The categories were slightly changed for the purpose of the RCT as the original categories were overlapping.
- Exercise capacity was measured by the Aastrand cycle test (91) using a calibrated bicycle (Monark 928 G3, Sweden) and performed by physiotherapists using a standardised performance-based protocol (92). There is moderate evidence of the reliability, validity and acceptability of the Aastrand cycle test in people with chronic pain (93).

Other variables:

- Cases of adverse events or death were extracted from electronic health records.
- Adherence was extracted from the electronic health records and defined as attending $\geq 80\%$ of the scheduled inpatient days, based on reasoning. Thus, adherence was defined as attending ≥ 12 inpatient days in the integrated programme and attending ≥ 17 inpatient days in the existing programme. Adherence to the home-based activities was not measured.
- Waiting time was measured by counting days from baseline to the start of the rehabilitation programme.

4.2.6 Data collection

Data on baseline characteristics and outcome variables (except for exercise capacity) were collected using a battery of questionnaires.

In the integrated programme, data were collected: 1) at baseline (before randomisation) (t₀), 2) before the pre-admission day (t₁), 3) before the 2-week inpatient stay (t₂), 4) at the end of the 2-week inpatient stay (t₃), 5) before the initial booster session (t₅), 6) before the second booster session (t₆), 7) before the 6-month follow-up visit (t₇), and 8) at 12 month follow up (t₈) (Figure 6).

For the existing programme, data were collected: 1) at baseline (before randomisation) (t₀), 2) before the 4-week inpatient stay (t₁), 3) at the end of the 4-week inpatient stay (t₄), 4) before the 6-month follow-up visit (t₇), and 5) at 12 month follow up (t₈) (Figure 6).

Patients were emailed the questionnaire link 10 days prior to the inpatient stay with a reminder after 5 and 8 days, if required. The research assistant ensured completion of questionnaires by handing out a tablet on the first inpatient day, if not completed beforehand. This procedure was repeated on the last inpatient day. Patients not attending a booster session, the 6-month follow-up visit and/or 12-month follow up were encouraged to complete the questionnaires via email. Patients unable to complete the electronic questionnaires were sent a paper version, which was double-entered into the database by the research assistant. The exercise capacity test was performed at the beginning of the 2-week (integrated programme) and the 4-week (existing programme) inpatient stay and, and again at the 6-month follow-up visit.

4.2.7 Statistical analyses

4.2.7.1 Sample size calculation

A difference in the change of 4 points in the ODI has been suggested as a minimum clinically important difference (78, 94). The sample size calculation for this RCT was based on an expected difference in change over time of 5 points on the ODI at the 6-month follow up (corresponding to a decrease of 10 points in the integrated programme compared to a decrease of 5 points in the existing programme (95)). The standard deviation (SD) was informed by our previous feasibility test including 12 patients attending the existing programme (SD on the difference of 10 points). With 80% power and a significance level of 0.05, 64 patients were required in each arm of the RCT, and allowing for lost to follow up of 20%, a total of 160 patients was needed.

4.2.7.2 Identical statistical analysis Studies 2 and 3

Descriptive statistics were presented with mean and SD or number and percentage depending on type of the variable.

Results were presented with 95% confidence intervals (CI) and p-values. A p-value ≤ 0.05 was considered statistically significant. For statistical analyses, STATA 15 (Study 2) and STATA 16 (Study 3) were used.

Statistical analysis plans was completed prior to the unblinding of data (Appendices 6 and 7).

4.2.7.3 Statistical analyses Study 2

Differences in sex, age and ICD-10 diagnosis codes were assessed for those patients randomised and those declining to participate. The primary analysis was a modified intention-to-treat (ITT) analysis according to originally allocated rehabilitation programme, excluding patients with missing outcome data at the 6-month follow up (complete case analysis). The between-group difference in mean change scores (integrated programme minus existing programme) from baseline to the 6-month follow up was analysed by multiple linear regression using change score as the dependent variable, rehabilitation programme as the independent variable, and the

corresponding baseline score as covariate. The within-group changes from baseline to the 6-month follow up were presented descriptively. Furthermore, the robustness of the complete case analysis in terms of the primary outcome was checked using ITT analysis with the last value carried forward. A per-protocol analysis was also conducted excluding patients who did not adhere to their rehabilitation programme. Finally, a secondary analysis including waiting time as a covariate was performed, as the process evaluation revealed that this variable by chance differed between the two rehabilitation programmes.

4.2.7.4 Statistical analyses Study 3

The primary analysis was an ITT analysis including the four similar measurement time point (t_0 , t_1 , t_7 and t_8). The effect of the rehabilitation programme on the primary and secondary outcomes was estimated by the difference in mean change between the two groups (integrated programme minus existing programme) from baseline to the 12-month follow up using a linear mixed model with a random intercept. The analysis included time (as a categorical variable), group and the interaction between group and time as explanatory variables. Furthermore, the linear mixed model was used to test if the development over time in the two rehabilitation programmes was similar (i.e. test of no interaction between group and time). The underlying assumptions behind a linear mixed model were checked by inspection of plots of random intercepts and residuals. For all outcomes except EQ-5D 5L and MDI, the assumptions were fulfilled, and hence the non-parametric bootstrap method with 1000 repetitions was used to compute p-values and 95% CIs for these two outcome measures. Three secondary analyses were conducted to examine the robustness of the primary analysis: 1) adding waiting time as a covariate, and 2) replacing missing values by the average of non-missing scores at the particular measurement time point, and 3) replacing missing values by the worst possible score (= 100) in the integrated programme and the best possible score (= 0) in the existing programme in an attempt to reveal a worst-case scenario. Additionally, graphs including means at all nine measurement time points (t_0 - t_8) were presented to illustrate mean changes over time for each rehabilitation programme.

4.2.8 Ethical approval and trial registration

The Central Denmark Region Committees on Biomedical and Research Ethics approved the study (journal number: 1-10-72-117-16). The study was assigned an international trial identification number (ClinicalTrials.gov; identifier NCT02884466).

5. Results

In this chapter, the results from the three studies are reported. First, the results from Study 1 are reported. Next, patient flow and baseline characteristics from Studies 2 and 3 are reported together. Then the results from Study 2 are reported, followed by the results from Study 3. Finally, the findings from the process evaluation are reported.

5.1 Study 1

The justification and description of the integrated programme was presented in accordance with the TIDieR checklist (57). The description corresponds with how the integrated programme was when the RCT started.

5.1.1 Item 1. Brief name

The integrated programme.

5.1.2 Item 2. Why

The justification of the integrated programme was described in depth in Chapter 2.

5.1.3 Item 3. What (materials)

The facilities at the rehabilitation centre included classrooms, learning laboratories (e.g. a fully equipped kitchen), conversation rooms, a small hot water pool, and indoor and outdoor fitness facilities with cardio exercise equipment and strength training equipment.

A welcome pamphlet describing the clinical activities was emailed before the first inpatient day.

Most importantly, a pamphlet was provided that contained individualised preparation material focusing on the facilitation of goal-setting as well as physical and psychological preparation before the next inpatient stay.

5.1.4 Item 4. What (procedures)

Activity sheets describing each of the 38 clinical activities in detail are provided in Appendix 8. In Table 9, a completed activity sheet using aqua gymnastics as an example is shown.

The 38 clinical activities were grouped into 10 key components (Table 3). Further, the interrelationship between the 10 key components were illustrated by a web and the person-centred focus of the intervention was emphasised by giving the patient a central position (Figure 7).

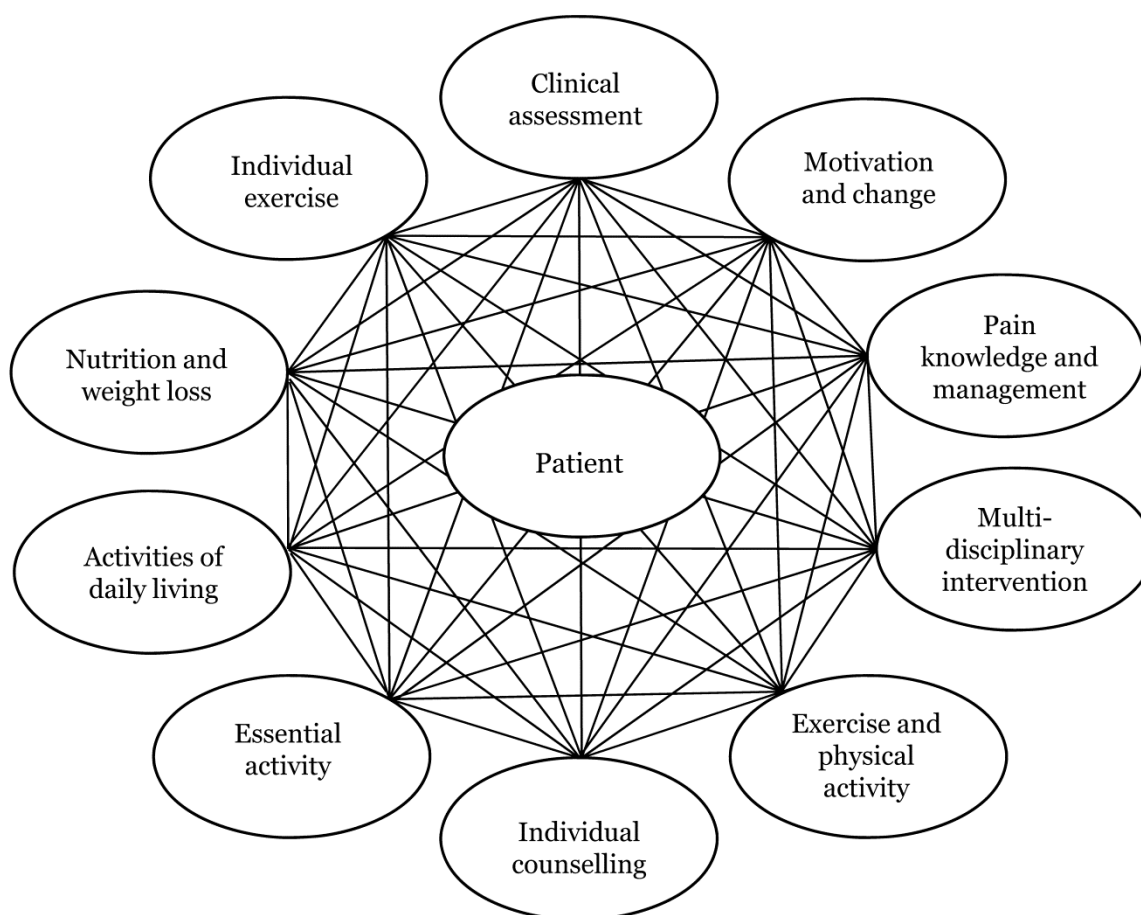


Figure 7. The interrelationship between the 10 key components.

5.1.5 Item 5. Who provided

The providers consisted of six physiotherapists, three occupational therapists, three nurses (educated as coaches and primarily focusing on the psychological aspect of the programme), a rheumatologist, and a nutritional counsellor. The majority of the providers were trained in the Motivational Interviewing approach.

5.1.6 Item 6. How

The mode was a combination of theory and practice using group lecture and dialogue, group sessions (supervised and non-supervised), individual counselling and non-supervised individual exercise; all delivered face-to-face except for the non-supervised exercise.

To support the integration of knowledge, skills and behaviours into the daily life of the patients, a pamphlet (Section 5.1.3) and a telephone call during each of the last two home-based periods was scheduled. Furthermore, the patients were allowed and encouraged to contact the providers via an exercise app twice until the second booster session.

5.1.7 Item 7. Where

This item is reported in Section 4.2.2.

5.1.8 Item 8. When and how much

During the inpatient stays, 38 clinical activities were provided, some of them more than once. To illustrate the different sequence in which the clinical activities were delivered, a graphical depiction of the integrated programme and the existing programme was composed, inspired by a method developed for that purpose (96) (Appendix 9).

5.1.9 Item 9. Tailoring

The integrated programme was partially standardised and partially tailored, the latter to ensure a degree of person-centeredness. The tailoring occurred primarily during: 1) the clinical assessment (including goal-setting), 2) the multidisciplinary conference, 3) individual counselling, 4) exercise and physical activity, and 5) individual exercise (Table 3). Thus, what the individual patient received differed slightly as a result of tailoring.

5.1.10 Item 10. Modifications

This item is reported as part of the process evaluation (Section 5.5.1).

5.1.11 Item 11. How well (planned)

Adherence to inpatient days, time between inpatient stays, and adherence to the clinical activities was recorded in each patient's electronic health record.

5.1.12 Item 12. How well (actual)

This item is reported as part of the process evaluation (Section 5.5.5).

Table 9. The activity sheet developed to describe the 38 clinical activities; using aqua gymnastics as an example (57).

What (procedures) (Item 4)		Aqua gymnastics consisting of aerobic and anaerobic exercises as well as exercises focusing on mobility and stability/balance.
Tailoring (Item 9)		The exercises were chosen and adjusted based on the individual patient.
Who provided (Item 5)		Physiotherapist
How (Item 6)		Group session
When and how much (Item 8)	Number of sessions	Supervised: four sessions Non-supervised: six sessions
	Duration	30 minutes incl. warm-up and cool-down
	Intensity	Borg Scale of Perceived Exertion equaling 11-15

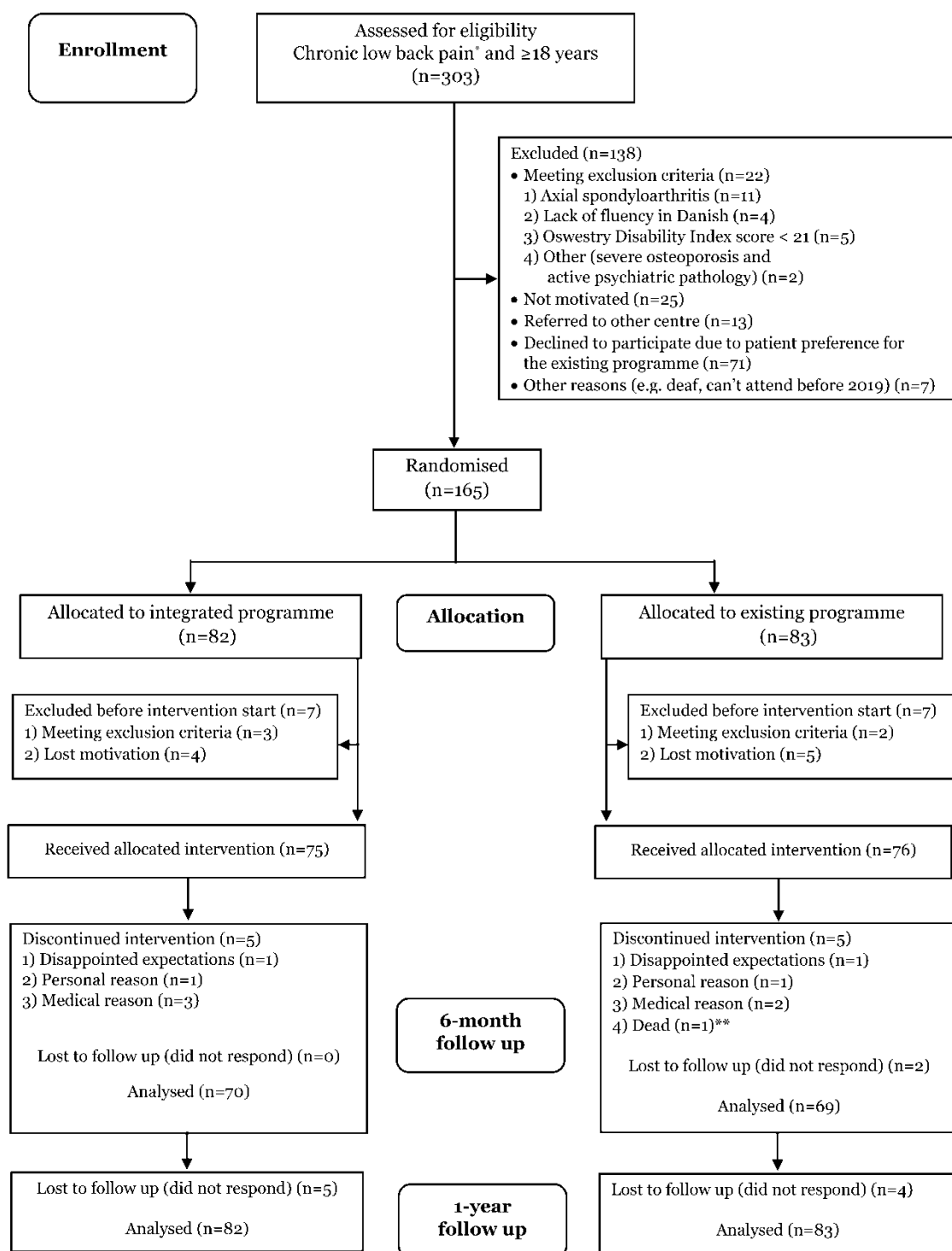
5.2 Patient flow and baseline characteristics in Studies 2 and 3

Recruitment and random allocation started in February 2016 and ended in August 2018. The first rehabilitation programme commenced in September 2016, and the last reached the 6-month follow up in May 2019 and 12-month follow up in November 2019.

In total, 303 patients were screened for eligibility, of whom 138 were excluded for various reasons. Those declining to participate ($n = 71$) did not differ from those randomised with respect to age (mean age 50, age range 22-79), sex (68% women) or diagnosis (data not presented). In total, 165 patients were randomly allocated to the integrated programme ($n = 82$) or to the existing programme ($n = 83$) (Figure 8).

There were no clinically relevant differences in baseline characteristics between patients lost to follow up and patients who completed the 6-month and 12-month follow up, respectively (data not shown).

At baseline, the patients had a mean age of 50 years (SD 13) and the majority (73%) of patients were women. The mean ODI score was 42 (SD 11) and the mean NRS score during the last two weeks was 6 (SD 2). Further baseline characteristics were also similar between the two rehabilitation programmes (Table 10).



* The following ICD-10 diagnosis codes for diseases, signs and symptoms related to chronic low back pain were used: DM40, DM41, DM42, DM43 (not 43.3, 43.4, 43.6), DM47, DM48, DM51, DM53 (not 53.0, 53.1), DM54 (not 54.0, 54.2, 54.6), DM96.1 and DT91.0.

** The reason for dead was unrelated to the trial.

Figure 8. Flow chart of patients in Studies 2 and 3. Combined from (63, 72).

Table 10. Characteristics of the population at baseline. Columns 1 and 2 show patients randomised (n = 165) and columns 3 and 4 show patients providing 6-month follow-up data (n = 139) (63).

	Patients allocated at baseline		Patients providing 6-month follow-up data	
	Integrated programme n = 82	Existing programme n = 83	Integrated programme n = 70	Existing programme n = 69
Sex (women) n (%)	60 (72%)	60 (73%)	52 (74%)	49 (71%)
Age (years)				
Mean (SD)	49 (13)	51 (13)	50 (12)	52 (12)
Range	22-72	25-84	28-72	25-84
Marital status n (%)				
Married	60 (73%)	58 (70%)	51 (73%)	50 (72%)
Single/widowed	22 (27%)	25 (30%)	19 (27%)	19 (28%)
Smokers n (%)				
Yes	24 (29%)	28 (34%)	17 (24%)	23 (33%)
No	58 (71%)	55 (66%)	62 (76%)	46 (67%)
Leg pain n (%)				
Yes	65 (79%)	59 (71%)	55 (79%)	49 (71%)
No	12 (15%)	17 (21%)	12 (17%)	14 (20%)
Don't know	5 (6%)	7 (8%)	3 (4%)	6 (9%)
Employment status* n (%)				
Self-supporting	17 (26%)	16 (25%)	17 (30%)	15 (29%)
Temporary social benefits	9 (14%)	11 (17%)	7 (13%)	8 (15%)
Permanent social benefits	29 (45%)	27 (42%)	24 (43%)	20 (39%)
Age-related pension	8 (12%)	10 (16%)	7 (13%)	9 (17%)
Others	2 (3%)	0 (0%)	1 (2%)	0 (0%)
Education level*				
Low (≤ 12 years)	10 (15%)	14 (22%)	6 (11%)	11 (21%)
Middle (≤ 16 years)	44 (68%)	44 (69%)	40 (71%)	36 (69%)
High (> 16 years)	11 (17%)	6 (9%)	10 (18%)	5 (10%)
Disability** ODI (0-100)				
Mean (SD)	42 (10)	43 (11)	41 (11)	43 (12)
Range	20-68	24-72	20-68	24-72
Back pain intensity*** NRS (0-10)				
Mean (SD)	6 (2)	6 (2)	6 (2)	6 (2)
Pain Self-efficacy PSEQ (0-60)				
Mean (SD)	28 (11)	27 (10)	28 (12)	27 (11)
HRQoL EQ-5D 5L (-0.624-1)				
Mean (SD)	0.567 (0.157)	0.603 (0.118)	0.578 (0.153)	0.599 (0.126)
Depression MDI (0-50)				
Mean (SD)	20 (12)	20 (11)	19 (11)	22 (11)

	Patients allocated at baseline		Patients providing 6-month follow-up data	
	Integrated programme n = 82	Existing programme n = 83	Integrated programme n = 70	Existing programme n = 69
Physical activity n (%)				
Minutes spent in physical exercise during a week				
< 30	42 (51%)	50 (60%)	34 (49%)	41 (59%)
≥30 ≤ 120	32 (39%)	31 (37%)	28 (40%)	26 (38%)
> 120	8 (10%)	2 (3%)	8 (11%)	2 (3%)
Minutes spent in physical activity during a week				
< 30	13 (16%)	20 (24%)	9 (13%)	17 (25%)
≥30 ≤ 300	52 (63%)	55 (66%)	45 (64%)	45 (65%)
> 300	17 (21%)	8 (10%)	16 (23%)	7 (10%)
Hours spent sitting in 24-hour period				
≥ 10	9 (11%)	14 (17%)	7 (10%)	6 (9%)
<10 ≥4	53 (65%)	52 (63%)	47 (67%)	51 (74%)
<4	20 (24%)	17 (20%)	16 (23%)	12 (17%)

*Due to technical issues in the database, data on employment status and education level was only available in $\approx 75\%$ of the patients.

**Due to technical issues in the database, one patient with an ODI score of 20 was included.

***Mean back pain intensity during the last two weeks.

ODI: Oswestry Disability Index; NRS: Numerical Rating Scale; PSEQ: Pain Self-Efficacy Questionnaire; HRQoL: Health-Related Quality of Life; EQ-5D 5L: EuroQol-5 Domain 5-level; MDI: Major Depression Inventory

5.3 Results Study 2

At 6-month follow up (63), the complete case analysis included 70 patients allocated to the integrated programme and 69 patients allocated to the existing programme, thus 12 patients in the integrated programme and 14 patients in the existing programme were lost to follow up (Figure 8).

The between-group difference in the mean change in ODI score revealed an estimate of -0.28 (95% CI: -4.02; 3.45) being neither statistically nor clinically significant (Table 11). Nor were there any statistically significant between-group differences in mean change in the secondary outcomes (Table 11).

Data on physical activity was omitted from the analysis due to post hoc awareness of inadequacies of the three questions measuring it (this issue is also relevant to Study 3). The inadequacies consisted of overlap of categories, the intervals between the categories not being linear, and question 3 (the number of hours spent sitting) was deemed of less relevance in a CLBP population where the alternative to sitting could be lying.

Among those allocated to the integrated programme, the average decrease in ODI scores was from 42 (95% CI: 39; 44) at baseline to 36 (95% CI: 33; 39) at the 6-month follow up. Among those allocated to the existing programme, the average ODI score decreased from 43 (95% CI: 40; 45) at baseline to 37 (95% CI: 34; 40) at the 6-month follow up.

Some secondary analyses were performed. Firstly, on average, patients in both rehabilitation programmes improved on all outcomes (within-group) (Table 11). Secondly, the ITT analysis on the ODI with the last value carried forward did not change the result (mean difference: 0.90 (95% CI: -2.63; 4.44), p-value = 0.614). Thirdly, including waiting time as a covariate did not change the result either (mean difference: -0.92 (95% CI: -4.73; 2.89), p-value = 0.633).

Lastly, no related adverse events or deaths occurred due to the RCT (this issue is also relevant to Study 3).

Table 11. Summary of 6-month follow-up data on primary and secondary outcomes. Between-group and within-group change; complete case analysis (63).

	Between-group*		Within-group	
	Mean (95% CI)	p-value	Integrated programme (n = 70) Mean (95% CI)	Existing programme (n = 69) Mean (95% CI)
Primary outcome				
Disability (ODI)	-0.28 (-4.02 ; 3.45)	0.881	-5.76 (-8.31 ; -3.20)	-5.64 (-8.45 ; -2.83)
Secondary outcomes				
Pain intensity** (NRS)	-0.02 (-0.64 ; 0.59)	0.937	-0.76 (-1.21 ; -0.31)	-0.64 (-1.08 ; -0.19)
Pain Self-Efficacy (PSEQ)	0.05 (-3.47 ; 3.57)	0.978	6.01 (3.48 ; 8.80)	6.22 (3.63 ; 8.80)
HRQoL (EQ-5D 5L)	0.01 (-0.03 ; 0.05)	0.670	0.05 (0.02 ; 0.08)	0.03 (0.00 ; 0.07)
Depression (MDI)	0.62 (-1.98 ; 3.21)	0.639	-3.3 (-5.27 ; -1.24)	-4.57 (-6.52 ; -2.62)

*Adjusted for corresponding baseline value. Existing programme as reference group.

**Mean pain intensity during the last two weeks.

ODI: Oswestry Disability Index; NRS: Numerical Rating Scale; PSEQ: Pain Self-Efficacy Questionnaire; HRQoL: Health-Related Quality of Life; EQ-5D 5L: EuroQol-5 Domain 5-level; MDI: Major Depression Inventory

5.4 Results Study 3

At the 12-month follow up (72), a mean between-group difference in change of -0.53 (95% CI: -4.08 to 3.02) was found in the ODI score; this was neither statistically nor clinically significant (Table 12 and Figure 9). Furthermore, no evidence of a difference in development in ODI score over time was found ($\chi^2(3) = 0.12$, p-value = 0.989) (Figure 9).

Similarly, no statistically significant between-group differences were found for any of the secondary outcomes (Table 12 and Figure 10).

Table 12. Summary of 12-month follow-up data on primary and secondary outcomes. Between-group and within-group change; intention-to-treat analysis (72).

	Between-group		Within-group	
	Mean difference* (95% CI)	p-value	Integrated programme Mean (95% CI)	Existing programme Mean (95% CI)
Primary outcome				
Disability (ODI)	-0.53 (-4.08 ; 3.02)	0.770	-4.55 (-7.08 ; -2.02)	-4.02 (-6.51 ; -1.53)
Secondary outcomes				
Pain intensity** (NRS)	-0.10 (-0.68 ; 0.48)	0.727	-0.58 (-1.00 ; -0.17)	-0.48 (-0.89 ; -0.07)
Pain Self-Efficacy (PSEQ)	0.01 (-3.34 ; 3.37)	0.994	4.43 (2.05 ; 6.82)	4.42 (2.07 ; 6.78)
HRQoL (EQ-5D 5L)***	0.02 (-0.04 ; 0.07)	0.558	0.04 (0.00 ; 0.07)	0.02 (-0.01 ; 0.06)
Depression (MDI)***	1.67 (-1.52 ; 4.85)	0.305	-1.79 (-3.96 ; 0.38)	-3.45 (-5.80 ; -1.11)

*Mean difference = integrated – existing

**Mean pain intensity during the last two weeks

***When analyzing EQ-5D 5L and MDI non-parametric bootstrap method with 1000 repetition two compute p-values and 95% CI's was used.

ODI: Oswestry Disability Index; NRS: Numerical Rating Scale; PSEQ: Pain Self-Efficacy Questionnaire; HRQoL: Health Related Quality of Life; EQ-5D 5L: EuroQol-5 Domain 5-level; MDI: Major Depression Inventory

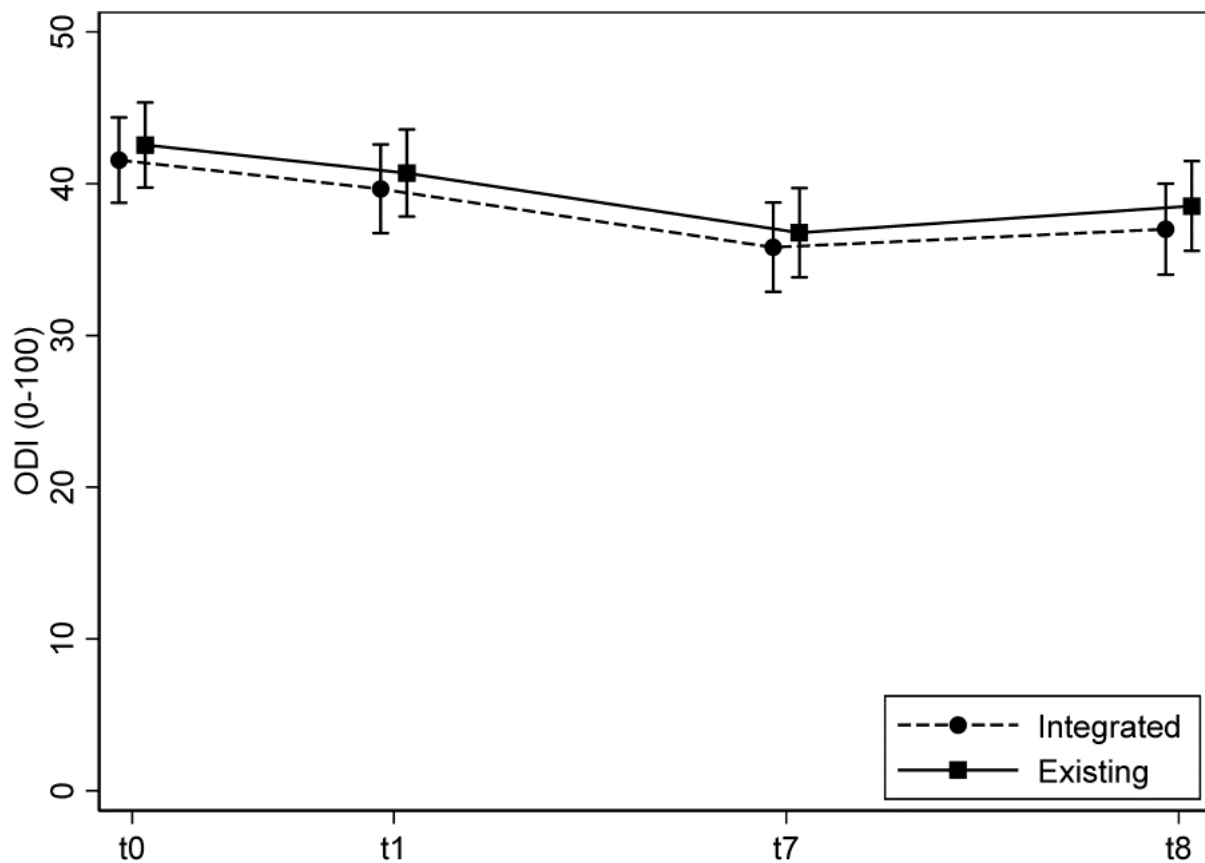


Figure 9. Mean change in back-specific disability measured by the Oswestry Disability Index (ODI) during the 12-month follow-up period, including the four identical measurement time points: t0 = baseline, t1 = before intervention start, t7 = before the 6-month follow-up visit, and t8 = the 12-month follow up (from the linear mixed model). The time between t0-t1 was 4 months (= mean waiting time), the time between t1-t7 and between t7-t8 was 6 months. (72)

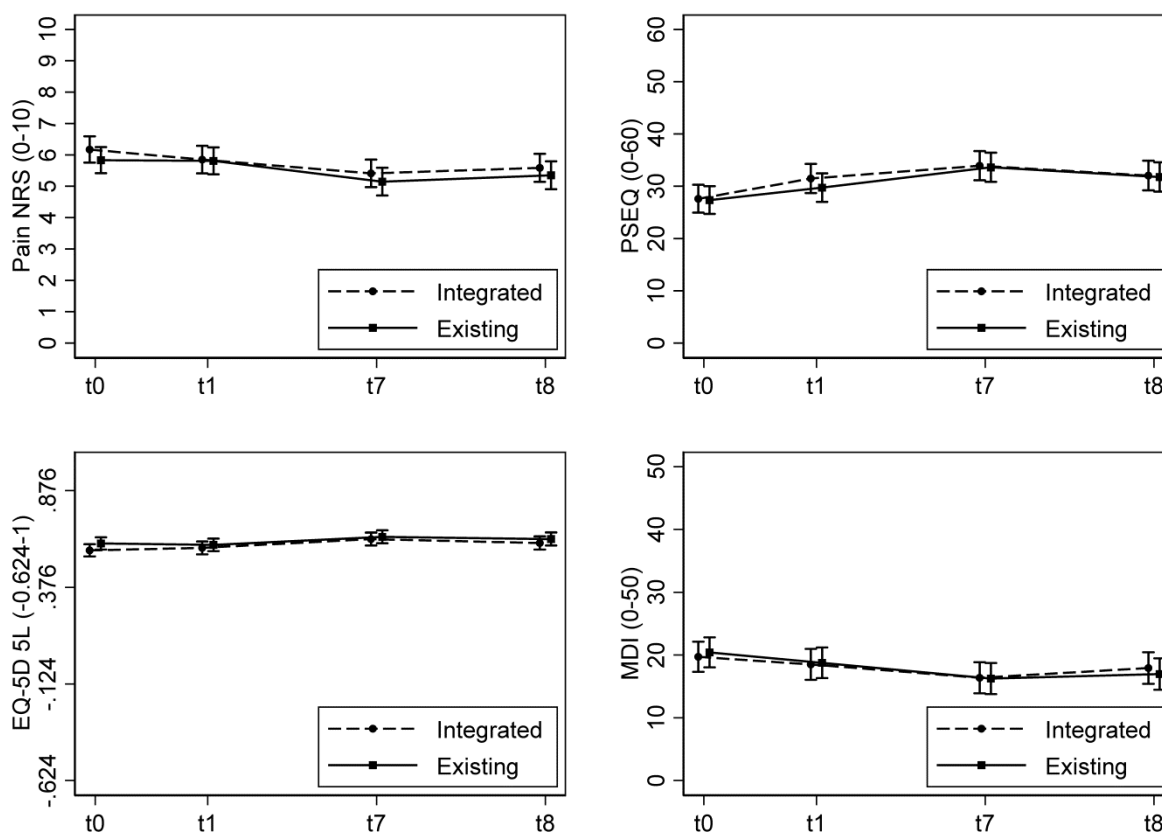


Figure 10. Mean changes in secondary outcomes during the 12-month follow-up period including the four identical measurement time points: t0 = baseline, t1 = before intervention start, t7 = before the 6-month follow-up visit, and t8 = the 12-month follow up. The time between t0-t1 was 4 months (= mean waiting time), the time between t1-t7 and between t7-t8 was 6 months.

NRS: Numerical Rating Scale; PSEQ: Pain Self-Efficacy Questionnaire; EQ-5D 5L: EuroQol-5 Domain 5-level; MDI: Major Depression Inventory. (72)

Among those allocated to the integrated programme, the average decrease in ODI scores was from 42 (95% CI: 39; 44) at baseline to 37 (95% CI: 34; 40) at the 12-month follow up. Among those allocated to the existing programme, the average ODI score decreased from 43 (95% CI: 40; 45) at baseline to 39 (95% CI: 36; 41) at the 12-month follow up.

Within-group improvements were found on most outcomes, except for depressive symptoms in the integrated programme and HRQoL in the existing programme (Table 12).

The secondary analysis adjusted for waiting time did not change the result (mean difference: -0.41 (95% CI: -4.02; 3.20), p-value = 0.824). Neither did the secondary analysis replacing missing values by the average of non-missing scores at the particular time point (mean difference: -0.60 (95% CI: -3.78; 2.58), p-value = 0.712). Finally, the secondary analysis replacing missing

values with the worst possible ODI score (= 100) for the integrated programme, and by the best possible ODI score (= 0) for the existing programme, changed the results (mean difference: 19.79 (95% CI: 13.80; 25.77), p-value = 0.000).

The outcome trajectories, including all measurement time points from t0-t8, revealed that changes over time were similar in the two rehabilitation programmes, regardless of which outcome measure was chosen (Figure 11).

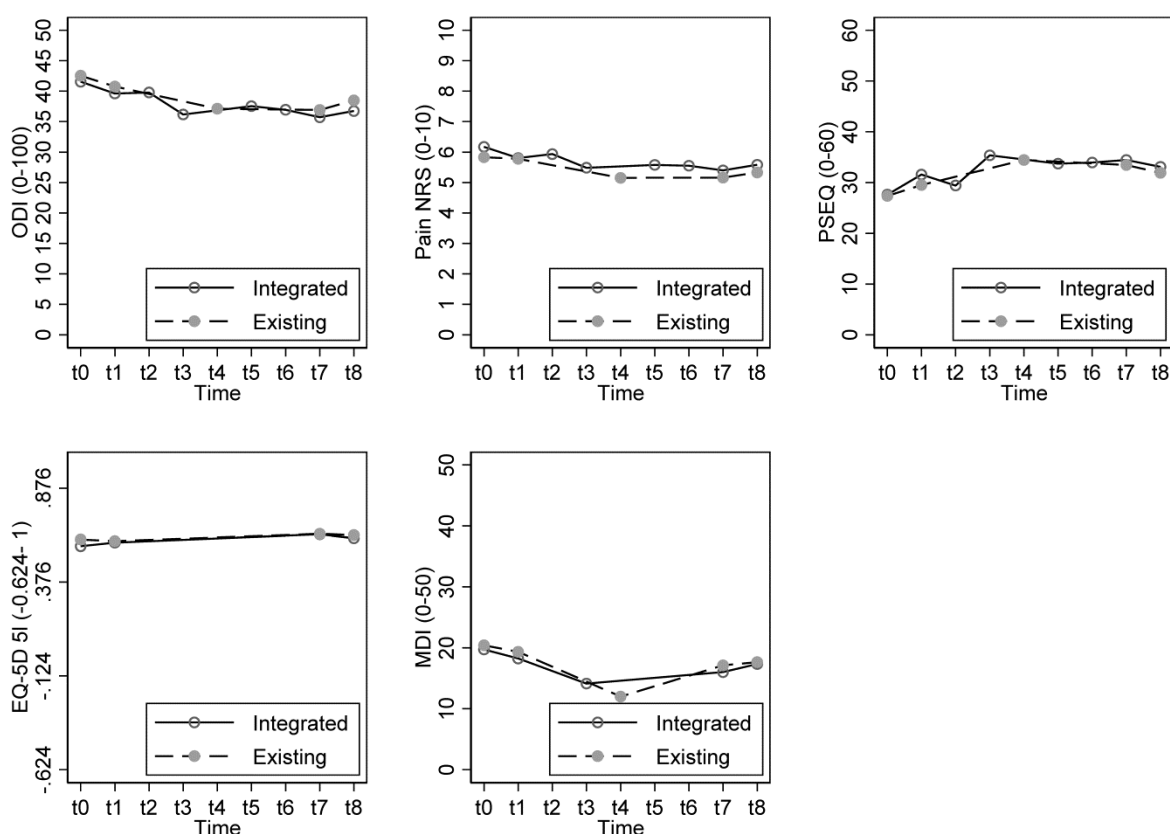


Figure 11. Mean changes over the 12-month follow up-period, including all nine measurement time points (t0-t8). Notice that the distances between the measurement time points are displayed as equidistant even though they were not. t0 = baseline, t1 = before the pre-admission day (integrated programme) and before the 4-week inpatient stay (existing programme), t3 = the end of the 2-week inpatient stay (integrated programme), t4 = the end of the 4-week inpatient stay (existing programme), t5 = before the initial booster session (integrated programme), t6 = before the second booster session (integrated programme), t7 = before 6-month follow-up visit (integrated programme + existing programme), and t8 = the 12-month follow up (integrated programme + existing programme).

ODI: Oswestry Disability Index; NRS: Numerical Rating Scale; PSEQ: Pain Self-Efficacy Questionnaire; EQ-5D 5L: EuroQol-5 Domain 5-level; MDI: Major Depression Inventory. (72)

5.5 Process evaluation

5.5.1 Process research question 1

“Was the integrated programme modified during the trial?”

Yes, it was modified, and the most relevant modifications were:

- Employment of a psychologist in September 2017.
- Change in the number of hours of aqua gymnastics from April 2018 (from 1 hour and 25 minutes per week to 1 hour per week). Furthermore, the setting was changed from the rehabilitation centre’s own hot water pool to a public hot water pool.
- Continuous challenges with implementing the exercise capacity test.

The modifications were identical to modifications in the existing programme.

5.5.2 Process research question 2

“What relevant changes to referral, administration, management etc. occurred during the trial?”

The most relevant changes were:

- Postponement of allocated inpatient stays initiated by the patients. This caused a more or less persistent low number of patients attending most of the integrated programmes despite overbooking. Therefore, fewer integrated programmes were scheduled in the last part of the RCT.
- Low adherence to the 6-month follow-up visit reaching approximately 20% non-adherence in each rehabilitation programme.

Examples of modifications to the integrated programme (process research question 1), and relevant changes to referral, administration, management etc. (process research question 2) were reported collectively in Appendix 10.

5.5.3 Process research question 3

“The stakeholders’ experiences with and perspectives of the integrated programme”.

Findings from a Masters student’s interview with the stakeholders in February 2018 are briefly summarised below.

The stakeholders:

- Felt involved in the development of the integrated programme.
- Had positive expectations about the integrated programme – though when implementing it, the overall aim (integrating knowledge, skills and behaviours) got lost in daily clinical

practice, and they felt that they either had to justify or apologise for the design of the integrated programme without being able to explain the overall aim to the patients.

- Emphasised the pamphlet and phone calls as important elements, nevertheless, they felt that using these supportive elements reduced time for other important tasks in daily clinical practice.
- Expressed concerns about whether they were able to make an actual difference to patients attending the integrated programme e.g. due to time spent on re-connecting with patients and on follow-up talks.
- Comparing the two rehabilitation programmes, they favoured the existing programme, as they were concerned about the differences in group cohesion, the different time periods to disconnect from daily lives, and the different abilities to informally recap with patients attending the integrated programme.

5.5.4 Process research question 4

“The patients’ experiences with the integrated programme and with integrating the programme into their daily life”.

Findings from a Masters student’s interview with patients in February 2019 are briefly summarised below.

The patients:

- Had various opinions on the delivery of the integrated programme; some found the inpatient days too compressed, whereas others felt they had too much time on their own; some felt there were too few inpatient days and wondered if the integrated programme was effective enough, whereas others found the combination of inpatient days and home-based activities meaningful in terms of their personal life.
- Felt that the alternation between inpatient stays and home-based activities, as well as the phone calls, provided them with relevant feedback and motivation, and for that reason, they experienced a relevant transition from the inpatient stays to their daily life. In contrast, some patients did not feel that they succeeded in integrating their efforts into their daily life.
- Profited from the group cohesion.
- Experienced that the efforts by the providers were meaningful, and they felt seen, heard and understood while attending the integrated programme.
- Felt that they were able to accept their situation and disability after attending the integrated programme.

5.5.5 Process research question 5

"Was the integrated programme delivered as planned?"

Yes, it was delivered as planned. Adherence to the inpatient days was excellent in both rehabilitation programmes, as the majority of patients attended $\geq 80\%$ of the inpatient days, shown by 99% of the patients in the integrated programme, and 100% of the patients in the existing programme. For that reason, the pre-specified per-protocol analysis in Study 2 was deemed unnecessary.

5.5.6 Process research question 6

"Did waiting time differ between the two programmes?"

Yes, by chance, waiting times differed between the two rehabilitation programmes with a mean of 105 days (SD 9) in the integrated programme and a mean of 141 days (SD 10) in the existing programme.

6. Discussion

This chapter begins with a section reflecting on substantial elements of the integrated programme including Study 1. Next, the PPI process is discussed. Following that, the main results from the RCT (Studies 2 and 3) are discussed in two parts. In part one, the main results will be compared with the literature, and in part two, the main findings from the process evaluation will be discussed in relation to the main results from the RCT. Finally, strengths, limitations and external validity will be discussed.

6.1 Reflections on the integrated programme including Study 1

To ensure a systematic approach, this section is divided into subsections following the stages outlined in the MRC's guidance (32) (Figure 1). Those of the 12 related elements in the MRC's guidance considered important for discussion are *italicised*.

6.1.1 Development stage

Considering *Identifying the evidence base* and *Identifying/developing theory*. The justification of this RCT was described in Study 1 based on the evidence, and the underlying theories being the biopsychosocial approach and the Chronic Care Model (57). The applicability of those theories could rightly be questioned given the null effect. While the biopsychosocial approach, responding to the aetiology of CLBP has been widely accepted and scientifically endorsed since the 1980s (15, 16, 18-20), the Chronic Care Model (51-53) does not have quite the same evidence base when it comes to musculoskeletal conditions (48, 97-100). The latter will be further elaborated in Section 6.3.

Considering *Modelling processes and outcomes*. Two aspects need elaboration, namely the detailed description of the integrated programme following the TIDieR checklist (58) published as Study 1 (57), and the chosen outcome domains and outcome measures.

Firstly, the experiences using the TIDieR checklist (58) ensured a systematic, detailed description of the integrated programme. As part of the description, activity sheets were developed (Table 9), allowing us to structure and standardise the description of each of the 38 clinical activities. Further details, e.g. about the specific exercises, could have been added. However, it was the type of clinical activities rather than the details about each of them that we wanted to capture and describe (101). Furthermore, it would have been ideal to describe the evidence and/or theory behind each of the 38 clinical activities, and add it either to the activity sheets and/or to the logic model. However, the clinical activities were designed to support and enhance each other (Table 3 and Figure 7), and it was the bundle of clinical activities, and not one specific clinical activity, that we wanted to assess (31). For this reason, the evidence and/or theory behind each clinical activity was not described.

Secondly, disability was chosen as the preferred outcome domain based on a combination of results from the PPI process and the international recommendations about core outcomes for trials in the field of LBP (81, 82). It could be argued that measuring the domain of integrating knowledge, skills, and behaviours into daily life, which was the intended target of the integrated programme, would have been more appropriate. To our knowledge, no outcome measure to capture this specific domain has been developed and validated. Therefore, disability was chosen as a proxy for this integration even though the relationship between disability and integration of knowledge, skills, and behaviours is unknown. Based on its widespread use and endorsement (80, 82) together with its validity (78), the ODI seemed an appropriate outcome measure.

An alternative outcome measure could have been the Patient Activation Measure (PAM) which assesses the domain of patient knowledge, skill, and confidence required for self-management (102). The use of the PAM in a CLBP population has been sparsely investigated, and the Danish version of the PAM has only been validated in patients with dysglycaemia (102). For these reasons, the PAM was considered irrelevant in this RCT.

Pain self-efficacy is a psychological domain (2) that seeks to capture a person's expectation to perform a particular behaviour or task, and their confidence in being able to do so despite pain (86). The domain of pain self-efficacy may be closer to the domain of integrating knowledge, skills, and behaviours, than the domain of disability. Despite that, no difference in pain self-efficacy was found between the two rehabilitation programmes, albeit the RCT was not powered to detect differences in this outcome.

Given the broad biopsychosocial dimensions of the chosen outcomes, and the consistent finding of no between-group differences in this RCT, it is unlikely that the results of the RCT would have been any different using any other outcomes.

6.1.2 Feasibility/piloting

Considering *Testing procedures*. With the value of hindsight, running a pilot RCT would have been preferable, as it could possibly have revealed (some of) the unexpected challenges established in the process evaluation.

Considering *Determining sample size*. As described in Section 4.2.7.1, the literature suggests a 4-point between-group difference in change as being of minimum clinical importance (78, 94). At the time we calculated our sample size, no RCTs comparing two or more rehabilitation programmes had used the ODI as a primary outcome measure, and only one RCT comparing a rehabilitation programme to outpatient physiotherapy had used the ODI in an RCT (95). The RCT found a 9-point decrease on the ODI at the 6-month follow up following a 3-week ambulatory programme from Monday to Friday, and a 4-point decrease on the ODI following outpatient

physiotherapy (95). Thus, we found it reasonable to detect a 10-point decrease in ODI at the 6-month follow up following participation in the integrated programme, and a 5-point decrease in ODI at the 6-month follow up following participation in the existing programme. The SD of the difference was estimated at 10 points, informed by our previous feasibility-test. This is at the lower end compared with that reported in the literature that estimates SD to be between 10-21 points in a population of patients with CLBP when using the ODI (78). Collectively, the chosen estimates correspond to an effect size of 0.5, which could be considered too ambitious in light of an effect size at 0.23 when pooling 16 RCTs comparing rehabilitation with usual care (7). In contrast, comparing two comprehensive rehabilitation programmes with an estimated effect size of 0.5 could be considered reasonable in light of an effect size of 0.68 when pooling 19 RCTs comparing rehabilitation with physical treatment (7). The same rationale and an estimated between-group difference in change of 5 points on the ODI was used in an RCT from 2018, comparing two rehabilitation programmes at the 12-week follow up (103), which supports our estimation. Finally, we were less ambitious than another RCT from 2018 comparing two rehabilitation programmes that looked for a between-group difference of 8 points on the ODI at the 12-month follow up (104). For practical reasons we had to allow for a degree of pragmatism by making a realistic estimation based on potentially eligible patients, inclusion- and exclusion criteria, and the number of patients willing to participate and complete the RCT (79).

6.1.3 Evaluation stage

Considering *Assessing effectiveness*. A single-centre, pragmatic, two-arm parallel RCT was chosen as an appropriate design to assess effectiveness. Strengths and limitations of this design will be discussed in Section 6.5, except for the pragmatic attitude and the related PRECIS-2 score, which will be discussed here.

To assess the degree to which the self-identified pragmatic attitude of our RCT was truly so, the PRECIS-2 tool was used retrospectively. Had the assessment been done prospectively, the scores would probably have been slightly different, especially in relation to domain 4 (Organisation). Prospectively, it was expected that the integrated programme could easily be adapted to the usual organisation, thus, an a priori score would probably have been 5. Retrospectively, it was learned that the integrated programme did not just fit in to the usual workflow at the rehabilitation centre; e.g. the realisation that the pamphlet and phone calls were new elements to the providers and not easily implemented, and the administrative changes related to booking. For these reasons, the retrospective score was 3 (Table 7 and Figure 5). Despite this difference between the hypothesised prospective and the actual retrospective score, the PRECIS-2 assessment is still thought to be useful to stakeholders when understanding applicability of the results (105). Considering and scoring design choices (Table 7) and depicting them in the PRECIS-2 wheel (Figure 5) confirmed

that the design was primarily pragmatic, thus, we were able to answer questions about real-world effectiveness (73, 74, 105).

Considering *Understanding change processes*. In the following, the methods used in the process evaluation will be discussed. The findings from the process evaluation and their relationship with the RCT results will be discussed in Section 6.4.

No gold standard exists as how to design and conduct a process evaluation (41, 43). The choice of process research questions and the methods used to clarify them in this research project were explicitly presented as recommended (43) (Figure 4). Ideally, it could have been valuable to address the experiences of patients attending the existing programme. Many other process research questions could have been addressed as well, but it is considered preferable to answer a few closely selected process research questions thoroughly than to answer several poorly (41, 61).

In this research process, the PhD candidate analysed the process data before the outcome data. Arguments exist for both separation and integration of the process evaluator and the outcome evaluator (42). A disadvantage of integration is the risk of inducing biased interpretations of the outcome data (42). Thus, the PhD candidate was blinded when performing the analyses of outcome data. It was impossible for the PhD candidate to be independent of the RCT when analysing the process data, and therefore, it is seen as a strength that two Masters students performed the in-depth analyses of process research questions 3 and 4. An advantage of the integration is that the PhD candidate was close enough to the clinic to record the problems and understand why they occurred, as well as being able to contribute with effective solutions to the problems in collaboration with the stakeholders.

In terms of the process evaluation, some essential methodological considerations relate to: 1) recruitment of patients and stakeholders, and 2) timing of data collection (41, 42).

Firstly, in terms of process research question 4, patients from different groups were included, and process research questions 1, 2 and 3 included every stakeholder employed at the rehabilitation centre. The broad inclusion of different groups of patients and every stakeholder was thought to give a wide perspective on implementation, mechanisms of change, and context affecting and interacting with the integrated programme.

Secondly, process research questions 1, 2 and 4 were addressed at multiple time points. The in-depth assessment of process research questions 3 and 4 were performed at a single time point for practical reasons. Data collection at multiple times points is valuable as experiences and perceptions may shift over time, and the conclusions drawn will be situated in the time in which they were collected (41, 42).

Some of the data gathered from the process evaluation were formative, being reported back quickly in order to improve the quality of the integrated programme, and some of it was summative, being reported after the evaluation of the RCT to help understand the effectiveness of the integrated programme (61). Each attempt to improve the quality of the integrated programme (e.g. the attempts to support implementation of the pamphlet, the phone calls, and the exercise capacity test) was openly documented in a process evaluation document (Appendix 10). For that reason, the potential improvements were not thought to compromise the results nor the external validity (41, 42).

In all, the systematic, rigorous and transparent approach to the process evaluation is considered to improve the validity of the RCT results (43).

Considering *Assessing cost-effectiveness*. Due to the time frame of the research project, an economic evaluation was not conducted even though it certainly could have contributed with an important aspect, and thus, made the results more useful for decision-makers (32). By choosing the EQ-5D as a secondary outcome measure, the possibility for an economic evaluation remains.

6.2 PPI

PPI was not incorporated into this research project according to a pre-specified plan. The project team had little experience, and neither the patients nor the stakeholders had any experience with PPI in research. Research training was not done, and whether to perform it or not is widely discussed (66, 106). Despite that, we experienced that all patients had introspection and were able to be critical in a constructive way, thus contributing with their lived experiences rather than being “professional patient representatives”. The PhD candidate facilitated and encouraged PPI to the best of her ability, but she was not specifically trained in this. Fortunately, the collaboration became two-way, as patients and stakeholders addressed the PhD candidate directly if they had ideas or insights they wanted to share. This maintained a continuous connection to knowledge, perspectives and lived experiences obtained in daily clinical practice.

Evidence regarding the impact of PPI remains weak (107), but PPI is thought to reduce research waste, and is proposed to be done for moral reasons (65). It is definitely the experience that PPI has improved the current research project by offering new ideas, improving understanding of patients’ and stakeholders’ perspectives, and thus, helped us to identify and hopefully avoid problems we might not have anticipated.

6.3 Main results from the RCT compared with the literature

Remembering that the optimal dose, content or delivery of a multidisciplinary rehabilitation programme remain unknown (7) (Chapter 1), the results from Studies 2 and 3 will be compared with those reported in the literature. Firstly, the results will be compared with the evidence base

consisting of RCTs comparing two or more rehabilitation programmes in patients with CLBP. Secondly, the results will be compared with the findings from RCTs assessing the effect of booster sessions in the musculoskeletal field.

To our knowledge, this is the first RCT comparing two different ways of delivering a rehabilitation programme while keeping the content stable. Six RCTs, with a total sample size of 945 participants, were of special interest when undertaking comparison with our results (46-49, 103, 104). The similarities between the six RCTs and our RCT were that they included CLBP patients, they compared two or more different rehabilitation programmes, and they used disability as the primary outcome domain. However, they had some substantial differences:

- The populations included had different levels of disability at baseline (46-49, 103, 104).
- Four different outcome measures were used to measure disability, namely the ODI (103, 104), the Roland Morris Disability Questionnaire (46, 49), the Quebec Back Pain Disability Scale (47), and the Pain Disability Index (48).
- Some of the RCTs compared different dose and content (47, 49, 103), and some compared identical rehabilitation programmes plus an additional component in one programme e.g. involvement of spouses (46), more specifically tailored interventions (104), or subsequent booster sessions (48). The last-mentioned RCT was referred to in Chapter 2, as it was the only previous RCT in the CLBP field aiming to assess the effect of adding booster sessions (seven phone calls) (48).
- Two RCTs used short-term follow up (3 months or less) (103, 104), two RCTs used medium-term follow up (3 to less than 12 months) (47, 49), and yet four RCTs incorporated long-term follow up (12 months or more) (46, 48, 49, 104).

These differences limit direct comparison with our RCT. However, regardless of the differences, the results of the six RCTs were similar to the current RCT, and found no significant differences in disability when comparing two or more rehabilitation programmes.

In general, evidence supports that rehabilitation is a beneficial process (33). As well, when it comes to research into rehabilitation and CLBP, evidence supports the effect of rehabilitation (7, 14, 22-24). However, when adding our results to the current evidence base in the CLBP research field, it is obvious that there is an unsolved Gordian knot, as it seems difficult to demonstrate if and how much dose, content and delivery of a specific rehabilitation programmes matters (33). It could be discussed whether the reasons for the repeated null effect in the RCTs could be due to sample sizes, differences in population, outcome measures, designs of the rehabilitation programmes, follow up lengths, or general quality of the trials. Rather, it should be considered whether the null effect are due to the identical comprehensive nature of the rehabilitation programmes being compared in the

RCTs, or whether the complex interplay between individually biopsychosocial factors driving disability in patients with CLBP makes it difficult to improve disability with programmes that more or less have one-size-fits-all approaches (33).

While only the aforementioned single RCT assessed booster sessions in the CLBP and rehabilitation field (48), a review including three RCTs evaluated the effect of booster sessions in addition to exercise therapy in patients with hip and/or knee osteoarthritis. It found moderate evidence for long-term effectiveness with respect to physical function (97). This result is in line with one additional RCT (98), but in contrast to two RCTs (99, 100) in the same field, and as well in contrast to our results. Overall, the evidence base in terms of the effect of adding booster sessions to exercise interventions in the musculoskeletal field remains equivocal.

6.4 Main findings from the process evaluation and their relationship to the main results from the RCT

Overall, the initial analyses of process data provided prospective insight into what we might subsequently find in terms of the overall effect. The findings from changes in referral, administration, and management occurring during the RCT (process research question 2), and the stakeholders' experiences with and perspectives of the integrated programme (process research question 3) are collectively thought to be a drawback of the integrated programme, thereby attenuating any difference between the two rehabilitation programmes.

An exhaustive discussion of all nuances in the process evaluation is not presented, but key examples are given below according to each process research question.

6.4.1 Process research question 1

“Was the integrated programme modified during the trial?”

Allowing for a high degree of flexibility in the delivery of the clinical activities as a part of the pragmatic attitude of the RCT (Table 7 and Figure 5) meant that two major modifications occurred during the RCT. Those were the employment of a psychologist and the change in number of hours and setting of aqua gymnastics, both of which occurred in the integrated and the existing programmes. For this reason, the modifications are not suspected to have influenced the results (108).

Further, the continuous challenge with implementing the exercise capacity test resulted in lack of valid data. Thus, before opening up the data, it was decided to exclude data on exercise capacity from the statistical analyses. In this non-blinded RCT, data on a relative objective outcome such as exercise capacity would have been interesting, as a part of the rehabilitation programme aimed to improve physical outcomes. On the other hand, exercise capacity was not a major component in the

rehabilitation programmes and it may not on its own have revealed a positive effect of successful integration of knowledge, skills and behaviours.

6.4.2 Process research question 2

“What relevant changes to referral, administration, management etc. occurred during the trial?”

The patient-initiated postponements of allocated inpatient stays and the low adherence to the 6-month follow-up visits (despite good overall adherence) is not thought to have a direct effect on the results. However, the effect these events induced amongst the providers, and thus the indirect effect they could have caused, is a cause for concern. This is further discussed in Section 6.4.3.

6.4.3 Process research question 3

“The stakeholders’ experiences with and perspectives of the integrated programme”

The changes to referral, administration, and management (Section 6.4.2) spread frustration across the whole setting and could potentially have obstructed the success of the integrated programme by inducing a lack of endorsement for, and belief in, the integrated programme.

The continuous challenges with implementing the pamphlet and the phone calls, which were thought to support integration of knowledge, skills, and behaviours, was rather concerning as it could potentially have served to attenuate any differences between the two rehabilitation programmes.

As these findings emerged, it became obvious that the logistics of implementing the integrated programme in daily clinical practice was a challenge. This could potentially have caused bias in favour of the existing programme, and thereby potentially induce a negative effect on the results.

On reflection, the changes required to implement the integrated programme in daily clinical practice might have been overlooked despite the thorough development and feasibility-testing including PPI; pointing back to the lack of a pilot RCT (Section 6.1.2).

6.4.4 Process research question 4

“The patients’ experiences with the integrated programme and with integrating the programme into their daily life”

The patients were interviewed 1 year after the stakeholders, and thus, in general, their experiences did not seem to be affected by the experiences and perspectives of the stakeholders. Thus, the patients’ experiences were not thought to be a drawback of the integrated programme.

6.4.5 Process research question 5

“Was the integrated programme delivered as planned?”

Adherence to the inpatient days was excellent, and for that reason, the planned per-protocol analysis was not performed. The lack of adherence, therefore, was not thought to have had an effect on the main results.

6.4.6 Process research question 6

"Did waiting time differ between the two programmes?"

Adjustment for waiting time performed in Studies 2 and 3 did not affect the estimates, thus, waiting time can be rejected to have influenced the main results.

6.5 Strengths and limitations in Studies 2 and 3

When interpreting the results of the RCT, some considerations about the design and the circumstances that occurred during the RCT need elaboration.

The RCT has some strengths according to its design:

- The RCT design is considered the gold standard for studies about effectiveness of interventions, as, if conducted properly, the basic premise is that they minimise known and unknown confounding (109, 110).
- The parallel design, implying that the two rehabilitation programmes were evaluated simultaneously, is expected to ensure that changes in the clinical activities and/or the organisation during the RCT would have an equal impact on the two rehabilitation programmes.
- Stratified randomisation on the basis of baseline disability was performed to achieve approximate balance in mean disability levels and thereby keep this variable equally distributed between the two rehabilitation programmes (79, 108, 109).
- The database that generated the random allocation sequence was set up to use random blocks of six. This allocation concealment prevents selection bias (79, 108-110). Furthermore, the block randomisation ensured that the number of patients could not differ more than three patients between groups (79, 108).
- The PhD candidate performing the statistical analysis was blinded to treatment allocation thereby also preventing bias (31, 79, 108, 109).
- In Study 3, an ITT analysis was performed. This handles bias introduced by missing data by keeping every patient randomised in the RCT, thereby supporting the strengths of the RCT design (108).
- Data were analysed in accordance with pre-specified statistical analysis plans (Appendices 6 and 7) thus, ensuring that the statistical analyses and thus, the results, were not data-driven (111).

Furthermore, some additional strengths in terms of the circumstances that occurred during the RCT need highlighting.

Firstly, patient characteristics were comparable between rehabilitation programmes at baseline. This is vital as a serious imbalance could potentially influence the clinical course and thereby the results (108, 110).

Secondly, patients dropping out from the RCT were similar in numbers and baseline characteristics between groups, hence attrition bias is not suspected (108).

Thirdly, baseline characteristics were comparable between patients lost to follow up and patients who completed the 6-month and 12-month follow up, respectively, thus minimising the risk of selection bias (110).

Fourthly, high adherence to the inpatient days confirmed that the patients actually participated in the rehabilitation programmes (109).

Lastly, the targeted sample size was reached.

This RCT also has a number of limitations which could possibly have distorted the results:

- Theoretically, due to the lack of blinding of patients and providers, knowledge about treatment allocation could have introduced information bias especially in light of all outcome measures being self-reported. Firstly, patients were aware of the comparator programme and the allocation when answering the questionnaires (except from the baseline questionnaire). Secondly, providers were similarly aware and the process evaluation revealed that they compared the two rehabilitation programmes in favour of the existing programme. This could have affected the providers' interaction and communication with the patients. However, since there was no difference between the two rehabilitation programmes either at 6-month or at 12-month follow up, the risk of information bias was not suspected to have biased the results.
- Lack of blinded providers delivering the rehabilitation programmes to non-blinded patients in the same rehabilitation centre, at the same time induced the risk of contamination. This meant that the concept of integration of knowledge, skills, and behaviours could have been transferred to patients in the existing programme by themselves or by the providers. Contamination can dilute any differences between the two rehabilitation programmes, and thereby reduce the point estimate of effectiveness of the integrated programme leading to the risk of a type II error (112). Despite the risk of contamination, the parallel design was

preferred rather than a staggered design where external circumstances could have been different between the rehabilitation programmes (see strengths).

- The degree to which the integration of knowledge, skills, and behaviours led to contamination could potentially have been revealed by measuring adherence to home-based activities, which was not done. Lack of this data precluded us from assessing whether patients in the integrated programme actually integrated the acquired knowledge, skills and behaviours into their daily lives to a higher extent than patients in the existing programme. Thus, we failed to have valuable information to improve our understanding of the mechanisms of treatment failure (108). Measuring adherence to home-based activities is a challenge as there are no obvious good measurement of that domain, it often requires a huge amount of effort from the patients, it can easily be subject to information bias, and it can produce myriad of data.

Internal validity is determined by how well the design, data collection, and analyses are carried out, and can be threatened by bias and confounding (113). Due to their nature, well-conducted RCTs are known to have a strong internal validity as a starting point. When counterbalancing the strengths and limitations according to the design and the circumstances occurring during this RCT, risk of bias and confounding are limited, which supports strong internal validity.

6.6 External validity

A prerequisite for making inferences about external validity, also known as generalisability, is that the internal validity needs to be strong (113). However, there will always be a trade-off between internal and external validity, and thus using a pragmatic attitude maximises external validity but compromises internal validity (32, 105). We attempted to conduct this RCT with a high degree of rigour while also providing strong external validity by applying a pragmatic attitude.

Some aspects with respect to external validity need elaboration. In theory, it could be assumed that those being referred, and those willing to participate differed from the general population of patients with CLBP. Differences in those being referred or differences between those randomised and those declining to participate ($n = 71$) would affect external validity. We have no reason to suspect differences in referral patterns and there were no differences between those randomised and those declining to participate ($n = 71$) with respect to baseline characteristics (Section 5.2). These factors along with the broad inclusion- and exclusion criteria, due to the pragmatic attitude of the RCT, support the external validity.

A multicenter study could have enhanced the external validity further as the integrated programme then would have been assessed in various contexts, by different providers, and included a more diverse sample (79). A multicentre study involving all three departments of the rehabilitation centre in Denmark was considered, but unfortunately not possible due to available resources.

7. Conclusion

Firstly, a thorough justification and description of the integrated programme for patients with chronic low back pain was performed.

Secondly, the integrated programme did not lead to improved back-specific disability or other outcomes for patients with chronic low back pain when compared with the existing programme at 6-month follow up.

Thirdly, the integrated programme did not lead to improved back-specific disability or other outcomes for patients with chronic low back pain when compared with the existing programme at 12-month follow up.

8. Implications

This properly planned research process represents the culmination of many decisions, resulting in a successful evaluation of the integrated programme. The dissertation contributed new knowledge about the delivery of multidisciplinary rehabilitation programmes in patients with CLBP. In spite of the null effect, this is still an important finding and some implications deserve consideration.

8.1 For clinical practice

Study 1 succeeded in justifying and describing the integrated programme in detail, allowing for implementation in clinical practice.

The results of Studies 2 and 3 did not provide evidence to support that changing the delivery of a multidisciplinary rehabilitation programme improved the outcome in terms of disability in a population of patients with CLBP after 6 and 12 months. This is useful knowledge in clinical practice where multiple rehabilitation programmes for patients with CLBP exist. Thus, clinicians referring to and clinicians providing, a rehabilitation programme can rightly have doubts about whether a specific rehabilitation programme is effective or not. This dissertation provided an evidence base from which clinicians with their mind at rest can tell their patients that as long as the content stays the same, it is without consequence to change the delivery of the rehabilitation programme. This would induce a positive effect on the patients referred to the rehabilitation programme and for whom it can be necessary to be reassured about the effectiveness of the rehabilitation programme they are referred to and into which they are expected to put a lot of personal effort.

8.2 For research

The thorough description of the integrated programme provided in Study 1 enables other researchers to replicate or build on our research findings. It also enables reviewers to synthesise evidence. Thus, Study 1 is of value for the research community.

While the effect of multidisciplinary rehabilitation in patients with CLBP is well established (7, 22-24), the results from Studies 2 and 3 provided evidence that the two ways of delivering rehabilitation programmes were equal in terms of effect on the chosen outcomes. These results could give rise to a discussion about whether patient preferences should guide the choice of rehabilitation programme. This of course has some implications for clinical practice. Firstly, are the services delivering the rehabilitation programmes able to deliver more than one rehabilitation programme? With this RCT, even though having a highly pragmatic attitude, we experienced some organisational difficulties in delivering more than one rehabilitation programme, at the same time, in the same setting, by the same providers. This implies that the delivery of flexible rehabilitation

programmes may be an enormous challenge in clinical practice requiring huge efforts from the services including the providers delivering the rehabilitation programmes. Many different rehabilitation programmes exist worldwide, but it is questionable if we know enough about what matters to the patients. This could possibly be explored further by qualitative work followed by patient preference clinical trials. However, we need to be aware that patient-preferences are not the only thing, as the context also plays a key role. For example, legislative, social and cultural structures are often different from country to country, and thus, difficult to standardise in clinical practice as well as in research.

In terms of CLBP being driven by biopsychosocial factors, and in light of the plethora of research in the field, it seems like a challenging task to target each individual patient with a primarily one-size-fits-all approach (33). This implies that future research should focus on how to subgroup patients with CLBP to receive more targeted treatment for their complex problems (7, 14, 114).

8.3 For public health

The management of CLBP is vital from a public health perspective both nationally and internationally in light of the burden experienced by the individuals affected, their relatives and society in general. A major weakness of the current RCT and many other RCTs in the same field (7, 44, 103, 104) is the lack of a cost-effectiveness analysis. It is imperative that future research includes cost-effectiveness analyses, as there is an obvious societal interest in achieving the best possible outcome with the least costly rehabilitation programmes.

CLBP should be understood as a condition of life. With the huge amount of available intervention research, new intervention studies are probably not the most optimal solution to stem the challenges facing these patients, their relatives and society. Maybe we need to gain a wider perspective and look into new actions required by the political, public health, and health care systems (21). Firstly, in terms of the political arena, national and international policy-makers ought to increase the public recognition of the effects and burden of low back pain. Secondly, according to the public health challenge, the onset and persistence of disability associated with LBP needs to be prevented. This can be done by changing priorities, systems and practices. Lastly, with respect to the health care system, a shift away from the emphasis on the biomedical and fragmented model of care is required. This implies a change in culture, clinical behaviour, and systems, as well as a need to tackle vested interests.

9. References

1. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* (London, England). 2016;388(10053):1545-602.
2. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. *Lancet* (London, England). 2018;391(10137):2356-67.
3. Vlaeyen JWS, Maher CG, Wiech K, Van Zundert J, Meloto CB, Diatchenko L, et al. Low back pain. *Nature reviews Disease primers*. 2018;4(1):52.
4. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ* (Clinical research ed). 2006;332(7555):1430-4.
5. Hoy D, March L, Brooks P, Blyth F, Woolf A, Bain C, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Annals of the rheumatic diseases*. 2014;73(6):968-74.
6. Waddell G. The biopsychosocial model. *The Back Pain Revolution*. Second edition ed: Churchill Livingstone; 2004. p. 265-82.
7. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RW, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review and meta-analysis. *BMJ* (Clinical research ed). 2015;350:h444.
8. O'Sullivan P. It's time for change with the management of non-specific chronic low back pain. *British journal of sports medicine*. 2012;46(4):224-7.
9. Waddell G. The epidemiology of back pain. *The Back Pain Revolution* Second edition ed: Churchill Livingstone; 2004. p. 71-90.
10. Waddell G. Risk factors for back pain. *The Back Pain Revolution*. Second edition ed: Churchill Livingstone; 2004. p. 91-114.
11. Engel GL. The need for a new medical model: a challenge for biomedicine. *Science* (New York, NY). 1977;196(4286):129-36.
12. Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. *Spine*. 1987;12(7):632-44.
13. Pincus T, Kent P, Bronfort G, Loisel P, Pransky G, Hartvigsen J. Twenty-five years with the biopsychosocial model of low back pain-is it time to celebrate? A report from the twelfth international forum for primary care research on low back pain. *Spine*. 2013;38(24):2118-23.
14. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. *Lancet* (London, England). 2018;391(10137):2368-83.
15. Wade D. Rehabilitation - a new approach. Part two: the underlying theories. *Clinical rehabilitation*. 2015;29(12):1145-54.
16. Organization WH. International classification of functioning, disability and health : ICF 2001. p. 1-26.
17. White Book on Physical and Rehabilitation Medicine (PRM) in Europe. Chapter 1. Definitions and concepts of PRM. *European journal of physical and rehabilitation medicine*. 2018;54(2):156-65.
18. Wade D. Rehabilitation--a new approach. Overview and Part One: the problems. *Clinical rehabilitation*. 2015;29(11):1041-50.
19. Wade D. Rehabilitation - a new approach. Part three: the implications of the theories. *Clinical rehabilitation*. 2016;30(1):3-10.
20. Wade D. Rehabilitation - a new approach. Part four: a new paradigm, and its implications. *Clinical rehabilitation*. 2016;30(2):109-18.
21. Buchbinder R, van Tulder M, Oberg B, Costa LM, Woolf A, Schoene M, et al. Low back pain: a call for action. *Lancet* (London, England). 2018;391(10137):2384-8.

22. Qaseem A, Wilt TJ, McLean RM, Forciea MA. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. *Annals of internal medicine*. 2017;166(7):514-30.
23. Excellence UNiFHaC. Low back pain and sciatica in over 26s: assessment and management. 2016 [Available from: www.nice.org.uk/guidance/ng59].
24. Sundhedsstyrelsen. Anbefalinger for tværsektorielle forløb for mennesker med kroniske lænderygsmærter København Sundhedsstyrelsen; 2017 [updated 08.06.17. 1,1:[Available from: <https://www.sst.dk/da/sygdom-og-behandling/kronisk-sygdom/faglige-anbefalinger/~media/D995B05669A44B319734DD8206975B16.ashx>].
25. Siegert RJ, Taylor WT, Dean SG. Introduction. In: Dean SG, Siegert RJ, Taylor WJ, editors. *Interprofessional rehabilitation : A person-centred approach*: Wiley-Blackwell; 2012. p. 1-8.
26. Kathryn McPherson e, Barbara EGe, Alain Leplège e. Rethinking rehabilitation : Theory, Practice, History - And the Future. In: McPherson K, Gibson BE, Leplège A, editors. *Rehabilitation Science in Practice Series*. Boca Raton, Florida: CRC Press; 2015. p. 3-20.
27. Bank WW. *World report on disability*. 2011.
28. Wade DT. Describing rehabilitation interventions. *Clinical rehabilitation*. 2005;19(8):811-8.
29. Levack W, Dean SG. Processes in rehabilitation. In: Dean SG, Siegert RJ, Taylor WJ, editors. *Interprofessional rehabilitation : A person-centred approach*: Wiley-Blackwell; 2012. p. 79-108.
30. Stucki G, Cieza A, Melvin J. The International Classification of Functioning, Disability and Health (ICF): a unifying model for the conceptual description of the rehabilitation strategy. *Journal of rehabilitation medicine*. 2007;39(4):279-85.
31. Wade DT, Smeets RJ, Verbunt JA. Research in rehabilitation medicine: methodological challenges. *Journal of clinical epidemiology*. 2010;63(7):699-704.
32. Council TMR. *Developing and evaluating complex interventions: new guidance 2008* [Available from: <https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/>].
33. Wade DT. What is rehabilitation? An empirical investigation leading to an evidence-based description. *Clinical rehabilitation*. 2020;269215520905112.
34. Dean SG, Ballinger C. An interprofessional approach to rehabilitation. In: Dean SG, Siegert RJ, Taylor WJ, editors. *Interprofessional rehabilitation : A person-centred approach*: Wiley-Blackwell; 2012. p. 45-70.
35. Pryor J, Dean SG. The person in context. In: Dean SG, Siegert RJ, Taylor WJ, editors. *Interprofessional rehabilitation : A person-centred approach*: Wiley-Blackwell; 2012. p. 135-66.
36. 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.
37. Mohler R, Kopke S, Meyer G. Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline (CReDECI 2). *Trials*. 2015;16:204.
38. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *BMJ (Clinical research ed)*. 2017;358:j3453.
39. Sermeus W. *Complex interventions in health - An overview of research methods*. First ed. New York: Routledge; 2015.
40. Giangregorio LMT, L. Pilot studies and feasibility studies for complex interventions - An introduction. In: Richards DAH, I. R., editor. *Complex interventions in health - An overview of research methods*. New York: Routledge; 2015. p. 127-35.
41. Moore G. Process evaluation of complex interventions - a summary of Medical Research Councils guidance. In: Richards DA, Hallberg IR, editors. *Complex Interventions in Health - An overview of research methods*. New York: Routledge; 2015. p. 222-31.

42. 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.
43. Grant A, Treweek S, Dreischulte T, Foy R, Guthrie B. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials*. 2013;14:15.
44. Linden M, Scherbe S, Cicholas B. Randomized controlled trial on the effectiveness of cognitive behavior group therapy in chronic back pain patients. *Journal of back and musculoskeletal rehabilitation*. 2014;27(4):563-8.
45. Tavafian SS, Jamshidi AR, Mohammad K. Treatment of low back pain: randomized clinical trial comparing a multidisciplinary group-based rehabilitation program with oral drug treatment up to 12 months. *International journal of rheumatic diseases*. 2014;17(2):159-64.
46. Abbasi M, Dehghani M, Keefe FJ, Jafari H, Behtash H, Shams J. Spouse-assisted training in pain coping skills and the outcome of multidisciplinary pain management for chronic low back pain treatment: a 1-year randomized controlled trial. *European journal of pain (London, England)*. 2012;16(7):1033-43.
47. Leeuw M, Goossens ME, van Breukelen GJ, de Jong JR, Heuts PH, Smeets RJ, et al. Exposure in vivo versus operant graded activity in chronic low back pain patients: results of a randomized controlled trial. *Pain*. 2008;138(1):192-207.
48. Mangels M, Schwarz S, Worringen U, Holme M, Rief W. Evaluation of a behavioral-medical inpatient rehabilitation treatment including booster sessions: a randomized controlled study. *The Clinical journal of pain*. 2009;25(5):356-64.
49. Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, Knottnerus JA. Chronic low back pain: physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. *Pain*. 2008;134(3):263-76.
50. Siegert RJ, McPherson KM, Dean SG. Theory development and a science of rehabilitation. *Disability and rehabilitation*. 2005;27(24):1493-501.
51. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *The Milbank quarterly*. 1996;74(4):511-44.
52. Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? *Effective clinical practice : ECP*. 1998;1(1):2-4.
53. Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: translating evidence into action. *Health affairs (Project Hope)*. 2001;20(6):64-78.
54. Taylor WJ, Geyh S. A rehabilitation framework: the International Classification of Functioning, Disability and Health. In: Dean SG, Siegert RJ, Taylor WJ, editors. *Interprofessional rehabilitation : A person-centred approach*: Wiley-Blackwell; 2012. p. 9-44.
55. Tousignant-Laflamme Y, Martel MO, Joshi AB, Cook CE. Rehabilitation management of low back pain - it's time to pull it all together! *Journal of pain research*. 2017;10:2373-85.
56. Siegert RJ, McPherson KM, Dean SG. Theory development and a science of rehabilitation: Authors' response to commentaries. *Disability and rehabilitation*. 2005;27(24):1517-9.
57. Schmidt AM, Terkildsen Maindal H, Laurberg TB, Schiottz-Christensen B, Ibsen C, Bak Gulstad K, et al. The Sano study: justification and detailed description of a multidisciplinary biopsychosocial rehabilitation programme in patients with chronic low back pain. *Clinical rehabilitation*. 2018;269215518780953.
58. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ (Clinical research ed)*. 2014;348:g1687.
59. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Process evaluation in randomised controlled trials of complex interventions. *BMJ (Clinical research ed)*. 2006;332(7538):413-6.
60. Craig P. Taking account of context in population health intervention research: guidance for producers, users and funders of research. Southampton2018.

61. Foundation WKK. Logic Model Development Guide Michigan2004 [Available from: <https://www.bttop.org/sites/default/files/public/W.K.%20Kellogg%20LogicModel.pdf>.
62. Köpke S. Exploring complexity in systematic reviews of complex interventions. In: Richards DA, Hallberg IR, editors. *Complex Interventions in Health - An overview of research methods*. New York: Routledge; 2015. p. 73-9.
63. Schmidt AM, Schiøttz-Christensen B, Foster NE, Laurberg TB, Maribo T. The effect of an integrated multidisciplinary rehabilitation programme alternating inpatient interventions with home-based activities for patients with chronic low back pain: a randomized controlled trial. *Clinical rehabilitation*. 2020;269215519897968.
64. INVOLVE. [Available from: <https://www.invo.org.uk/resource-centre/resource-for-researchers/>.
65. Richards DA. The critical importance of patient and public involvement for research into complex interventions. In: Richards DA, Hallberg IR, editors. *Complex interventions in health - An overview of research methods*. First ed. Oxon and New York: Routledge; 2015. p. 46-50.
66. de Wit MP, Berlo SE, Aanerud GJ, Aletaha D, Bijlsma JW, Croucher L, et al. European League Against Rheumatism recommendations for the inclusion of patient representatives in scientific projects. *Annals of the rheumatic diseases*. 2011;70(5):722-6.
67. Slade SC, Dionne CE, Underwood M, Buchbinder R, Beck B, Bennell K, et al. Consensus on Exercise Reporting Template (CERT): Modified Delphi Study. *Physical therapy*. 2016;96(10):1514-24.
68. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *Journal of clinical epidemiology*. 2010;63(8):834-40.
69. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS medicine*. 2008;5(1):e20.
70. 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.
71. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ (Clinical research ed)*. 2008;337:a2390.
72. Schmidt AM, Laurberg TB, Moll LT, Schiøttz-Christensen B, Maribo T. The effect of an integrated multidisciplinary rehabilitation programme for patients with chronic low back pain: long-term follow up of a randomised controlled trial. . Submitted to *Clinical Rehabilitation* 300320 2020.
73. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ (Clinical research ed)*. 2015;350:h2147.
74. Ford I, Norrie J. Pragmatic Trials. *The New England journal of medicine*. 2016;375(5):454-63.
75. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *Journal of clinical epidemiology*. 2009;62(5):499-505.
76. PRECIS-2. [Available from: <https://www.precis-2.org/>.
77. Anesthesiologists ASo. ASA Physical Status Classification System [updated 2014. Available from: <https://www.asahq.org/resources/clinical-information/asa-physical-status-classification-system>.
78. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine*. 2000;25(22):2940-52; discussion 52.
79. Hart T, Bagiella E. Design and implementation of clinical trials in rehabilitation research. *Archives of physical medicine and rehabilitation*. 2012;93(8 Suppl):S117-26.
80. Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LOP, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. *Pain*. 2018;159(3):481-95.
81. Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, et al. Core outcome domains for clinical trials in non-specific low back pain. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*. 2015;24(6):1127-42.

82. Chapman JR, Norvell DC, Hermsmeyer JT, Bransford RJ, DeVine J, McGirt MJ, et al. Evaluating common outcomes for measuring treatment success for chronic low back pain. *Spine*. 2011;36(21 Suppl):S54-68.
83. Clement RC, Welander A, Stowell C, Cha TD, Chen JL, Davies M, et al. A proposed set of metrics for standardized outcome reporting in the management of low back pain. *Acta orthopaedica*. 2015;86(5):523-33.
84. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Danish version of the Oswestry Disability Index for patients with low back pain. Part 1: Cross-cultural adaptation, reliability and validity in two different populations. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*. 2006;15(11):1705-16.
85. Rasmussen MU, Rydahl-Hansen S, Amris K, Samsoe BD, Mortensen EL. The adaptation of a Danish version of the Pain Self-Efficacy Questionnaire: reliability and construct validity in a population of patients with fibromyalgia in Denmark. *Scandinavian journal of caring sciences*. 2016;30(1):202-10.
86. Nicholas MK. The pain self-efficacy questionnaire: Taking pain into account. *European journal of pain (London, England)*. 2007;11(2):153-63.
87. Group TE. EQ-5D-5L [Available from: <https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/>].
88. Chiarotto A, Terwee CB, Kamper SJ, Boers M, Ostelo RW. Evidence on the measurement properties of health-related quality of life instruments is largely missing in patients with low back pain: A systematic review. *Journal of clinical epidemiology*. 2018;102:23-37.
89. Olsen LR, Jensen DV, Noerholm V, Martiny K, Bech P. The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. *Psychological medicine*. 2003;33(2):351-6.
90. Fysioterapeuter D. Hvordan spørger du ind til patienternes aktivitetsniveau? 2014 [updated 19-12-2014. Available from: <https://www.fysio.dk/nyheder/2014/Hvordan-sporger-du-ind-til-patienternes-aktivitetsniveau>].
91. Åstrand P-O. Textbook of work physiology: physiological bases of exercise. 4 ed 2003. p. 649.
92. Fysioterapeuter D. Et-punktstest (Aastrands test) [Available from: <https://fysio.dk/fafo/maleredskaber/et-punktstest/>].
93. Ratter J, Radlinger L, Lucas C. Several submaximal exercise tests are reliable, valid and acceptable in people with chronic pain, fibromyalgia or chronic fatigue: a systematic review. *Journal of physiotherapy*. 2014;60(3):144-50.
94. Meade TW, Dyer S, Browne W, Townsend J, Frank AO. Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. *BMJ (Clinical research ed)*. 1990;300(6737):1431-7.
95. Henchoz Y, de Goumoens P, So AK, Paillex R. Functional multidisciplinary rehabilitation versus outpatient physiotherapy for non specific low back pain: randomized controlled trial. *Swiss medical weekly*. 2010;140:w13133.
96. Perera R, Heneghan C, Yudkin P. Graphical method for depicting randomised trials of complex interventions. *BMJ (Clinical research ed)*. 2007;334(7585):127-9.
97. Pisters MF, Veenhof C, van Meeteren NL, Ostelo RW, de Bakker DH, Schellevis FG, et al. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review. *Arthritis and rheumatism*. 2007;57(7):1245-53.
98. Abbott JH, Chapple CM, Fitzgerald GK, Fritz JM, Childs JD, Harcombe H, et al. The Incremental Effects of Manual Therapy or Booster Sessions in Addition to Exercise Therapy for Knee Osteoarthritis: A Randomized Clinical Trial. *The Journal of orthopaedic and sports physical therapy*. 2015;45(12):975-83.
99. Fitzgerald GK, Fritz JM, Childs JD, Brennan GP, Talisa V, Gil AB, et al. Exercise, manual therapy, and use of booster sessions in physical therapy for knee osteoarthritis: a multi-center, factorial randomized clinical trial. *Osteoarthritis and cartilage*. 2016;24(8):1340-9.

100. Bennell KL, Kyriakides M, Hodges PW, Hinman RS. Effects of two physiotherapy booster sessions on outcomes with home exercise in people with knee osteoarthritis: a randomized controlled trial. *Arthritis care & research*. 2014;66(11):1680-7.
101. Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ (Clinical research ed)*. 2004;328(7455):1561-3.
102. Maindal HT, Sokolowski I, Vedsted P. Translation, adaptation and validation of the American short form Patient Activation Measure (PAM13) in a Danish version. *BMC public health*. 2009;9:209.
103. Iversen VM, Vasseljen O, Mork PJ, Gismervik S, Bertheussen GF, Salvesen O, et al. Resistance band training or general exercise in multidisciplinary rehabilitation of low back pain? A randomized trial. *Scandinavian journal of medicine & science in sports*. 2018;28(9):2074-83.
104. Verra ML, Angst F, Brioschi R, Lehmann S, Benz T, Aeschlimann A, et al. Effectiveness of subgroup-specific pain rehabilitation: a randomized controlled trial in patients with chronic back pain. *European journal of physical and rehabilitation medicine*. 2018;54(3):358-70.
105. Neta G, Johnson KE. Informing real-world practice with real-world evidence: the value of PRECIS-2. *BMC medicine*. 2018;16(1):76.
106. INVOLVE. Briefing note five: Offer training and support [Available from: <http://www.invo.org.uk/offer-training-and-support/>].
107. Brett J, Staniszewska S, Mockford C, Herron-Marx S, Hughes J, Tysall C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. *Health expectations : an international journal of public participation in health care and health policy*. 2014;17(5):637-50.
108. Lamb S, Altman DG. Individually and cluster-randomized trials. In: Richards DA, Hallberg IR, editors. *Complex interventions in health - An overview of research methods*. New York: Routledge; 2015. p. 191-9.
109. Fletcher RH. Treatment. *Clinical Epidemiology : The essentials 5th ed*: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2014.
110. Jadad AR. *Randomised controlled trials - A user's guide*: BMJ Books; 1998.
111. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Dore C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *Jama*. 2017;318(23):2337-43.
112. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ (Clinical research ed)*. 2001;322(7282):355-7.
113. Fletcher RH FS, Fletcher GS. Introduction. *Clinical Epidemiology - The essentials*. Fifth ed. Baltimore: Lippincott Williams & Wilkins; 2014. p. 1-16.
114. Foster NE. Barriers and progress in the treatment of low back pain. *BMC medicine*. 2011;9:108.

10. Appendices

Appendix 1

Study 1

Supplementary file: Appendix 8

Appendix 2

Study 2

Supplementary file 1: Appendix 5

Supplementary file 2: Figure 4 in this dissertation

Appendix 3

Study 3

Supplementary file 1: Appendix 7

Appendix 4

Permission of the UK Medical Research Council to reproduce Figure 1.

Appendix 5

A completed GRIPP2 short form

Appendix 6

Statistical analysis plan – Study 2

Appendix 7

Statistical analysis plan – Study 3

Appendix 8

Activity sheets describing each of the 38 clinical activities

Appendix 9

Graphical depiction comparing the integrated programme and the existing programme

Appendix 10

Examples process research question 1 and 2