

Anvendeligheten til fire måleinstrumenter for
bruk til evaluering av pasientopplæringskurs
for pasienter med inflammatorisk revmatisk
sykdom

- *En anvendelighetsstudie*

Renate Røe



Masteroppgave ved Institutt for helse og samfunn,
Seksjon for helsefag/ Det medisinske fakultet

UNIVERSITETET I OSLO

Dato: 24.11.17

Anvendeligheten til fire måleinstrumenter for bruk til evaluering av pasientopplæringskurs for pasienter med inflammatorisk revmatisk sykdom

- *En anvendelighetsstudie*

Renate Røe

Masteroppgave ved Institutt for helse og samfunn,
Seksjon for helsefag/ Det medisinske fakultet

UNIVERSITETET I OSLO

Dato: 24.11.17

© Forfatter Renate Røe

År 2017

Tittel

Forfatter: Renate Røe

<http://www.duo.uio.no>

Trykk: Representralen, Universitetet i Oslo

Sammendrag

Oppgavens formål: Å teste anvendeligheten til fire identifiserte måleinstrumenter for bruk til evaluering av pasientopplæringskurs.

Teoretisk forankring: Pasientopplæringskurs er en viktig del av den helhetlige behandlingen av pasienter med inflammatoriske revmatiske sykdommer. Kursene er ikke systematisk evaluert og det brukes flere ulike utfallsmål. EULAR anbefalinger for opplæringskurs til pasienter med inflammatorisk revmatisk sykdom understreker at eksisterende måleinstrumenter må evalueres. Et ekspertpanel med både brukerrepresentanter og helsepersonell har deltatt i tre Delphi-prosesser. Det ble identifisert syv aspekter som pasientopplæring anses å ha innflytelse på, og fire pasientrapporterte måleinstrumenter som fanger opp aspektene.

Metode: Pasienter med inflammatorisk revmatisk sykdom som deltok på kurs på seks revmatiske avdelinger fylte ut spørreskjema før og tre måneder etter kursdeltakelse. To sett av spørreskjemaer ble testet, sett A og sett B. Sett A: Arthritis Self-Efficacy Scale (ASES) for smerter og andre symptomer, Effective Consumer Scale (EC-17) og Patient Activation Measure (PAM). Sett B: Arthritis Self-Efficacy Scale (ASES) for smerter og andre symptomer og Health Education Questionnaire (heiQ, åtte separate undergrupper). Måleinstrumentene ble testet for manglende data, gulv- og takeffekter, intern konsistens (Cronbachs alfa) og før- og etterskår.

Resultater: 104 pasienter ble inkludert ved baseline, 63 i sett A og 41 i sett B. 78 (75%) fullførte spørreskjemaet tre måneder etter kursdeltakelse. Det var få manglende data i alle måleinstrumentene (0-2.2%). Alle måleinstrumentene hadde takeffekter i ett eller flere spørsmål ved baseline, ingen gulveffekter ble funnet i noen av instrumentene. Cronbachs alfa var akseptabel i ASES smerter (0.75) og ASES andre symptomer (0.81), EC-17 (0.91) og PAM (0.78). Seks av undergruppene i heiQ hadde akseptabel Cronbachs alfa (0.72-0.86). Statistisk signifikante endringer ble funnet i EC-17 og tre av undergruppene i heiQ.

Konklusjon: EC-17 og tre av undergruppene i heiQ var de mest anvendelige måleinstrumentene i forhold til gulv- og takeffekter, intern konsistens og før- og etterskår. ASES er det eneste måleinstrumentet som måler mestringstro og bør vurderes selv om det ikke oppnådde signifikante endringer etter tre måneder.

Abstract

Purpose: To test the feasibility of four identified Patient reported outcome measures (PROMs) to evaluate patient education programmes.

Literature framework: Patient education (PE) is an important part of the comprehensive treatment of patients with inflammatory rheumatic diseases. The programmes are not systematically evaluated and several different outcomes are used. EULAR recommendations for PE for people with inflammatory arthritis (IA) states that existing outcomes need to be evaluated. A Norwegian expert panel comprising patients and health professionals participated in a three-round Delphi-process. Seven domains that can be influenced by PE were identified and four patient-reported outcome measures that capture these outcome domains were identified by a systematic literature review and further Delphi-rounds.

Methods: Patients with IA who attended to PE programmes at six rheumatology departments completed questionnaires before and three months after participation. Two sets of questionnaires were tested. Set 1: Arthritis Self-Efficacy Scale (ASES) for pain and other symptoms, Effective Consumer Scale (EC-17) and Patient Activation Measure (PAM). Set 2: Arthritis Self-Efficacy Scale (ASES) for pain and symptoms and Health Education Questionnaire (heiQ, eight separate categories). The instruments were tested for missing values, internal consistency (Cronbach's alpha) and ability to detect change.

Results: 104 patients completed the questionnaires before intervention, 63 set A and 41 set B, respectively. 78 (75%) responded at three-month follow-up. Missing values were low in all PROMs (0-2.2%). All PROMs had ceiling effects in one or more questions at baseline. Cronbach's alpha was acceptable in ASES pain (0.75) and ASES other symptom (0.81), EC-17 (0.91) and PAM (0.78). Six heiQ categories had acceptable Cronbach's alpha (0.72-0.86). Statistically significant improvements were found in EC-17 and three heiQ categories after 3 months.

Conclusion: Based on this feasibility study in patients with IA, EC-17 and three heiQ categories were the most feasible PROMs in terms of ceiling effects, internal consistency and ability to detect change. ASES is the only instrument to measure self-efficacy and should be considered although it did not exhibit significant changes at three months.

Forord

Denne masteroppgaven har blitt til over tre år med intensivt arbeid, høy læringskurve og avgjørende veiledning og støtte. Jeg har alltid hatt et stort ønske om å bidra til forskning og utvikling av mitt fagfelt. Som sykepleier på Revmatologisk avdeling på Diakonhjemmet sykehus har jeg vært en del av et tverrfaglig arbeidsmiljø der klinisk arbeid og forskning er tett knyttet til hverandre. Gjennom mine kollegaer ved Revmatologisk avdeling fant jeg den faglige inspirasjonskilden og støtten jeg trengte for å realisere drømmen. I mars 2017 submitterte vi et abstrakt til EULAR-kongressen (European League Against Rheumatism). Abstraktet ble antatt og fremvist via posterpresentasjon på kongressen i Madrid i juni, se vedlegg 1.

Jeg er veldig takknemlig for muligheten jeg har fått til å gjennomføre masterforløpet gjennom økonomisk støtte fra Revmatologisk avdeling på Diakonhjemmet sykehus. Jeg vil rette en stor takk til min fantastiske veileder Heidi Andersen Zangi for uvurderlig støtte og veiledning i arbeidet med masteroppgaven. Din kunnskap og erfaring har vært av stor betydning, og ikke minst har din tilgjengelighet og positivitet hjulpet meg igjennom perioder hvor prosjektet har virket uoverkommelig. Du har også støttet og veiledet meg gjennom flere foredrag av resultatene og fremlegg av posterpresentasjonen, noe som har vært til stor betydning for resultatet av masteroppgaven. Jeg vil takke min enhetsleder Elisabeth Røssum-Haaland og fagsykepleier Torill Kristin Andreassen Karlsen for å ha lagt til rette forholdene og gjort det mulig for meg å gjennomføre utdanningen ved siden av min stilling som sykepleier. Jeg vil også takke min fagsykepleier Christin Lunner Olsen og resten av mine kollegaer for god støtte, faglig innspill og oppmuntring gjennom masterforløpet. Sist, men ikke minst, vil jeg takke min samboer Hans Gunnar Busland for hans tålmodighet og hans alltid like positive innstilling og støtte i tider hvor jeg selv har vært mindre positiv.

Oslo, november 2017

Renate Røe

Innholdsfortegnelse

Sammendrag	VI
Abstract	VII
Forord	IX
1 Introduksjon	1
1.1 Opgavens disposisjon	2
2 Formål og problemstilling	3
2.1 Formål	3
2.2 Problemstillinger	3
2.3 Utdyping og avgrensning av oppgaven	3
3 Teori	5
3.1 Inflammatoriske revmatiske sykdommer	5
3.2 Pasientopplæringskurs	5
3.3 Pasient- og pårørendeopplæring i revmatologi	7
3.4 Anvendelighetsstudier	10
4 Metode	11
4.1 Design	11
4.2 Deltakere	11
4.3 Spørreskjemaene	11
4.4 Datainnsamling	12
4.5 Variabler og målemetoder	13
4.5.1 Bakgrunnsvariabler	13
4.5.2 Arthritis Self-Efficacy Scale (ASES)	13
4.5.3 Effective Musculoskeletal Consumer scale (EC-17)	15
4.5.4 The Health Education Impact Questionnaire (heiQ)	16
4.5.5 Patient Activation Measure (PAM)	18
4.5.6 Spørreskjemaenes forståelighet og relevans	20
4.6 Pasientopplæringskursene	21
4.7 Statistiske analyser	23
4.7.1 Bakgrunnsvariabler	23
4.7.2 Manglende data	23
4.7.3 Gulv- og takeffekter	23
4.7.4 Intern konsistens	24
4.7.5 Instrumenters evne til å fange opp endring	25
4.7.6 Spørreskjemaenes forståelighet og relevans	25
4.8 Etsiske avveininger	25
5 Resultater	27
5.1 Flytdiagram over respondenter	27
5.2 Bakgrunnsvariabler	28
5.3 Manglende data	29
5.4 Gulv- og takeffekter	29
5.5 Intern konsistens	31
5.6 Endringer i skår etter 3 måneder	32
5.7 Spørreskjemaenes forståelighet og relevans til kursene	35
6 Presentasjon og diskusjon av funn	36
6.1 Diskusjon av sentrale funn	36

6.1.1	Manglende data.....	36
6.1.2	Gulv- og takeffekter.....	37
6.1.3	Intern konsistens	39
6.1.4	Skår før og etter gjennomført kurs.....	40
6.2	Metodiske styrker og svakheter	44
6.3	Teoretiske implikasjoner av funn	45
7	Konklusjon og anbefaling av instrumentvalg.....	46
	Litteraturliste	47
	Vedlegg	53

1 Introduksjon

Jeg har arbeidet som sykepleier på Revmatologisk avdeling på Diakonhjemmet sykehus i fire år. På denne avdelingen undersøkes, diagnostiseres og behandles pasienter med inflammatoriske revmatiske sykdommer. I mitt arbeid møter jeg pasienter på ulike stadier i sykdomsforløpet. En av flere fellesnevnerne for de aller fleste pasienter, uavhengig av sykdomsvarighet, er informasjonsbehovet. Inflammatoriske revmatiske sykdommer er kroniske sykdommer. Nye legemidler og fokus på tidlig innsettelse av effektiv behandling de siste 15 årene har gjort det mulig å dempe sykdomsaktiviteten og redusere symptomene. Likevel lever fortsatt mange pasienter med smerter, tretthet, redusert fysisk funksjon og livskvalitet. Europeiske kunnskapsbaserte anbefalinger fremhever derfor informasjon og undervisning som et viktig supplement til medikamentell behandling (1-3).

På revmatologisk avdeling på Diakonhjemmet er pasientopplæringskurs en del av behandlingsforløpet og tilbys alle nydiagnostiserte pasienter. Pasienter som har hatt sykdommen over lengre tid tilbys også kurs etter ønske og behov. Alle revmatologiske avdelinger i Norge tilbyr systematisk opplæring i form av ulike kurs for pasienter med revmatiske sykdommer. Det er imidlertid stor variasjon i hvem som får tilbud om kurs og hvordan disse struktureres i form av lengde, innhold og evaluering. Kursene blir ikke evaluert systematisk. Oversiktsstudier viser at det brukes flere ulike utfallsmål i studier gjort av pasientopplæringskurs for pasienter med inflammatoriske revmatiske sykdommer. Dette vanskeliggjør sammenlikning av resultater på tvers av kursene (4-6).

På veien mot en mer systematisk evaluering av pasientopplæringen er kvalitetsutviklingsprosjektet Pasient- og pårørende-opplæring i revmatologi (POPOPP revma) opprettet ved Nasjonal kompetansetjeneste for revmatologisk rehabilitering (NKRR) på Diakonhjemmet sykehus. I dette prosjektet er det identifisert flere aktuelle måleinstrumenter som kan benyttes i en systematisk evaluering av pasientopplæringskurs ved revmatologiske avdelinger i Norge (7). Før en systematisk evaluering utføres, er det ønskelig å gjennomføre

en anvendelighetsstudie for å vurdere hvilket eller hvilke av måleinstrumentene som skal benyttes.

For meg har det siden oppstart av mastergradsutdanningen vært ønskelig at min masteroppgave skal komme min avdeling og pasientgruppe til gode. Derfor, når jeg fikk tilbudet om å gjennomføre anvendelighetsstudien og dermed bidra til utviklingen av en mer systematisk evaluering av pasientopplæringskurs, grep jeg anledningen med begge hender.

1.1 Oppgavens disposisjon

Oppgaven består av seks kapitler og starter med introduksjonen ovenfor. Deretter kommer formål, problemstilling og utdyping og avgrensning av oppgaven. Her blir oppgavens fokusområde klargjort. Videre kommer teorikapittelet som knytter oppgaven til en teoretisk og empirisk ramme. Det gir en dypere forståelse av oppgavens formål og problemstilling. For å tydeliggjøre både oppgavens betydning og begrensninger, er det inkludert teori og definisjon av anvendelighetsstudier. Metodekapittelet redegjør for valg av metode og beslutninger tatt i forskningsforløpet. Måleinstrumentenes validitet og reliabilitet blir gjennomgått og valg av analyser forklart. Teori av begreper som benyttes i problemstillingen er lagt til metodekapittelet da begrepene er tett knyttet til valg av statistiske analyser. Metodekapittelet avrundes med en etisk avveining av metodene og valgene i forskningsforløpet. Etter metode følger resultatkapittelet hvor resultater av analysene blir beskrevet. Resultatene forklares i både tekst, tabell og figurer for å gi leseren en visuell fremstilling av funnene. Deretter kommer diskusjon av resultatene. Først diskuteres sentrale funn og intern validitet. Videre blir metodiske styrker og svakheter og teoretiske implikasjoner av funn belyst. Oppgaven avrundes med en konklusjon der funnene blir summert og anbefaling av instrumentvalg til en hovedstudie blir fremlagt.

2 Formål og problemstilling

2.1 Formål

Denne masteroppgaven er knyttet til kvalitetsutviklingsprosjektet POPOPP revma og er en anvendelighetsstudie av måleinstrumenter identifisert i dette prosjektet. Formålet med studien er å vurdere anvendeligheten til de fire identifiserte måleinstrumentene Arthritis Self-Efficacy Scale, Effective Consumer Scale, Health Education Questionnaire og Patient Activation Measure med tanke på gulv- og takeffekter, manglende data, intern konsistens og skåring før og etter kurs.

2.2 Problemstillinger

- I hvilken grad har de fire måleinstrumentene manglende data?
- I hvilken grad har de fire måleinstrumentene gulv- og takeffekter ved baseline?
- Hvordan er den interne konsistensen i hvert enkelt måleinstrument?
- Hvordan er før- og etterskår for de fire instrumentene?
Hvor stor er endringen for de forskjellige instrumentene?

2.3 Utdyping og avgrensning av oppgaven

Problemstillingene undersøker anvendeligheten fra en forskers ståsted. Pasientopplæringskurs skal bygge på både forskningsbasert kunnskap og pasienters erfaringskunnskap (8), og dette kravet bør også overholdes ved valg av måleinstrumenter til å evaluere kursene med. Problemstillingene ovenfor vil derfor bli sett i lys av respondentenes egne meninger om måleinstrumentene ved at spørsmål om måleinstrumentenes forståelighet og relevans til kurset er inkludert i enden av hvert spørreskjema de har mottatt. Kvalitetsprosjektet POPOPP revma omhandler pasientopplæring for både pasienter og

pårørende. Fordi dette er en masteroppgave er det behov for å avgrense omfanget av oppgaven. Vi har derfor valgt å fokusere kun på pasientene.

3 Teori

3.1 Inflammatoriske revmatiske sykdommer

Inflammatoriske revmatiske sykdommer er autoimmune sykdommer dominert av inflammasjon i ledd og bindevev. Verdens helseorganisasjon (WHO) klassifiserer revmatiske sykdommer i fire hovedgrupper hvorav inflammatoriske revmatiske sykdommer utgjør en av hovedgruppene. Revmatoid artritt, primær sjøgrens syndrom, ankyloserende spondylitt og psoriasisartritt er de vanligste inflammatoriske revmatiske sykdommene (9).

Inflammatoriske revmatiske sykdommer domineres av inflammasjon i ledd og flere andre organer (9). Ubehandlet kan sykdommene gi feilstillinger i ledd, leddskader og funksjonsbegrensninger. Pasientene behandles medikamentelt med både syntetiske og biologiske sykdomsmodifiserende antirevmatiske legemidler (sDMARDs og bDMARDs), ikke-steroid antiinflammatoriske legemidler (NSAIDs), kortikosteroider og ren smertestillende medikamenter. Målet med den medikamentelle behandlingen er å oppnå remisjon eller lav sykdomsaktivitet. Det finnes foreløpig ingen kurativ behandling, og mange pasienter opplever fortsatt smerter, tretthet, redusert fysisk funksjon og livskvalitet. Informasjon og undervisning fremheves som et viktig supplement til den medikamentelle behandlingen i europeiske kunnskapsbaserte anbefalinger om behandling av inflammatoriske revmatiske sykdommer (1-3). Opplæring av kronisk syke er også viktig i et helseøkonomisk perspektiv. Man regner med at minst 300 000 personer har en revmatisk sykdom i Norge. Muskel- og skjelettlidelser er en av de hyppigste årsakene til at folk oppsøker helsetjenester. I Norge er disse sykdommene også dominerende årsaker til sykefravær, rehabilitering og uførepensjonering, og revmatiske sykdommer alene svarte for 1% av sykepengetilfellene i 2008 (10).

3.2 Pasientopplæringskurs

I mange år har vestlig medisin hatt hovedfokus på biologi og patofysiologi, både når det gjelder akutt og langvarig sykdom. Tilsvarende har pasientrollen vært passiv og autoritetstro. Når pasientrollen for flere tiår tilbake begynte å endre seg til å bli langt mer aktiv og medvirkende, endret også behovene seg. Pasienters økende medinnflytelse over egen behandling fordret at

de hadde kompetanse til å gjøre egne valg. Dette førte til et internasjonalt fokus på utvikling av pasientopplæringskurs. I forbindelse med innføringen av lov om spesialisthelsetjenester i 2001 ble norske sykehus lovpålagt å drive pasient- og pårørendeopplæring (11).

Lov om Spesialisthelsetjenester definerer pasient- og pårørendeopplæring som en av spesialisthelsetjenestens fire hovedoppgaver, på lik linje med pasientbehandling, utdanning av helsepersonell og forskning (11). Pasientopplæring er innlemmet i loven spesielt med tanke på behovet for, og betydningen av, opplæringstilbud til pasienter med kroniske sykdommer og nedsatt funksjonsevne (12). Omlag en fjerdedel av den norske befolkningen har i dag en alvorlig kronisk sykdom og må daglig treffe beslutninger om egen tilstand og behandling. I takt med økende levealder i befolkningen vil dette antallet kunne øke, noe som understreker betydningen av opplæring (8, 13). I forarbeidene til loven understrekes det at god opplæring og mestring av sykdommen er like viktig som god behandling, og at mangel på opplæring er like alvorlig som mangel på medisiner og gode behandlingsformer (8).

Pasientopplæring er definert som planlagte opplæringsaktiviteter som igangsettes for å bedre pasienters helseatferd og helsestatus (14). Opplæringen skal styrke pasientens evne til å håndtere hverdagen med sykdommen, gjennomføre behandlingen, forebygge komplikasjoner og opprettholde helse og livskvalitet. Opplæringstilbudet til kronisk syke skal baseres på tverrfaglig kunnskap og innsats av helsepersonell fra flere spesialiteter i tillegg til pasienter og/eller pasientorganisasjoner (8). Det er også av betydning at pasientopplæringen er kunnskapsbasert. Kunnskapsbasert praksis er å ta faglige avgjørelser basert på systematisk innhentet forskningsbasert kunnskap, profesjonell og klinisk erfaringsbasert kunnskap og pasienters og pårørendes erfaringskunnskap. Helse Sør-Øst har de siste årene hatt en regional strategi for pasient- og pårørendeopplæring med mål om et mer kunnskapsbasert opplæringstilbud, blant annet gjennom å styrke forskningsaktiviteten på fagområdet (15).

Pasientopplæringskurs for pasienter med inflammatoriske revmatiske sykdommer er innlemmet i et helhetlig behandlingsopplegg. Opplæringen foregår ved de enkelte sykehusavdelingene, lærings- og mestringssentra (LMS) tilknyttet sykehusene og på rehabiliteringssentre. Det finnes ingen nasjonale retningslinjer for hvordan pasientopplæringen skal gjennomføres. Det er publisert systematiske oversiktsstudier som

har vist at opplæringsprogram for pasienter med inflammatoriske revmatiske sykdommer gir små, men signifikante forbedringer på oppfølging av behandling, mestringsstrategier og fysisk og psykisk helse (4-6). Pasienter rapporterer også erfaringsutveksling i møte med andre personer som betydningsfullt (16, 17). Til tross for flere publiserte studier om opplæringsprogram også generelt, er forskningsfeltet uoversiktlig og ofte med dårlig beskrevne tiltak (18). I systematiske oversikter påpekes det at flere studier av pasientopplæring for pasienter med revmatoid artritt har lav kvalitet, mangelfull eller uklar teoretisk forankring og ulike utfallsmål. Dette vanskeliggjør evaluering og sammenlikning av resultater av ulike pasientopplæringskurs (4-6). Det er gjennomført få studier av pasientopplæring innenfor revmatologi i Norge og ingen norske studier er representert i de systematiske oversiktene. For å oppnå en mer kunnskapsbasert opplæring anbefaler The European League Against Rheumatism (EULAR) at det gjennomføres systematiske evalueringer med standardiserte måleinstrumenter (2).

3.3 Pasient- og pårørendeopplæring i revmatologi

Pasient- og pårørendeopplæring i revmatologi (POPOPP revma) er et kvalitetsutviklingsprosjekt som ble igangsatt ved Nasjonal kompetansetjeneste for revmatologisk rehabilitering (NKRR) på Diakonhjemmet sykehus i 2014. Prosjektet ble igangsatt som en respons på behovet for kvalitetssikring av opplæringstilbudet og for å sørge for at pasienter og pårørende over hele landet får et likeverdig tilbud. Prosjektet ble ledet av forsker Heidi A. Zangi og har bestått av en ekspertgruppe på til sammen 28 brukerrepresentanter og tverrfaglig helsepersonell fra ulike helseregioner. Alle i ekspertgruppen har erfaring fra pasientopplæring innen revmatologi.

Prosjektets hovedmål var å identifisere måleinstrumenter (spørreskjemaer) som kan benyttes i en systematisk evaluering av pasient- og pårørendekurs ved revmatologiske avdelinger i Norge, samt å gjennomføre en anvendelighetsstudie med de identifiserte måleinstrumentene (7). Prosessen hadde tre delmål:

- (I) Gjennomføre en Delphi-prosess for å oppnå konsensus blant brukere (pasienter) og helsepersonell om åtte sentrale aspekter som opplæringen kan ha innflytelse på

- (II) Identifisere relevante måleinstrumenter (spørreskjemaer) som er egnet til å måle endringer i de sentrale aspektene som er identifisert i delmål I
- (III) Gjennomføre evalueringen (anvendelighetsstudie) med måleinstrumentene som ble identifisert i delmål II (7)

I delmål I ble sentrale aspekter identifisert gjennom en Delphi-prosess. En Delphi-prosess er en fellesbetegnelse på konsensusmetoder der man søker å oppnå enighet innenfor et valgt tema (19). Prosessen foregår igjennom flere runder der ekspertgruppen først kommer med forslag til aspekter. Aspektene blir kategorisert og de med samme meningsinnhold slått sammen. De redigerte aspektene sendes ut til ekspertgruppen som rangerer de ulike aspektene etter forhåndsbestemte kriterier. Svarene blir analysert og på ny utsendt til ekspertgruppen med opplysninger om hvilken oppslutning aspektene fikk ved forrige runde. Ekspertgruppen rangerer igjen aspektene, og denne prosessen gjentas til det er oppnådd konsensus (19).

Via delmål I ble det identifisert åtte sentrale aspekter som opplæringen anses å ha innflytelse på. To av aspektene, "bevissthet om egne styrker" og "ta styring over livet", ble sammenslått fordi det var vanskelig å skille disse aspektene når det kom til måleinstrumentene. De endelige aspektene ble:

1. Bevissthet om egne styrker
2. Forståelse for sykdom og behandling
3. Kunnskap om helsefremmende livsstil
4. Kommunikasjon mellom pasient og helsepersonell
5. Mestringsstrategier
6. Mestringstro
7. Dele erfaringer

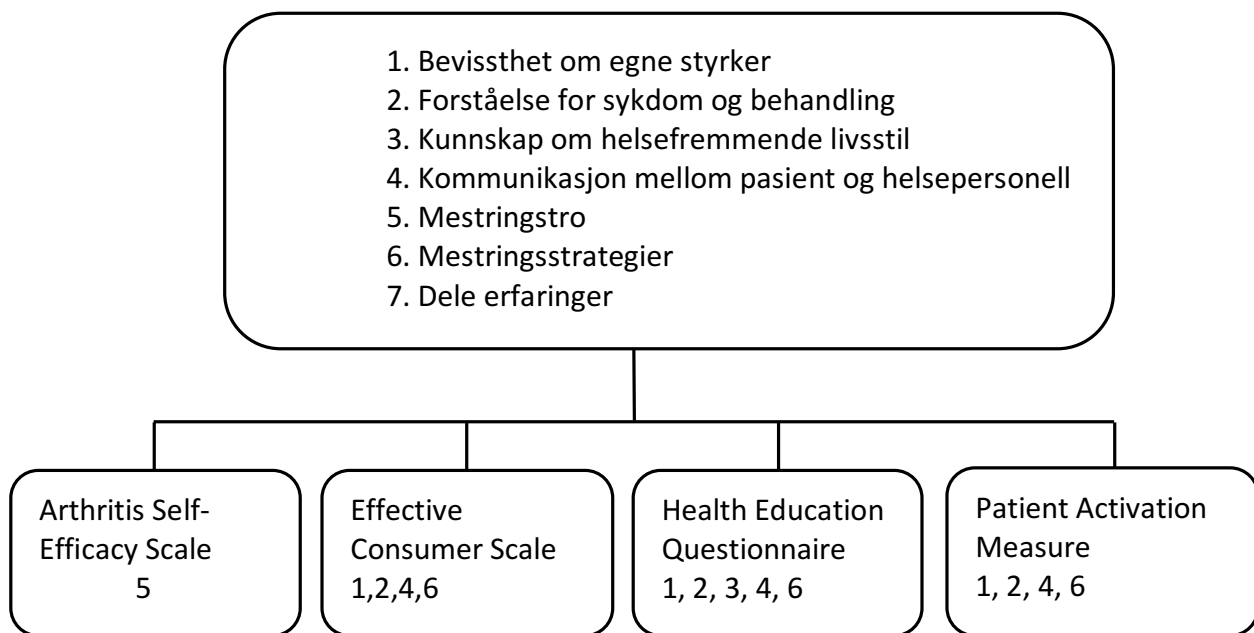
I delmål II ble aspektene benyttet til å identifisere aktuelle måleinstrumenter gjennom et systematisk litteratursøk. Et hovedkriterium var at måleinstrumentene måtte være validerte og oversatt til norsk. Via litteratursøket ble det funnet 21 måleinstrumenter. Av disse ble 13 instrumenter ekskludert fordi de ikke var oversatt til norsk, og ni instrumenter ble tatt med

inn i en ny Delphiprosess i ekspertpanelet (vedlegg 2). I denne prosessen ble fire av måleinstrumentene utvalgt som de mest aktuelle:

1. Arthritis Self-Efficacy Scale (ASES)
2. Effective Consumer Scale (EC-17)
3. Health Education Questionnaire (heiQ)
4. Patient Activation Measure (PAM)

De fire måleinstrumentene fordeler seg ulikt over seks av de syv aspektene, se figur 1. Ingen av instrumentene dekker aspektet "dele erfaringer".

Figur 1: Måleinstrumentenes fordeling over de syv ulike aspektene:



Det er ønskelig at det i delmål III, som et resultat av denne anvendelighetsstudien, fremkommer en anbefaling om hvilket eller hvilke av de fire måleinstrumentene som anses som mest anvendelige for evaluering av pasientopplæringskurs for pasienter med inflammatoriske revmatiske sykdommer. Det er også ønskelig at det fremkommer en oversikt over hvilke måleinstrumenter som til sammen dekker flest mulig av de seks aspektene.

3.4 Anvendelighetsstudier

Anvendelighetsstudier er forstudier til en hovedstudie for å estimere viktige parametere som er nødvendig for å designe hovedstudien (20). Det primære formålet er å sikre at hovedstudien er praktisk gjennomførbar og redusere trusler mot gyldigheten av studiens resultater (21). Eksempler på dette er tid som trengs for å samle inn og analysere data, undersøke egenskaper ved foreslåtte utfallsmål og oppfølgingsfrekvenser og svarfrekvenser på spørreskjemaer (20). Hypotesen(e) av interesse i hovedstudien testes ikke. Det har lenge vært uklare grenser mellom anvendelighetsstudier og pilotstudier, og det har de senere årene vært økende fokus på å definere og skille disse fra hverandre. Det nasjonale instituttet for helseforskning (NIHR) i Storbritannia har utarbeidet en definisjon som nå er bredt anerkjent. I følge NIHR er en anvendelighetsstudie forskjellig fra en pilotstudie ved at en anvendelighetsstudie prøver ut deler av en hovedstudie, mens en pilotstudie er en miniatyrutgave av hovedstudien som tester om alle komponentene i hovedstudien kan fungere sammen (22). Anvendelighetsstudier trenger ikke være randomiserte og forventes ikke å ha den utvalgsstørrelsen som trengs for finne statistisk signifikante endringer (20). Resultatene av de fleste studiene måles med beskrivende statistikk og kvalitativ analyse (21).

4 Metode

4.1 Design

Dette er en anvendelighetsstudie med et prospektivt før-/etter-design i form av en spørreundersøkelse, der anvendeligheten til ASES, EC-17, heiQ og PAM undersøkes.

4.2 Deltakere

Inklusjonskriteriene for deltakelse i studien var at pasienten måtte være over 18 år og diagnostisert med en inflammatorisk revmatisk sykdom. Pasientene måtte delta i et pasientopplæringskurs for pasienter med inflammatorisk revmatisk sykdom. Pasienter ble ekskludert fra studien hvis de ikke var i stand til å lese norsk på et nivå hvor de kunne forstå spørreskjemaene.

4.3 Spørreskjemaene

De fire måleinstrumentene ASES, EC-17, heiQ og PAM er pasientrapportert spørreskjemaer. Det ble laget spørreskjemaer i papirversjon som skulle besvares av pasienter før og tre måneder etter pasientopplæringskurs. For å unngå et uhandterlig langt spørreskjema ble de fire måleinstrumentene fordelt på to ulike sett; spørreskjema A (sA) og spørreskjema B (sB). ASES er det eneste instrumentet som måler mestringstro. Det ble derfor inkludert i begge spørreskjemaene. sA inneholdt ASES, EC-17 og PAM, mens sB inneholdt ASES og heiQ. Begge spørreskjemaene bestod i tillegg av demografiske variabler som kjønn, alder, utdanning og livsstil (vedlegg 3-8). Siste side i både sA og sB utlevert før kurs bestod av et spørsmål om hvor forståelige respondentene syntes spørsmålene i spørreskjemaet var. Siste siden i spørreskjemaene tre måneder etter kurs bestod av et spørsmål om hvor relevante respondentene syntes spørsmålene var med tanke på kurset de hadde deltatt på.

4.4 Datainnsamling

For å få en oversikt over antall tilgjengelige opplæringskurs sendte vi en mail til alle revmatologiske avdelinger i Norge. Vi informerte om studien og etterspurte aktuelle kurs i tidsrommet mars-juli 2016 og muligheter for bistand med inklusjon av pasienter. Av 13 forespurte avdelinger var det seks som både skulle gjennomføre aktuelle kurs og som kunne bidra med inkludering av pasienter til studien. Dette var revmatologisk avdeling ved Diakonhjemmet sykehus, Nordlandssykehuset, Revmatismesykehuset Lillehammer, Sykehuset Levanger, Universitetssykehuset i Nord-Norge og Ålesund sykehus. Avtalte kontaktpersoner ved avdelingene fikk tilsendt skriftlig informasjon om studien, anonymiserte spørreskjemaer, id-nøkkel, samtykkeerklæring og ferdigfrankerte svarconvolutter i forkant av hvert kurs. Det ble utlevert sA til kurs som skulle gjennomføres i mars og halve april, mens sB ble utlevert til kursene i siste halvdel av april og ut juli. Aktuelle pasienter fikk skriftlig og muntlig informasjon om studien ved oppmøte første kursdag og ble forespurt studiedeltakelse av kursleder. Pasienter som samtykket til deltakelse fikk utlevert et spørreskjema som skulle besvares før kurset startet. Ved Ålesund og Diakonhjemmet sykehus, fikk kurspåmeldte pasienter informasjonsskriv, samtykkeerklæring og første spørreskjema tilsendt i posten 1-2 uker før kursoppstart (vedlegg 9 og 10). Besvarte spørreskjemaer ble innsamlet av kursleder og postet til prosjektleder Heidi A. Zangi på NKRR. Spørreskjemaer og id-nøkkel ble sendt i separate convolutter. Pasientene besvarte spørreskjema før deltakelse på pasientopplæringskurs og tre måneder etter endt kurs. Spørreskjema tre måneder etter endt kurs ble tilsendt pasientene via post. De besvarte skjemaene ble lagt i de ferdigfrankerte svarconvoluttene og returnert til prosjektlederen. Påminnelsesbrev med et nytt spørreskjema ble tilsendt pasienter som ikke responderte innen to uker. Mottatte spørreskjemaer og id-nøkler ble lagt i separate låsbare skap.

4.5 Variabler og målemetoder

For å beskrivelse måleinstrumentenes psykometriske egenskaper og målemetoder har vi benyttet valideringsstudier av det originale måleinstrumentet samt oversatte versjoner fra europeiske land.

4.5.1 Bakgrunnsvariabler

Spørreskjemaet inneholder spørsmål om alder, kjønn, sivilstatus, arbeidsstatus og livsstil (mosjon, røyk og snus).

4.5.2 Arthritis Self-Efficacy Scale (ASES)

ASES ble utviklet i USA i 1989 og var det første artritt-spesifikke måleinstrument av self-efficacy (norsk: mestringstro eller mestringsforventning). Det ble laget for å måle pasienters tro på at de kunne utføre bestemte aktiviteter eller handlinger for å oppnå et resultat og takle konsekvenser av inflammatoriske revmatiske sykdommer (23). ASES er et pasientrapportert spørreskjema hvor formålet er å evaluere utfall av pasientopplæringsprogram og rehabilitering (23). Spørsmålene ble utviklet gjennom fokusgrupper med pasienter med revmatoid artritt. Validering av spørsmål og undergrupper foregikk gjennom to ulike studier (23).

Den originale ASES består av 20 spørsmål fordelt på tre selvstendige undergrupper som måler self-efficacy relatert til fysisk funksjon (9 spørsmål), smerter (5 spørsmål) og andre symptomer (6 spørsmål). Eksempel på spørsmål er; "hvor sikker er du på at du kan minske smerten ganske mye?" (fra smerter) og "hvor sikker er du på at du kan takle giktsymptomene dine slik at du kan gjøre de tingene du liker å gjøre?" (fra andre symptomer). Det originale spørreskjemaet besvares med en 10-punkts (10-100) ende- og midtpunktsforankret skala; (10) meget usikker, (50) nokså sikker og (100) meget sikker. ASES finnes i flere ulike versjoner der enten tre, to eller bare en av undergruppene benyttes. Det er også utviklet versjoner spesielt rettet mot pasienter med fibromyalgi og ankyloserende spondylitt. De ulike versjonene har forskjellige skåringsskalaer. Den originale skåringsskalaen til ASES har blitt kritisert for å være kompleks, og det er derfor utviklet en ny skåringsskala med en fem-punkts skala fra 1-5 i den norske versjonen av ASES. Dette er en skåringsskala som er vanlig å benytte i pasientrapporterte

spørreskjemaer (24, 25). Skalaen har fem svaralternativer; (1) meget usikker, (2) litt usikker, (3) verken sikker eller usikker, (4) litt sikker og (5) meget sikker. Denne versjonen er nylig validert i Norge (25). Ved utregning av resultater lages en sumskår av alle spørsmålene i hver enkelt skala der høyere skår representerer høyere nivå av mestringstro. Hvis mer enn 25% av spørsmålene mangler, skal ikke undergruppen skåres. I denne studien benyttes ASES-11 som inneholder undergruppene smerter og andre symptomer med skåringsskalaen fra 1-5.

Flere av de ulike utgavene av ASES er validert, men her vises det til resultater fra evalueringsstudier av den originale ASES og ASES-11. Disse to utgavene av ASES gir oversikt over de psykometriske egenskapene til undergruppene smerter og andre symptomer som er av interesse for denne studien. Originalversjonen av ASES med skåringsskala fra 10-100 er evaluert i USA og Sverige, mens ASES-11 er evaluert i England og Norge (23, 25-28). I den engelske studien benyttes en 10-punkts endepunksforankret skala fra 1-10 med ankrene "veldig usikker" til "veldig sikker" i hver sin ende, mens den norske studien benytter skåringsskalaen fra 1-5 (25, 27). Studienes utvalg består av pasienter med inflammatoriske revmatiske sykdommer og artrose, og et av studiene har også inkludert pasienter med fibromyalgi. Den strukturelle validiteten med henholdsvis tre- og todelt struktur bekreftes i samtlige studier. Den interne konsistensen målt med Cronbachs alfa er mellom 0.75-0.94 for undergruppene i alle studiene. Test-retest målt med Pearsons korrelasjonskoeffisient er rapportert i to av studiene og ligger over 0.85 i alle tre skalaene (27, 28). Manglende data og gulv- og takeffekter rapporteres kun i en studie (25). Studien rapporterer manglende data under 5%, ingen gulveffekter, et spørsmål med takeffekt i ASES smerter (24.8%) og to spørsmål i ASES andre symptomer (16.3% og 19.4%) (25). Ved summering av alle spørsmålene, finnes ingen gulv- eller takeffekter. Responsiviteten til ASES med skåringsskalaen fra 1-5 er foreløpig ikke undersøkt.

4.5.3 Effective Musculoskeletal Consumer scale (EC-17)

EC-17 er et pasientrapportert spørreskjema utviklet på oppdrag fra «The OMERACT (Outcome Measures in Rheumatology) Effective Musculoskeletal Consumer Project». Prosjektets formål var å imøtekomme behovet for et passende utfallsmål for evaluering og sammenlikning av pasientopplæringskurs for pasienter med muskel- og skjelettlidelser (29). Spørreskjemaet ble utviklet gjennom litteraturgjennomgang, intervjuer med pasienter og helsepersonell, tilbakemeldinger fra OMERACT-medlemmer og pretest med pasienter.

EC-17 inneholder 17 påstander om ferdigheter, holdninger og kunnskap som er viktig for effektivt å håndtere, delta i og ta styring over egne helsetjenester. To eksempler er "jeg forstår informasjonen jeg får om min sykdom" og "jeg kan sette realistiske mål for mestring av sykdommen min". For å gjøre spørreskjemaet oversiktlig og enkelt å besvare, er spørsmålene fordelt over fem underoverskrifter (30);

1. hvordan jeg bruker helseinformasjon
2. hvordan jeg avklarer og avveier verdier og prioriteringer
3. hvordan jeg kommuniserer med andre
4. hvordan jeg forhandler om roller og tar kontroll
5. hvordan jeg tar beslutninger og handler

Hver underoverskrift inneholder 3-4 spørsmål og besvares på en Likert-skala fra 0-4; (0) aldri, (1) sjelden, (2) noen ganger, (3) vanligvis og (4) alltid. Desto høyere skår, desto mer selvstendig og effektiv er pasienten i håndtering av egen helse og helsetjenester. Selv om EC-17 ble utviklet med tanke på en spesifikk pasientgruppe, er spørreskjemaet generisk. Spørsmålene er ikke knyttet opp mot en spesifikk diagnose, men ordet "sykdom" brukes i flere av spørsmålene. EC-17 er endimensjonal og gir en samlet skår. Ved skåring av spørreskjema omregnes råskårene til en skala fra 0-100 hvor 0=0, 1=25, 2=50, 3=75 og 4=100. Deretter beregnes en gjennomsnittsskår. Ved mer enn ett ubesvart spørsmål i en av underoverskriftene, eller mer enn tre ubesvarte spørsmål i hele spørreskjemaet, kan ikke skjemaet skåres (31).

Den engelske originalversjonen av EC-17 er oversatt og validert i Norge, Sverige og Nederland. Studienes utvalg består av pasienter med ulike muskel- og skjelettsykdommer. Utvalget i den

svenske studien består utelukkende av pasienter med ulike inflammatoriske revmatiske sykdommer, mens den engelske og norske studien i tillegg har inkludert pasienter med fibromyalgi og artrose. Utvalget i den nederlandske studien består utelukkende av pasienter med fibromyalgi og artrose. Den endimensjonale strukturen er bekreftet i flere studier (29, 30, 32, 33). I alle studiene har EC-17 en Cronbachs alfa over 0.90. I test-retest er Intraclass korrelasjonskoeffisienten (ICC) på >0.9 for to av studiene og 0.71 for den nederlandske studien (30, 32, 33). Alle valideringsstudiene rapporterer få manglende data, men faktisk antall manglende data for hver av de 17 spørsmålene er kun oppgitt i to av studiene og ligger på under 5% (29, 32). Det er ikke funnet gulv- og takeffekter i noen av studiene. Før- og etterstudier har funnet evidens for responsivitet, men det anbefales videre undersøkelser av instrumentets følsomhet i randomiserte kontrollerte studier, da dette foreløpig ikke er gjort (29-31, 33).

4.5.4 The Health Education Impact Questionnaire (heiQ)

HeiQ er et generisk pasientrapportert spørreskjema som ble utviklet i Australia i 2007 (34). Spørreskjemaet ble utviklet for å dekke mangelen på et måleinstrument med evne til å fange opp tidlige effekter av undervisningskurs for et bredt spekter av grupper med ulike kroniske sykdommer (34, 35). Det er utviklet gjennom flere stadier som inkluderte både helsepersonell, ledere for pasientopplæringskurs, akademikere og pasienter.

HeiQ består av 40 påstander fordelt på 8 selvstendige undergrupper. Hver undergruppe består av 4-6 påstander, se tabell 1.

Tabell 1: Navn og beskrivelse av undergruppene i heiQ

Undergruppe (antall påstander)	Nøkkelord
1. Positivt og aktivt engasjement i livet (5)	Bli engasjert i livet, intensjoner om handling, indikatorer som viser at man er engasjert og involvert i livet
2. Helserettet atferd (4)	Sunn atferd inkludert forebygging, diett, trening, avslapping; følbare endring
3. Tilegning av ferdigheter og teknikker (4)	Ferdigheter for å lette symptomer, ferdigheter og teknikker for å ta vare på egen helse
4. Konstruktive holdninger og tilnæringsmåter (5)	Begrense sykdommen; "jeg vil ikke la sykdommen kontrollere livet mitt"
5. Egenovervåking og innsikt (6)	Egenovervåking av kliniske indikatorer, egenstyring, sette fornuftige grenser eller mål og innsikt i å leve med et helseproblem
6. Orientering i helsetjenesten (5)	Kommunikasjon, beslutningsprosesser, forbindelser, forståelse; interaksjon med, og orientering, i helsesystemet, inkludert egen tiltro til at man kan snakke med helsepersonell; forståelse av måter å bruke helsevesenet på for å dekke behovene
7. Sosial integrasjon og støtte (5)	Sosial interaksjon, følelse av sosial isolasjon som et resultat av sykdommen, "fellesskap" i en gruppe som leder til en følelse av støtte, søke støtte fra andre.
8. Emosjonelt velvære (6)	Total helsereelatert negativ påvirkning; holdning til livet; engstelse, stress, sinne og depresjon.

Alle påstandene besvares på en Likert-skala fra 1 til 4: (1) "veldig uenig", (2) "uenig", (3) "enig" og (4) "veldig enig". Ved utregning av resultater blir påstandene i hver undergruppe summert og delt på antall påstander. Høyere skår indikerer bedre status i alle skalaene unntatt "emosjonelt velvære", der høyere skår indikerer mindre velvære. Ved utregning av gjennomsnitt snues denne skåren. HeiQ er lisensiert, og ved godkjent søknad mottas syntax for utregning av resultatene. Ved manglende svar kan det kun gjøres analyser med respondenter som har besvart mer enn 50% av påstandene innenfor hver enkelt undergruppe. I påstander med manglende data skal det kalkuleres en skår ved å addere skåren på alle besvarte påstander og så dividere denne skåren på antall besvarte påstander (36).

HeiQ er oversatt til over 20 ulike språk, men kun et fåtall av studiene rapporterer psykometriske egenskaper (35). I Europa er heiQ validert i Frankrike, Nederland, Norge og Tyskland (34, 35, 37-39). Utvalget i studiene består av pasienter med ulike kroniske sykdommer. I tre av studiene utgjør 18%-50% av studiens samlede utvalg pasienter med muskel- og skjelettlidelser (34, 35, 38). Det er ikke gjort validering av heiQ primært basert på pasienter med inflammatoriske revmatiske sykdommer som forfatter av denne studien er kjent med. Valideringsstudier av den originale versjonen av heiQ bekrefter en åttedelt faktorstruktur og gode psykometriske egenskaper (34, 40). Faktoranalyser bekrefter god validitet for alle åtte undergrupper (35, 37). Resultater fra åttedelt faktoranalyse i tre studier viser imidlertid en korrelasjonskoeffisient som varierer fra $r=0.69$ til $r=0.90$ mellom undergruppene "positivt og aktivt engasjement i livet" og "konstruktive holdninger og tilnærmingsmåter" (34, 38, 39). Dette reiser spørsmål om disse to skalaene konseptuelt og empirisk måler ulike konstruksjoner (37). En-modells faktoranalyse for hver av undergruppene viser faktorladning under 0.6 i tre eller flere spørsmål i undergruppen "egenovervåking og innsikt" i flere studier, noe som antyder en begrenset relasjon mellom spørsmålene (35, 37, 39). Denne undergruppen har også en intern konsistens, målt med Cronbachs alfa, på under 0.6 i flere studier, mens de resterende skalaene viser verdier over 0.7 (34, 35, 37). Test-retest er kun rapportert i en studie med en ICC som varierer mellom 0.61 og 0.84 for de ulike undergruppene (37). Verdier for manglende data er kun rapportert i to av studiene og ligger mellom 0.5-6.6% (34, 35). Resultater for gulv- og takeffekter er ikke rapportert i noen av studiene. Ingen av studiene har evaluert responsivitet. En randomisert kontrollert studie med den norske utgaven av heiQ har funnet signifikante endringer etter intervensjon i de to undergruppene "tilegning av ferdigheter og teknikker" og "konstruktive holdninger og tilnærmingsmåter" (41). Utvalget i studien består av pasienter med psoriasis. Ingen studier til nå har testet responsivitet i utvalg bestående av pasienter med inflammatoriske revmatiske sykdommer.

4.5.5 Patient Activation Measure (PAM)

Konseptet "å være aktivert" ble definert av Hibbard med flere i USA i 2004. Gruppen utviklet så det lisensierte måleinstrumentet PAM (42). Bakgrunnen for arbeidet var både helsepolitiske føringer med fokus på pasienters aktive deltakelse i egen helse og

helsetjenester, og the Chronic Illness Care Model som vektla pasientrettet omsorg med pasienter som aktive deltakere av behandlingsteamet. Pasientaktivering refererer til personens evne og vilje til å håndtere egen helse og helsevesenets tjenester. Begrepet aktivering fokuserer på kunnskaper, ferdigheter og selvtillit som kreves for den daglige håndteringen av egen helse. Pasienter med høy grad av aktivering forstår sin rolle i behandlingsprosessen og føler seg i stand til å håndtere denne rollen (42, 43).

PAM er et endimensjonalt pasientrapportert spørreskjema der pasientene blir bedt om å vurdere i hvor stor grad de er enig eller uenig i 13 ulike utsagn. Et eksempel på et utsagn er: "Jeg vet hvordan jeg skal forebygge forverring av min helsetilstand". Utsagnene er rangert etter vanskelighetsgrad der utsagn 1 er lettest og utsagn 13 er vanskeligst. Spørreskjemaet finnes i to utgaver med henholdsvis 13 og 22 utsagn. Begge utgavene finnes i en generisk versjon og en versjon for kronisk syke. Versjonen for kronisk syke med 13 utsagn er den mest brukte utgaven, og benyttes også i denne studien. Hvert utsagn skåres på en Likert-skala fra 1 til 5: (1) "helt uenig", (2) "nokså uenig", (3) "nokså enig", (4) "helt enig" og (5) "ikke aktuelt". I henhold til Insigna Health LLC sine retningslinjer for anvendelse av PAM, utregnes det en gjennomsnittsskår for utsagnene (44). Denne skåren transformeres til en standardisert aktiveringsskår fra 0 til 100 hvor 0=lavest aktivering og 100=høyest aktivering. Ved manglende data beregnes en poengsum ved å oppsummere svarene, dividere med antall ubesvarte spørsmål og multiplisere produktet med 13. Dette regnes ut via en matematisk formel (en syntax i SPSS) som man mottar ved søknad om lisens. Resultatene kan om ønskelig deles inn i fire nivåer fra lavt til høyt nivå av aktivering. Disse nivåene gir en konkret beskrivelse av pasientens kapasitet, evner, tro og sannsynlig atferd på forskjellige punkter langs skalaen og anvendes til å støtte hver enkelt pasient mer hensiktsmessig. Poengsummen fra 0-100 er mest aktuell for effektivvurdering av intervensjoner (45). En av de vanligste bruksområdene til PAM er å skreddersy støtte i henhold til de fire pasientaktiveringsnivåene (45). Stadig flere benytter imidlertid PAM for evaluering av intervensjoner som pasientopplæringskurs, og i 2008 ble PAM valgt som måleinstrument for evaluering av pasientopplæringskurs ved lærings- og mestringssentrene i Helse Midt-Norge (45, 46).

PAM er omfattende validert og forsket på. Oversettelse og validering er utført i fem europeiske land; Italia, (47), Danmark (48), Nederland (49), Norge (46, 50) og Tyskland (51). I

tillegg er det utført en psykometrisk sammenlikningsstudie av valideringsstudiene utført i Norge, Danmark, Nederland og Tyskland (52). Utvalget i den danske studien består av pasienter med diabetes, mens utvalget i et av de to norske studiene består av pasienter med psykisk sykdom. Resten av studiene har utvalg bestående av pasienter med et bredt spekter ulike kroniske sykdommer. Sammenlikningsstudien konkluderer med gode psykometriske egenskaper i alle fire utgavene av PAM, og liknende psykometriske funn på tvers av studiene. Funnene er også sammenliknbare med valideringsstudien til originalversjonen av PAM. Endimensjonal struktur bekreftes i flere av studiene (49-51). Intern konsistens målt med Cronbachs alfa ligger over 0.80 i alle studiene. Studiene viser gjennomgående få manglende data fra 0% til 3.2%. Det er rapportert gulveffekt i to utsagn i den ene norske valideringsstudien, og takeffekt på mellom 16.4% til 65.8% i 9-13 utsagn i flere av studiene (46-48, 50, 51). En randomisert kontrollert studie har funnet signifikante endringer i PAM fire måneder etter intervensjon (53). Studien bestod av pasienter med RA og PsA og intervensjonen bestod av et pasientopplæringskurs med fire gruppebaserte møter og et individuelt møte.

4.5.6 Spørreskjemaenes forståelighet og relevans

Spørsmålene som tar for seg forståelighet og relevans til spørsmålene i spørreskjemaene er selvkomponerte. Forståelighet ble undersøkt med spørsmålet "hvor forståelig synes du spørsmålene var?". Spørsmålet besvares med en 10-punkts (1-10) endepunktsforankret skala; (1) helt uforståelig og (10) svært lette å forstå. Respondentene ringer rundt et av tallene fra en til 10. Spørsmålenes relevansen ble undersøkt med spørsmålet "når du tenker på kurset du har deltatt på, hvor relevante synes du spørsmålene var?". Spørsmålet besvares på samme måte som ovenfor hvor skalaen har endepunktene (1) ikke relevante i det hele tatt og (10) svært relevante. Ved skåring av spørsmålene omgjøres skalaen til en ordinal skala med tre nivåer; (1) helt uforståelig/ikke relevante i det hele tatt, (2) noe forståelig/noe relevant og (3) veldig lette å forstå/svært relevante. Valg av cut-off må være gjennomtenkt da det kan påvirke og manipulere resultatene (54). For at ytterpunktene skulle veie likt, ble cut-off satt på fire og seks; nivå en (1-4), nivå to (5-6) og nivå tre (8-10). Spørsmålene er selvlagde og derfor ikke validerte, men spørsmålene brukes kun til å belyse problemstillingene og vil ikke ha en avgjørende betydning for resultatet.

4.6 Pasientopplæringskursene

Til sammen 13 pasientopplæringskurs er inkludert i studien. Kursene var gruppebaserte opplæringskurs i regi av revmatologiske avdelinger og/eller lærings- og mestringsentre (LMS) i Norge. De hadde ulike fokus, men en fellesnevner var å tilegne pasienten kunnskap og styrke deres evne til å håndtere sin inflammatoriske revmatologiske sykdom, se tabell 2.

tabell 2: Innholdet i hver enkelt kurs og hvilket spørreskjema som ble benyttet.

sA/sB	Kursholder/ diagnosegrupper		Varighet, dager	Diagnose/ sykdomslære	Medikamenter	Å leve med sykdom	Kosthold	Fysisk aktivitet/ fysioterapeut	Hjelpemidler/ ergoterapeut	Rettigheter og muligheter i NAV/ sosionom	Brukererfaringer/ brukerorganisasjoner	Gruppesamtale deltakere imellom	Utmattelse/sitenhet (Fatigue)	Smertes/ smertemestring	Behandlings-alternativer	Mulighet for pårørende å delta på kurset
sA	Ålesund sykehus	SpA	1													
sA	Ålesund sykehus	RA	2													
sA	Nordlandssykehuset	AS	3													
sA	Sykehuset Levanger	RA	2													
sA	UNN	SA	3													
sA	Revmatismesykehuset Lillehammer	RA	7													*
sA	Revmatismesykehuset Lillehammer	PMR	7													*
sA/sB [#]	DHS	RA	2													
sB	Revmatismesykehuset Lillehammer	Sjøgrens	7													*
sB	Nordlandssykehuset	RA/PsA	3													
sB	UNN	SLE/lupus	3													

[#]Kurset ble holdt tre ganger der sA ble utgitt en gang, sB to ganger. *Kurset har plass til 16 pasienter og 5 pårørende. SpA=Spondyloartritt, RA=revmatoid artritt, PsA=psoriasisartritt, SLE=systemisk lupus erytematosus, PMR=polymyalgia revmatika. DHS=Diakonhjemmet sykehus, UNN=Universitetssykehuset i Nord-Norge.

4.7 Statistiske analyser

Alle statistiske analyser er gjennomført med SPSS (Statistical Package for the Social Sciences) versjon 21.0. Datafilene er sjekket for usannsynlige verdier og normalfordeling før analyse. P-verdier <0.05 er valgt som kriterium for statistisk signifikans i analyse av før- og etterskår, og konfidensintervall er satt til 95%.

4.7.1 Bakgrunnsvariabler

Bakgrunnsvariabler presenteres med deskriptiv statistikk. Det benyttes gjennomsnitt og standardavvik (SD) eller min-maks for kontinuerlige variabler (median og range der data ikke er normalfordelt) og frekvens og prosent for kategoriske variabler. Min-maks velges til fordel for standardavvik for alder slik at man kan se nedre og øvre alder i utvalget.

4.7.2 Manglende data

Manglende data defineres her som ubesvarte spørsmål eller spørsmål besvart med to kryss, et kryss mellom to ulike svaralternativer eller andre avkryssninger som er vanskelig å tolke. Manglende data kan medføre skjevfordelte data, tap av informasjon, redusert statistisk styrke, økte standardfeil og svekket generaliserbarhet av funnene. Til tross for en direkte relasjon til kvalitet på statistisk slutningen, finnes ingen etablert cutoff-verdi for akseptabel prosentandel manglende data i et datasett i litteraturen (55). Det er likevel vanlig å operere med mindre enn 5% manglende data som akseptabelt. Det er ønskelig å finne minst mulig manglende data i et spørreskjema. En høy prosentandel manglende data for ett eller flere spørsmål kan for eksempel bety at spørsmålet er vanskelig å besvare, at spørsmål og svaralternativer ikke passer sammen eller at svaralternativene ikke er optimale. Manglende data undersøkes med frekvensanalyser for hvert enkelt spørsmål i hvert av spørreskjemaene ved baseline.

4.7.3 Gulv- og takeffekter

Gulv- og takeffekter er et begrep som beskriver hvor mange av respondentene som har besvart et spørsmål med henholdsvis laveste eller høyeste skåralternativ. Gulv- og takeffekter foreligger der $\leq 15\%$ av respondentene har besvart med henholdsvis lavest eller høyest

skåralternativ (56). Forekomst av gulv- og takeffekter indikerer at det finnes mer varians i konseptet av interesse enn det måleinstrumentet (spørsmål og/eller svaralternativ) er i stand til å fange opp. Forekomst av takeffekter kan derfor indikere at det finnes lavere eller høyere skår for spørsmålet, men at svaralternativene setter et "tak" for laveste eller høyeste mulige observasjon. Svaralternativene strekker seg ikke langt nok til å avdekke variansen. Forekomst av gulv- og takeffekter ved baseline kan medføre begrenset mulighet til å fange opp endringer før og etter intervensjon (56). Gulv- og takeffekter undersøkes med frekvensanalyser for hvert enkelt spørsmål i hvert av spørreskjemaene ved baseline.

4.7.4 Intern konsistens

Intern konsistens er et mål på hvor mye spørsmålene innen samme skala korrelerer og måler det samme konseptet (57). Mange begreper i helsevitenskap er "ikke-observerbare" og umulig å måle med kun et spørsmål. Eksempler på dette er mestringstro som måles med ASES og pasientaktivering som måles med PAM. Det konstrueres derfor et sammensatt mål med flere variabler relatert til ulike aspekter av begrepet. Samlet skaper spørsmålene en sammensatt skala som måler begrepet på en mer helhetlig måte enn hva ett spørsmål vil være i stand til (57). Siden spørsmålene skal måle et felles latent begrep, må de korrelere med hverandre. Til sammen skal de korrelerte spørsmålene gjenspeile individuell variasjon i forhold til et gitt fenomen. Cronbachs alfa er en av de mest brukte indikatorene på intern konsistens og benyttes også i denne studien. Cronbachs alfa har en korrelasjonskoeffisient som strekker seg fra 0 til 1 (56). Verdier mellom 0.7-0.95 er tilfredsstillende, men verdier på 0.8 eller mer foretrekkes (56, 58). Ved lav Cronbachs alfa bør det undersøkes hvor godt hvert enkelt spørsmål korrelerer med den totale skåren (58). Korrelasjonsverdier under 0.3 indikerer at spørsmålet måler noe annet enn skalaen i helhet. Cronbachs alfa påvirkes både av styrken i korrelasjonen mellom spørsmål og antall spørsmål. Desto sterkere korrelasjonen mellom spørsmålene og desto flere spørsmål, desto nærmere 1 vil Cronbachs alfa være (57). I skalaer med færre enn 10 spørsmål kan det ved lav alfa-verdi være mer korrekt å rapportere verdien for den gjennomsnittlige korrelasjonen mellom variablene. En optimal verdi for gjennomsnittlig korrelasjon mellom variablene strekker seg fra 0.2 til 0.4 (58). Intern konsistens måles ved å analysere Cronbachs alfa ved baseline.

4.7.5 Instrumenters evne til å fange opp endring

Forskjell i skår før og tre måneder etter kurs undersøkes med paret T-test for normalfordelte data og paret Wilcoxon-test for ikke-parametriske data. Det blir tatt hensyn til måleinstrumentenes instruksjon om håndtering av manglende data før analyse. For å finne effektstørrelsen på endringen benyttes standard response mean (SRM). En effektstørrelse sier noe om hvor stor effekten er og hvilken vei den går. SRM er en valid metode som er mye anvendt for å beregne effektstørrelse og sammenlikning av resultater på tvers av studier. SRM utregnes ved å dele gjennomsnittlig endring på standardavviket til endringen (59). Verdien som oppnås strekker seg fra 0-1 og tolkes med Cohen's verdier der 0-<0.2 er ingen effekt, 0,2-<0,5 liten effekt, 0,5-<0,8 moderat effekt og 0,8 og mer er stor effekt (60).

4.7.6 Spørreskjemaenes forståelighet og relevans

Forståeligheten til spørsmålene blir undersøkt ved baseline, mens relevansen undersøkes etter tre måneder. Resultatene presenteres med deskriptiv statistikk. Det benyttes frekvens og prosent i fremstillingen.

4.8 Etske avveininger

Det er sendt inn fremleggsvurdering til Regional etisk komité (REK) sør-øst (vedlegg 11). Prosjektet ble klarert for gjennomføring uten godkjenning av REK fordi formålet med prosjektet ikke var å fremskaffe ny kunnskap om helse, sykdom, diagnostikk eller behandling og derfor faller utenom helseforskningslovens virkeområde. Det ble også innsendt meldeskjema til personvernombudet ved OUS da det er identifiserbare opplysninger i studien (vedlegg 12). Selv om studien ikke er fremleggspliktig overfor REK er det likevel viktig at studien gjennomføres på en forsvarlig måte med hensyn til regler for taushetsplikt, personvern og samtykke. For å sikre informert samtykke, fikk pasientene både muntlig og skriftlig informasjon om studien (vedlegg 9). Her ble de informert om studiens formål, håndtering av forskningsdata, frivillig deltagelse og muligheten til når som helst å trekke seg fra studien eller takke nei til deltagelse uten å oppgi noen grunn og uten at det ville få konsekvenser for verken dem eller deres deltagelse på kurset. Alle deltakerne ga skriftlig samtykke til å delta i studien. Det er ønskelig at pasientene skal få god tid til å vurdere deltagelse. Kursdeltakere ved Diakonhjemmet og Ålesund sykehus fikk tilsendt skriftlig

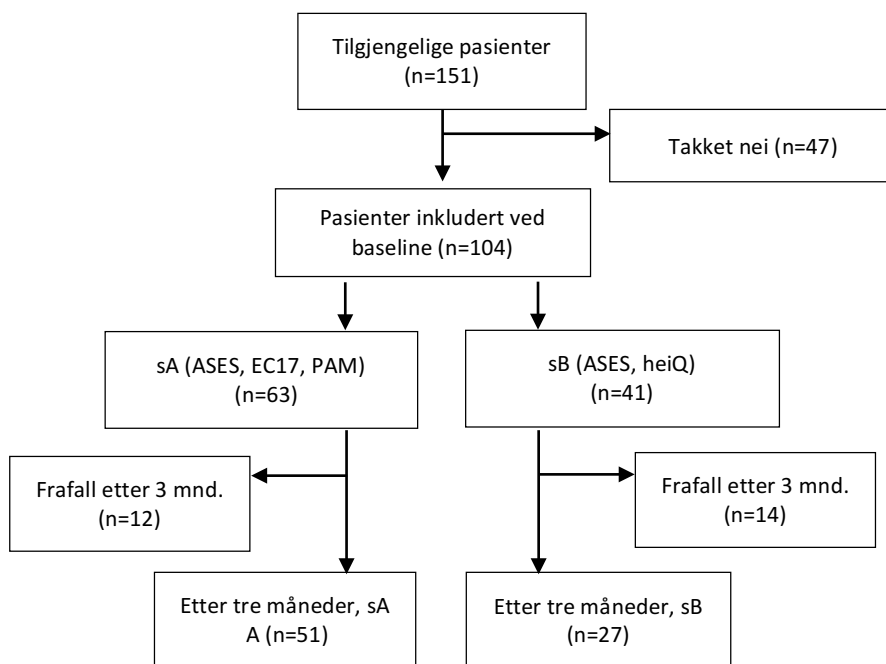
informasjon, samtykkeskjema og spørreskjema 1-2 uker før kursoppstart, noe som ga pasientene god tid til å tenke seg om. Ved de andre kursstedene fikk pasientene informasjon samme dag som kurset startet, da forfattere av studien ikke hadde tilgang til informasjon om kursdeltakere i forkant av kursoppstart. Dette gir langt kortere betenkningstid, men ble likevel vurdert som godkjent på bakgrunn av spørreskjemaenes innhold og studiens formål der anvendeligheten av spørreskjemaene er i hovedfokus og ikke pasientresultater. De revmatologiske avdelingene som bidro til utlevering/innlevering av skjema fikk et lite økonomisk vederlag som takk for hjelpen. Det er ingen åpenbare ulemper ved å delta i studien, bortsett fra tidsbruken det vil ta å besvare spørreskjemaet. Alle opplysningene ble behandlet uten navn og fødselsdato eller andre direkte gjenkjennende opplysninger. Spørreskjemaene ble merket med id-nummer som kan knyttes opp mot en navneliste. Spørreskjemaene og nøkkel mellom pasient-id og id-nummer ble og er oppbevart i adskilte låste skap som kun er tilgjengelig for meg og prosjektleder Heidi A. Zangi. Datafilene til masterprosjektet har foregått på PC tilknyttet en passordbeskyttet bruker-id på Diakonhjemmet sykehus. Filene ligger lagret på Diakonhjemmet sykehus sin forskningsserver i en mappe som kun jeg og prosjektleder har tilgang til.

5 Resultater

I første del av kapittelet presenteres flytdiagram over respondentene i studien. Deretter følger bakgrunnsvariabler og analyseresultater tilknyttet problemstillingene. Til slutt kommer resultatene tilknyttet spørreskjemaenes forståelighet og relevans.

5.1 Flytdiagram over respondenter

Totalt 151 pasienter deltok på 13 ulike pasientopplæringskurs, hvorav 104 (69%) ble inkludert i studien, se figur 2. Eksakt antall kursdeltakere ble innhentet i ettertid, noe som medførte at det ikke var mulig å innhente oversikt over antall kursdeltakere i tre av kursene. Det er tatt høyde at disse tre kursene var fulle (16 deltakere per kurs), selv om det ofte forekommer frafall i kursene grunnet blant annet sykdom. Av de 104 inkluderte fikk 63 deltakere utdelt sA og 41 deltakere fikk sB. Etter tre måneder besvarte henholdsvis 81% og 66% av disse sA og sB.



Figur 2: Flytdiagram over respondenter

5.2 Bakgrunnsvariabler

Tabell 3 viser bakgrunnsvariablene for de inkluderte pasientene. Flertallet av respondentene var kvinner. Alderen varierte fra 19-80 år i hele gruppen. Det var noen prosentvise forskjeller mellom gruppene i de fleste variablene.

Tabell 3: Demografiske bakgrunnsvariabler

	Alle N=104	sA n=63	sB n=41
Kjønn, n (%) kvinner	76 (73)	42 (67)	34 (83)
Alder i år, gj.sn. (min-maks)	50 (19-80)	52 (19-80)	46 (21-77)
Sivilstatus, n (%)			
Gift	45 (43)	27 (43)	18 (44)
Samboer	25 (24)	16 (25)	9 (22)
Andre, ugift/skilt/separert/enke/enkemann	28 (28)	16 (25)	12 (29)
Utdanning, n (%)			
Grunnskole	20 (19)	14 (22)	6 (15)
3 åring videregående skole	38 (37)	26 (41)	12 (29)
Utdanning over 3-åring videregående skole	45 (43)	23 (37)	22 (54)
Arbeid, n (%)			
I arbeid	57 (55)	31 (49)	26 (63)
Ikke i arbeid	44 (42)	30 (48)	14 (34)
Røyk og snus, n (%)			
Røyker ikke	90 (87)	52 (83)	38 (93)
Snuser ikke	90 (87)	57 (91)	33 (81)
Mosjon, n (%)			
Sjeldnere enn en gang i uka eller aldri	17 (16)	9 (14)	8 (20)
En gang i uka	14 (14)	8 (13)	6 (15)
2-3 ganger i uka	44 (42)	27 (43)	17 (42)
Omtrent hver dag	29 (28)	19 (30)	10 (24)

sA = spørreskjema A. sB = spørreskjema B.

Bakgrunnsvariabler med manglende data var alder (2.88%), kjønn (0.96%), sivilstatus (5.76%), utdanning (0.96%), arbeid (2.88%) og snus (4.80%). Det antas at manglende verdier ofte skyldes tilfeldige og praktiske forhold.

5.3 Manglende data

Alle måleinstrumentene hadde gjennomgående under 5% for manglende data ved baseline (tabell 4). HeiQ "konstruktive holdninger og tilnæringsmåter" var det eneste måleinstrumentet som ikke hadde manglende data, mens ASES andre symptomer hadde høyest forekomst på 2,2%.

Tabell 4: Manglende data i måleinstrumentene ved baseline

Måleinstrument (n)	Antall (%)
ASES smerter (104)	7 (1.3)
ASES andre symptomer (104)	14 (2.2)*
EC-17 (63)	3 (0.3)
PAM (63)	6 (0.7)
heiQ (41)	
Helserettet adferd	3 (1.8)
Positivt og aktivt engasjement i livet	4 (1.9)
Emosjonelt velvære	4 (1.6)
Egenovervåkning og innsikt	6 (2.4)
Konstruktive holdninger og tilnæringsmåter	0
Tilegning av ferdigheter og teknikker	2 (1.2)
Sosial integrasjon og støtte	3 (1.5)
Orientering i helsetjenesten	3 (1.5)

* Seks av de ubesvarte spørsmålene tilhører en enkelt respondent.

5.4 Gulv- og takeffekter

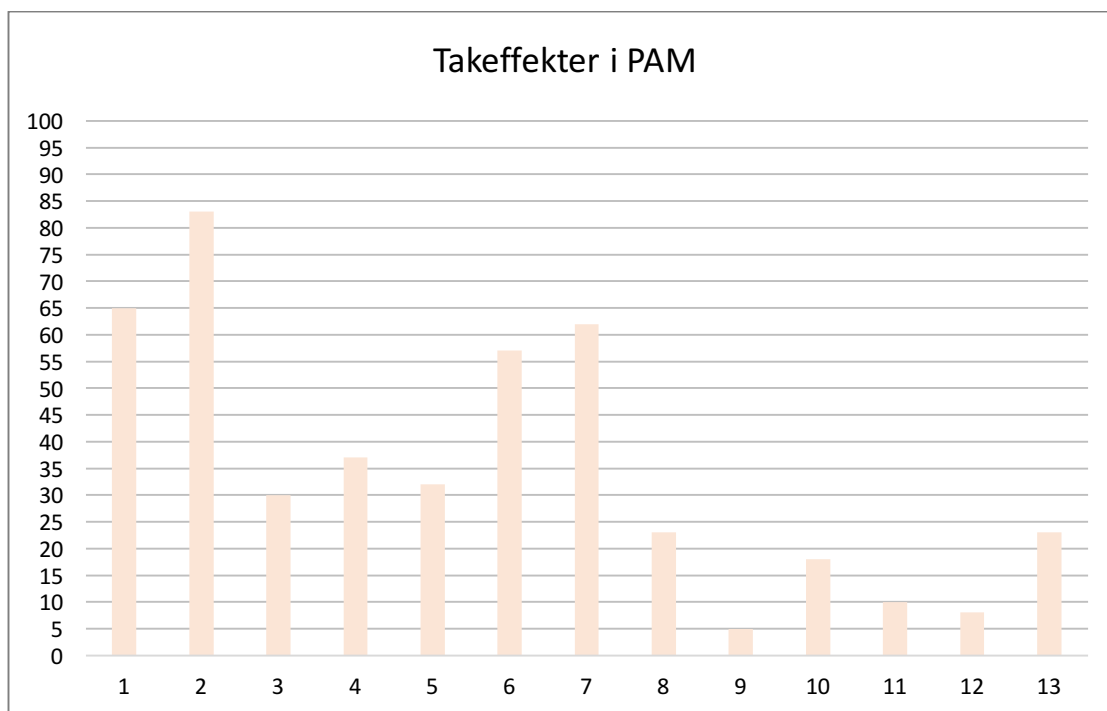
Ingen av måleinstrumentene hadde gulveffekter ved baseline. Med unntak av heiQ "emosjonelt velvære" og "tilegning av ferdigheter og teknikker", hadde alle måleinstrumentene takeffekter i et eller flere spørsmål (tabell 5).

Tabell 5: Oversikt over spørsmål med takeffekter ved baseline

Måleinstrument (antall spørsmål)	Spørsmål med takeffekt	Respondenter totalt	Respondenter med høyeste skår (%)
ASES smerter	2.	103	20 (19)
	4.	103	16 (16)
ASES andre symptomer EC-17 (17)	3.	102	22 (22)
	7.	63	14 (22)
	9.	63	19 (30)
	10.	63	10 (16)
	11.	63	11 (18)
heiQ helserettet adferd (4)	1.	40	9 (22)
	9.	41	7 (17)
	13.	41	7 (17)
heiQ positivt og aktivt engasjement i livet (5)	5.	40	14 (35)
	8.	41	13 (32)
	10.	41	9 (22)
heiQ egenovervåkning og innsikt (6)	3.	40	10 (25)
	6.	39	4 (10)
	11.	40	13 (33)
heiQ konstruktive holdninger og tilnærmingsmåter (5)	27.	41	10 (25)
	39.	41	7 (17)
	40.	41	10 (24)
heiQ sosial integrasjon og støtte (5)	22.	40	10 (24)
	35	41	10 (24)
heiQ orientering i helsetjenesten (5)	32	41	9 (22)
PAM (13)	1.	63	41 (63)
	2.	63	52 (83)
	3.	63	19 (30)
	4.	63	23 (37)
	5.	63	20 (32)
	6.	63	36 (57)
	7.	63	39 (62)
	8.	62	14 (22)
	10.	63	11 (18)
	13.	63	13 (21)

ASES=The Arthritis Self Efficacy Scale, EC17=Effective Consumer Scale, heiQ= The Health Education Impact Questionnaire, PAM=Patient Activation Measure.

PAM utpekte seg med takeffekter i 10 av 13 spørsmål, noe som utgjør 77% av alle spørsmålene i måleinstrumentet. Takeffektene strakk seg fra 21% til 83%, se figur 3. Dette var noe høyere enn den prosentvise forekomsten i de andre måleinstrumentene som strakk seg fra 10% til 35%.



Figur 3: Prosentvis fremstilling av takeeffekt i de 13 spørsmålene i PAM

Det ble også utført analyser av gulv- og takeffekter ved å summere alle spørsmålene i hvert enkelt spørreskjema. Det forelå da hverken gulv- eller takeffekter i noen av spørreskjemaene.

5.5 Intern konsistens

HeiQ "helserettet adferd" og "egenovervåkning og innsikt" hadde en Cronbachs alfa på henholdsvis 0.61 og 0.66. Den gjennomsnittlige korrelasjonen mellom variablene i disse to undergruppene var på henholdsvis 0.29 og 0.25. Dette er innenfor akseptable verdier. ASES smerter og ASES andre symptomer, EC-17, PAM og de seks resterende undergruppene i heiQ hadde en akseptabel til god Cronbachs alfa på mellom 0.71 til 0.91. EC-17 hadde den høyeste verdien (0.91) (tabell 6).

Tabell 6: Cronbachs alfa

Måleinstrument	N	Antall spørsmål	Cronbachs alfa
ASES smerter	101	5	0.75
ASES andre symptomer	100	5	0.81
EC-17	61	17	0.91
PAM	56	13	0.78
heiQ			
Helserettet adferd	39	4	0.61 (0.29*)
Positivt og aktivt engasjement i livet	38	5	0.81
Emosjonelt velvære	39	6	0.86
Egenovervåkning og innsikt	37	6	0.66 (0.25*)
Konstruktive holdninger og tilnæringsmåter	41	5	0.72
Tilegning av ferdigheter og teknikker	39	4	0.83
Sosial integrasjon og støtte	39	5	0.85
Orientering i helsetjenesten	38	5	0.72

*Gjennomsnittlig korrelasjon mellom variablene

5.6 Endringer i skår etter 3 måneder

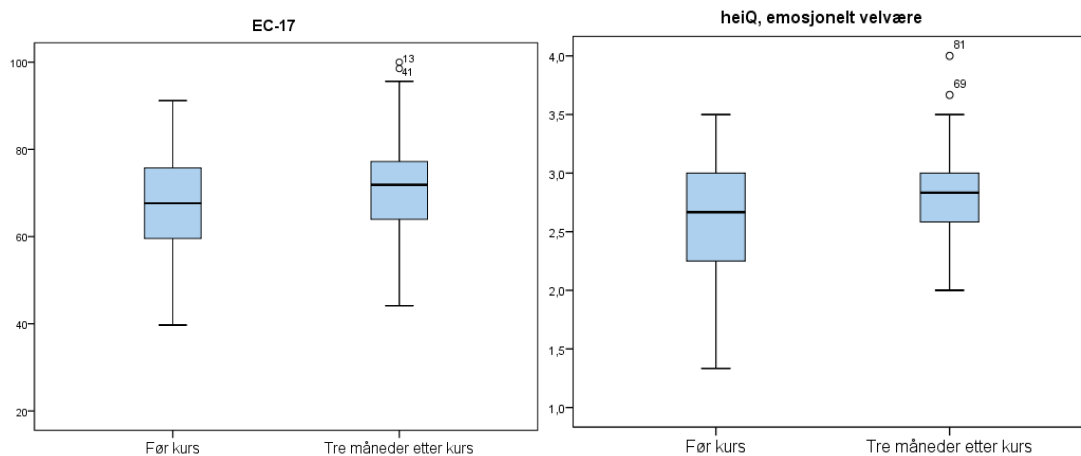
Alle måleinstrumentene hadde stigning i gjennomsnittlig skår fra baseline til tre måneder etter kurs, unntatt heiQ "positivt og aktivt engasjement i livet" som hadde en liten negativ endring på -0.12. Det var statistisk signifikant forbedring etter tre måneder på EC-17 og heiQ "emosjonelt velvære", "tilegning av ferdigheter og teknikker" og "orientering i helsetjenesten" (tabell 7).

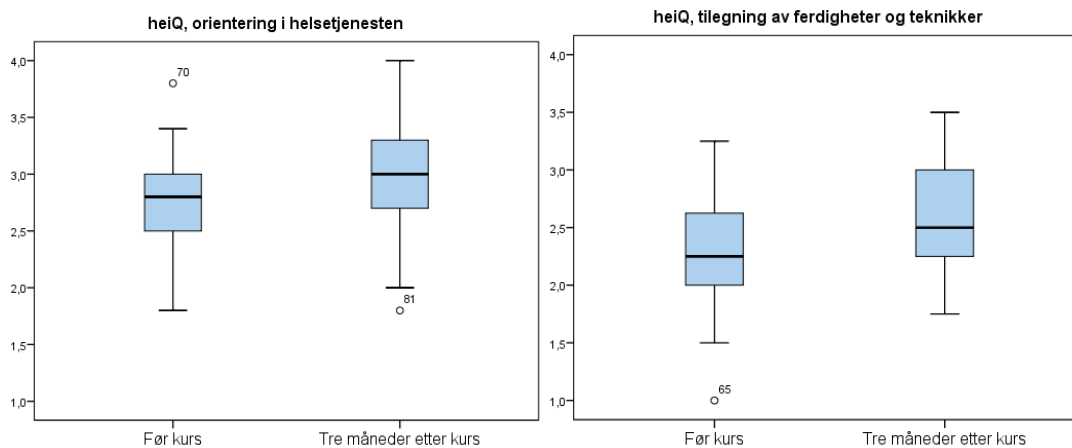
Tabell 7: T-test med skår før og tre måneder etter for de fire spørreskjemaene.

Måleinstrument (n)	skår	Gj.sn. (SD) før kurs	Gj.sn. (SD) etter kurs	Gj.sn. (SD) endring	95% KI av forskjell	P-verdi
ASES smerter (77)	1-5	3.21 (0.81)	3.33 (0.83)	0.12 (0.89)	-0.08, 0.32	0.227
ASES andre symptomer (76)	1-5	3.31 (0.69)	3.47 (0.70)	0.16 (0.82)	-0.24, 0.35	0.088
EC-17 (51)	0-100	67.7 (12.1)	71.3 (12.5)	3.6 (9.6)	0.9, 6.3	0.01
PAM (51)	0-100	60.0 (11.1)	63.7 (16.2)	3.7 (16.8)	-1.0, 8.4	0.122
heiQ (27)	1-4					
Helserettet adferd		2.84 (0.52)	2.86 (0.66)	0.02 (0.58)	-0.21, 0.25	0.869
Positivt og aktivt engasjement i livet		3.09 (0.46)	2.98 (0.47)	-0.12 (0.41)	-0.28, 0.45	0.151
Emosjonelt velvære		2.63 (0.52)	2.84 (0.50)	0.21 (0.41)	0.05, 0.37	0.012
Egenovervåkning og innsikt		2.72 (0.45)	2.89 (0.48)	0.18 (0.46)	-0.01, 0.36	0.058
Konstruktive holdninger og tilnæringsmåter		3.01 (0.47)	3.07 (0.39)	0.06 (0.40)	-0.10, 0.22	0.452
Tilegning av ferdigheter og teknikker		2.29 (0.53)	2.58 (0.47)	0.30 (0.61)	0.05, 0.54	0.018
Sosial integrasjon og støtte		2.61 (0.57)	2.80 (0.42)	0.20 (0.52)	-0.01, 0.41	0.06
Orientering i helsetjenesten		2.76 (0.43)	2.96 (2.96)	0.22 (0.40)	0.04, 0.36	0.015

ASES=The Arthritis Self Efficacy Scale, EC17=Effective Consumer Scale, heiQ= The Health Education Impact Questionnaire, PAM=Patient Activation Measure. SD=standard avvik, KI=konfidensintervall, SMR: Standard Response Mean. Nivå av signifikans $p < 0.05$.

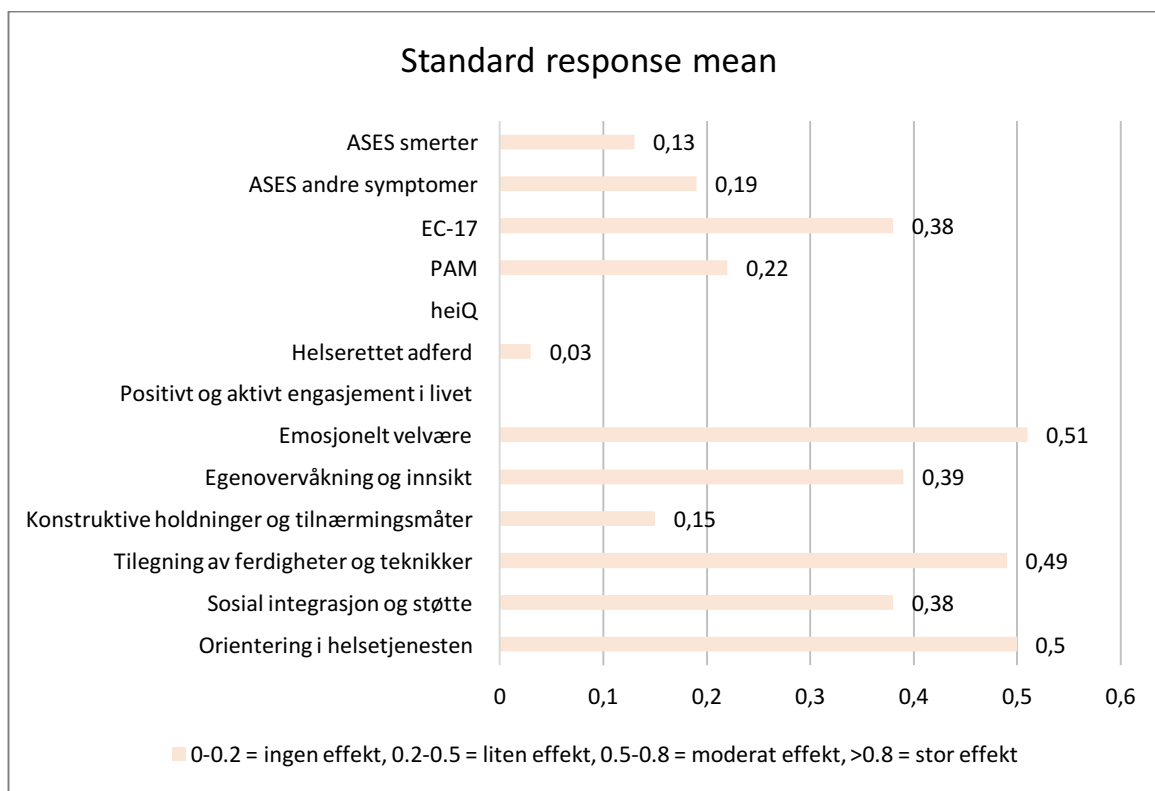
Boxplottet viser gjennomsnittlig skår for måleinstrumenter med statistisk signifikant endring ved begge måletidspunkt (Figur 4).





Figur 4: Gjennomsnittlig skår før og etter kurs i EC-17 og heiQ "emosjonelt velvære", "orientering i helsetjenesten" og "tilegning av ferdigheter og teknikker". EC17=Effective Consumer Scale, heiQ= The Health Education Impact Questionnaire.

SRM tilsvarende liten effekt ble funnet i EC-17, PAM, heiQ "egenovervåking og innsikt" og "sosial integrasjon og støtte". HeiQ "emosjonelt velvære" og "orientering i helsetjenesten" hadde en moderat effektstørrelse (SRM 0.51 og 0.5). Effektstørrelsen nærmet seg moderat på heiQ "tilegning av ferdigheter og teknikker" (SRM = 0.49). ASES andre symptomer nærmet seg "liten effekt" (SRM 0.19), mens ASES smerter og de resterende undergruppene i heiQ hadde effektstørrelse tilsvarende ingen effekt se (Figur 5).



Figur 5: Standard response mean for hvert enkelt måleinstrument. ASES=The Arthritis Self Efficacy Scale, EC17=Effective Consumer Scale, heiQ= The Health Education Impact Questionnaire, PAM=Patient Activation Measure.

5.7 Spørreskjemaenes forståelighet og relevans til kursene

73% av respondentene i sA og 68% i sB syntes spørsmålene var svært enkle å forstå. Mens 80% syntes spørsmålene i sA var svært relevante, var det 59% som syntes spørsmålene sB var svært relevante. Det var kun 2% som syntes spørsmålene i sA ikke var relevante i det hele tatt, mens det var 11% som syntes spørsmålene i sB ikke var relevante i det hele tatt (tabell 8).

Tabell 8: Respondentenes vurdering av spørsmålenes forståelighet og hvor relevante de er i forhold til kursenes innhold

Spørreskjema før kurs	Helt uforståelig	Noe forståelig	Svært enkle å forstå
sA (62)	5(8)	12(19)	45(73)
sB (40)	4(10)	9(22)	27(68)
Spørreskjema 3 måneder etter kurs	Ikke relevante i det hele tatt	Noe relevante	Svært relevante
sA (51)	1(2)	9(18)	41(80)
sB (27)	3(11)	8(30)	16(59)

sA: spørreskjema A, sB: spørreskjema B.

6 Presentasjon og diskusjon av funn

I denne studien undersøkte vi anvendeligheten til måleinstrumentene ASES smerter og andre symptomer, EC-17, heiQ og PAM med hensyn til manglende data, gulv- og takeffekter, intern konsistens og før- og etterskår. Alle måleinstrumentene viste tilfredsstillende lave forekomster av manglende data (<5%). Ingen av måleinstrumentene hadde gulveffekter. To av undergruppene i heiQ, "emosjonelt velvære" og "tilegning av ferdigheter og teknikker" hadde heller ingen takeffekter. Takeffekter forekom i et til ti spørsmål i de resterende måleinstrumentene med høyest forekomst i PAM. Cronbachs alfa var akseptabel i ASES smerter (0.75) og ASES andre symptomer (0.81), EC-17 (0.91), PAM (0.78) og seks av undergruppene i heiQ (0.72 - 0.86). HeiQ "helserettet adferd" og "egenovervåking og innsikt" hadde Cronbachs alfa på henholdsvis 0.61 og 0.66. EC-17 og heiQ "emosjonelt velvære", "tilegning av ferdigheter og teknikker" og "orientering i helsetjenesten" hadde statistisk signifikante endringer og en SRM tilsvarende liten til moderat effekt. PAM og heiQ "egenovervåking og innsikt" og "sosial integrasjon og støtte" hadde SRM verdier tilsvarende liten effekt, men ingen statistisk signifikante endringer.

6.1 Diskusjon av sentrale funn

6.1.1 Manglende data

Manglende data er mer regelen enn unntaket i kvantitativ forskning, og viktig å adressere da det er direkte relatert til kvaliteten på statistiske slutninger (55). Alle måleinstrumentene hadde forekomst av manglende data (0.3%-2.2%). Med en øvre akseptabel grense på 5% er alle måleinstrumentene godt innenfor godkjent verdi. I tillegg til å undersøke prosentvis forekomst av manglende data, bør man også undersøke eventuelle mønstre (55). Opphoping av antall manglende data i et eller flere spørsmål i et spørreskjema kan antyde problemer med de aktuelle spørsmålene. Vi foretok derfor analyser for å se på eventuelle mønstre i de manglende dataene. Vi fant at de manglende dataene var jevnt fordelt over spørsmålene i hvert av de fire måleinstrumentene. Foruten en respondent med seks ubesvarte spørsmål i ASES andre symptomer, var forekomst av manglende data spredt utover respondentene i alle måleinstrumentene. De seks ubesvarte spørsmålene i ASES andre symptomer stod alene på

en A4 side og respondenten hadde ingen andre ubesvarte spørsmål i resten av spørreskjemaet. Årsaken til dette kan være mange, men det er nærliggende å anta at dette var et uhell der respondenten hoppet over en side. Studien benytter en nyutviklet skåringsversjon av ASES som tidligere kun er benyttet i en valideringsstudie. Lav forekomst av manglende data er viktige funn som tyder på at svaralternativene passer til spørsmålene og er forståelige for respondentene. Denne antagelsen styrkes ved at valideringsstudien viser liknende funn (25). Lav forekomst av manglende data i EC-17 og PAM samsvarer med tidligere funn (29, 32, 46-52). Det er noe lavere forekomst av manglende data i heiQ i denne studien sammenliknet med tidligere rapporterte funn (34, 35). Det vil være fornuftig å anta en høy forekomst av manglende data hvis respondentene ikke forstår spørsmålene. Lave forekomster av manglende data samsvarer derfor med resultatene av respondentenes vurdering av spørsmålenes forståelighet, der 73% i sA og 68% i sB vurderte spørsmålene til å være svært enkle å forstå.

6.1.2 Gulv- og takeffekter

Vi fant ikke gulveffekter i noen av måleinstrumentene. Derimot var det takeffekter i et eller flere spørsmål i alle måleinstrumentene unntatt heiQ undergruppene "emosjonelt velvære" og "tilegning av ferdigheter og teknikker". Gulv- og takeffekter er i varierende grad rapportert i tidligere studier av ASES, EC17, heiQ og PAM. Fordi forekomst av takeffekter kan påvirke instrumenters evne til å fange opp endringer er dette sentralt å undersøke før man velger måleinstrument til en effektstudie.

PAM er det måleinstrumentet med høyest forekomst av takeffekter. Dette gjelder både antall spørsmål med takeffekter (10/13) og prosentandel takeffekter i spørsmålene (18%-83%). Takeffekter reduserer evnen til å fange opp endringer ved å begrense variasjonsområdet på svarene, svaralternativene legger "tak" på hva som kan måles og ikke (58, 61). Desto flere respondenter som skårer høyeste svaralternativ før intervensjon, desto vanskeligere er det å fange opp endring etter intervensjon. Dette øker sannsynligheten for å avvise nullhypotesen når den egentlig er sann (type 2-feil). Man oppnår et feilaktig resultat med ikke-signifikante endringer (58). PAM har en ikke-signifikant endring tre måneder etter kurs. Fordi PAM har høye forekomster av takeffekter kan dette ha innvirket på resultatet. Hvis dette er tilfellet kan

pasientopplæringskursene ha hatt effekt som ikke fremkommer i analysene på grunn av instrumentering. Instrumentering er forhold ved måleinstrumentet som kan resultere i kunstige resultater og svekke den indre validiteten i en studie (62). Slike begrensninger ved et måleinstrument er spesielt viktig å være klar over når man undersøker helserelaterte virkninger av en behandlingsintervensjon. Hvis flere studier feilaktig resulterer i manglende effekt, kan studiene til sammen påvirke helsepolitiske beslutninger på feil grunnlag (58, 61). Forekomst av gulv- og takeffekter kan variere avhengig av diagnose fordi ulike sykdommer gir ulike utfordringer. Takeffekter i PAM er rapportert i 10 til 13 utsagn i flere europeiske valideringsstudier (46-48, 50, 51). Disse studiene har utvalg bestående av pasienter med flere ulike kroniske sykdommer. Resultatene kan derfor tyde på at responskategoriene ikke dekker relevante svaralternativer for ulike pasientpopulasjoner med en rekke forskjellige kroniske sykdommer i Europa. På den andre siden er utsagnene i PAM bevisst utformet med stigende vanskelighetsgrad. Målet er at man enkelt skal kunne se hvor aktivert en pasient er kun ved å se på avkrysningene. Ut fra resultatene kan behandling og støtte skreddersys for den enkelte pasient. Den økende vanskelighetsgraden er til dels synlig på fordelingen av takeffektene som tenderer til å være høyere i de første spørsmålene, se figur 3. Ved bruk av PAM for kartlegging av aktivering hos enkeltpasienter er det ønskelig å se høyeste skår på flest mulig utsagn da dette betyr høy grad av aktivering og mindre behov for støtte. Spørsmålet er om denne utformingen av PAM er uheldig når instrumentet skal benyttes til evaluering av intervensjoner da flere studier har vist at det medfører høye takeffekter ved baseline (46-48, 50, 51).

ASES, EC-17 og seks undergrupper i heiQ har spørsmål med takeffekter, men det er både færre spørsmål og lavere prosentandel takeffekter i disse måleinstrumentene sammenliknet med PAM. Det finnes ingen teoretisk klarhet i hvor stor andel takeffekter som skal til for å svekke den indre validiteten og instrumentets evne til å fange opp endringer. Det må derfor benyttes sunn fornuft i vurderingen av takeffekter. En annen måte å måle gulv- og takeffekter er å summere alle svarene til hver enkelt respondent. Gulv- eller takeffekt defineres da ved at mer enn 15% av respondentene har besvart henholdsvis laveste eller høyeste skår på alle spørsmålene i måleinstrumentet ved baseline (56). En slik analyse viste ingen gulv- eller takeffekter i noen av måleinstrumentene i vår studie. Forekomst av takeffekter undersøkt på denne måten ville hatt en større påvirkning på instrumentenes evne til å fange opp endringer

enn den måten takeffekter er undersøkt på i denne studien. Det er derfor en fordel med denne analysemetoden at den kan resultere i negative funn.

6.1.3 Intern konsistens

ASES, EC17, seks av undergruppene til heiQ og PAM har alle en Cronbachs alfa over 0.7. Dette er akseptable verdier som samsvarer med resultater fra tidligere studier av instrumentene. Cronbachs alfa i PAM (0.78) er likevel noe lavere enn i andre studier (>0.8) (46-51). I motsetning til studier på ASES og EC17 har ingen av de tidligere studiene av PAM inkludert pasienter med revmatiske sykdommer. Selv om Cronbachs alfa er lavere i denne studien sammenliknet med tidligere studier, er resultatet over nedre akseptable verdi på 0.7 og PAM vurderes derfor til å ha en akseptabel intern konsistens i dette utvalget. EC-17 har en Cronbachs alfa på 0.91. I litteraturen er det ulike syn på hvor nedre og øvre grense for Cronbachs alfa bør være. Veldig høye alfa-verdier kan indikere høy korrelasjon mellom spørsmålene i skalaen. Dette betyr at et eller flere av spørsmålene kan være overflødige fordi de spør om det samme, bare med en annen ordlyd. Mens noen argumenterer for en øvre grense på 0.90, har andre satt en grense på 0.95 (56, 57). Cronbachs alfa i EC-17 er godt under 0.95 og så vidt over 0.90 og anses derfor som god.

HeiQ undergruppene "helserettet atferd" og "egenovervåking og innsikt" er de eneste måleinstrumentene med Cronbachs alfa under 0.7 (0.61 og 0.66). Mens "helserettet atferd" har vist akseptable Cronbachs alfa-verdier på over 0.7 i tidligere studier, har "egenovervåking og innsikt" vist lavere verdier (<0.7) enn de syv resterende undergruppene i flere andre studier (34, 35, 37). "Grensen" på 0.7 er glidende, og en verdi på 0.66 trenger ikke umiddelbart å forkastes. Målinger av intern konsistens bør fortrinnsvis støttes av teori, tidligere forskning og fornuftig argumentasjon (57). En svakhet ved Cronbachs alfa er at verdien påvirkes av antall spørsmål som inngår i analysen (57). Ved lav Cronbachs alfa i måleinstrumenter med mindre enn 10 spørsmål kan den gjennomsnittlige korrelasjonen mellom variablene være en bedre indikator. Begge undergruppene har færre enn 10 spørsmål. Spørsmålene i undergruppen "helserettet atferd" og "egenovervåking og innsikt" har en gjennomsnittlig korrelasjonen på henholdsvis 0.29 og 0.25. Dette er innenfor den ønskede verdien (0.2-0.4) og antyder at intern konsistens i begge tilfeller er tilfredsstillende. Grunnet en lav Cronbachs alfa har vi også

undersøkt hvor godt hvert enkelt spørsmål korrelerer med den totale skåren (58). Ett av spørsmålene i "helserettet atferd har ett spørsmål med korrelasjonsverdi under 0.3 (0.24), noe som betyr at dette spørsmålet måler noe annet enn de resterende spørsmålene. Ett av spørsmålene i "egenovervåking og innsikt" har også en korrelasjonsverdi under 0.3. Dette spørsmålet har også vist svak korrelasjon med de resterende spørsmålene i tidligere studier (35, 37, 39). De samme studiene har funnet vesentlig lavere Cronbachs alfa og flere spørsmål med korrelasjonsverdi under 0.3 i "egenovervåking og innsikt", sammenliknet med denne studien. Dette kan tyde på at ett eller flere spørsmål i undergruppen måler noe annet enn de resterende spørsmålene. Samlet viser dette at videre undersøkelser av konseptet er nødvendig, noe som også er påpekt i andre studier (35, 37). Den interne validiteten vurderes som akseptabel i begge undergruppene, men på bakgrunn av resultatene i denne studien og i tidligere studier anbefales det likevel å følge med på resultater av intern validitet fra fremtidige studier med liknende utvalg. HeiQ brukes i flere land og en potensiell endring av instrumentet vil få konsekvenser for de ulike oversettelsene. Derfor er det viktig med flere og grundige undersøkelser før en eventuell endring foretas.

6.1.4 Skår før og etter gjennomført kurs

I denne studien ble det funnet statistisk signifikante endringer i EC-17 og heiQ undergruppene "emosjonelt velvære", "tilegning av ferdigheter og teknikker" og "orientering i helsetjenesten". For å sikre en valid undersøkelse av pasientopplæringskurs er det viktig at måleinstrumentene reflekterer kursenes innhold (2). Dette er også en betingelse for at måleinstrumentene skal kunne fange opp endring. I respondentenes vurdering av spørreskjemaenes relevans til kursenes innhold, er det 80% som synes sA er svært relevant mens det er 59% som synes sB er svært relevant. Det er interessant at det er så stor forskjell i skår av relevans mellom de to ulike gruppene. Fordi det er vanskelig å relatere skåren direkte til hvert enkelt måleinstrument, er det av interesse så se videre på innholdet i måleinstrumentene og hvorvidt de reflekterer kursenes innhold.

ASES smerter og andre symptomer viste ingen statistisk signifikante endringer og hadde lave SRM-verdier (0.13 og 0.19). ASES smerter handler om respondentenes tro på egen evne til å mestre smerter. Alle kursene var innom ulike aspekter av det å leve med inflammatorisk

revmatisk sykdom, som sykdomslære, legemidler, kosthold og fysisk aktivitet, se tabell 2. Kun åtte av 13 kurs (62%) inneholdt spesifikk undervisning om smerter, noe som kan påvirke mulighetene til å finne endring i før- og etterskår i ASES smerter. Fordi kursene inneholdt flere ulike temaer, er det også mulig at det ble for lite fokus på områder ASES andre symptomer dekker; for eksempel håndtering av humør, fatigue og håpløshet. Kursenes varighet kan også være en relevant faktor. De fleste kursene ble holdt over en til tre dager. En studie som resulterte i statistisk signifikante endringer i ASES, undersøkte effekten av et kurs som ble holdt over en lengre periode med møter over flere måneder (63). I henhold til teori om self-efficacy, kreves både tid og trening for å opprettholde mestringstro (64). Tre av kursene i denne studien bestod av et mestringsopphold over syv dager der det i tillegg til undervisning også ble arbeidet med mestring av utfordringer sykdommen gir. Det kunne vært interessant å undersøke om disse kursene isolert gav statistisk signifikante endringer i ASES, men på grunn av få respondenter var ikke dette gjennomførbart.

Kursinnholdet kan også ha hatt betydning for funn av endringer i enkelte av undergruppene i heiQ som viste ikke-signifikante endringer. De åtte undergruppene i heiQ måler ulike aspekter ved pasientopplæringen og kan anvendes isolert fra hverandre avhengig av innholdet i kursene. HeiQ undergruppen "positivt og aktivt engasjement i livet" er det eneste måleinstrumentet som viste negativ endring etter kurs og det instrumentet som hadde lavest SRM-verdi. Undergruppen omhandler engasjement i livet og består av fem påstander; "de fleste dagene gjør jeg noen av tingene jeg virkelig liker", "jeg prøver virkelig å få mest mulig ut av livet", "jeg gjør interessante ting i livet mitt", "jeg er aktivt engasjert i livet" og "jeg har planer om å gjøre ting jeg liker i løpet av de neste dagene" Det er vanskelig å vite om informasjonen som ble gitt i kursene omfatter "engasjement i livet", se tabell 2. Om man antar at kursene i større eller mindre grad omfatter det å opprettholde engasjement i livet med en inflammatorisk revmatisk sykdom, er det likevel ikke sikkert respondentene relaterer dette til påstandene ovenfor som omhandler engasjement til *livet*. Engasjement i livet trenger heller ikke å være knyttet til den revmatiske sykdommen. Komorbiditet som depresjon, eller hendelser i livet som for eksempel medfører tap og sorg kan ha en større innflytelse på dette aspektet. Da vil et pasientopplæringskurs for inflammatoriske revmatiske sykdommer ikke nødvendigvis være nok til å medføre en endring i denne undergruppen av heiQ. Det samme kan gjelde for undergruppen "sosial integrasjon og støtte" som handler om å ha nok

pårørende som forstår og tar ivaretar pasienten ved sykdom, forståelse fra helsepersonell og muligheten til å snakke om helseproblemer med noen som forstår. Dette er ikke et tema som det direkte undervises i, og det er lite sannsynlig at et kurs i seg selv kan gjøre noe med antallet pårørende pasienten har. Hvis pårørende deltar på kursene kan kurset derimot øke deres forståelse for pasienten, sykdommen og hvordan de kan bidra til støtte i vanskelige perioder. Dette fordrer at venner og pårørende deltar på kurset. Pårørende kunne delta på 11 av 13 kurs i denne studien, men forfatter har ingen oversikt over antall deltakende pårørende som faktisk deltok. Kursene kan også ha bidratt til at flere deltakere kom i snakk med hverandre, noe som kan ha åpnet for muligheter til å bli kjent med personer som har liknende erfaringer. Dette kan igjen ha økt opplevelsen av sosial integrasjon og støtte. Undergruppen har en SRM tilsvarende liten effekt (0.38) og en p-verdi som nærmer seg signifikant (0.06). Fordi det kun var 28 respondenter i sB etter tre måneder kan utvalgsstørrelsen ha vært for liten til å finne statistisk signifikant endring og medført sjanse for type 2-feil. Dette kan også ha vært tilfellet i undergruppen "egenovervåking og innsikt". Undergruppen inneholder påstander som tilhører temaer alle kursene underviser i; "medikamenter" og "å leve med sykdom", men resulterte i ikke-signifikante endringer. Igjen er det vanskelig å vite om undervisningen fanget opp innholdet i påstandene om regelmessige legebesøk, kunnskap om hva som utløser helseproblemene og hvorfor medisiner er viktig. Fordi undergruppen har en SRM verdi tilsvarende liten effekt (0.39) og en p-verdi som nærmer seg signifikant (0.058), kan det likevel ha vært en type 2-feil som medførte ikke-signifikante endringer.

Ved sammenlikning av måleinstrumentene med statistisk signifikante endringer, finnes likhetstrekk mellom EC-17 og heiQ undergruppene "tilegning av ferdigheter og teknikker" og "orientering i helsetjenesten". Både "tilegning av ferdigheter og teknikker" og EC-17 undersøker ferdigheter i å ta vare på egen helse. For eksempel har heiQ påstanden "jeg vet godt hvordan jeg kan håndtere helseproblemene mine", og EC-17 har spørsmålet "jeg kan vurdere fordeler og ulemper vedrørende avgjørelser om min sykdom". Flere av påstandene i heiQ undergruppen "orientering i helsetjenesten" og spørsmålene i EC-17 omhandler hvordan bruke helsevesenet for å dekke helsebehov og kommunikasjon og tillitt mellom pasient og helsepersonell. Funn av signifikante endringer i måleinstrumenter med likhetstrekk i sA og sB samsvarer med at kursene i de to ulike gruppene har forholdsvis likt innhold, se tabell 2. EC-17 og PAM har også likhetstrekk, og i en studie er de funnet å korrelere moderat med

hverandre (31). Det ville derfor vært naturlig å anta statistiske signifikante endringer også i PAM. Det kan være ulike årsaker til at dette ikke er tilfellet. En mulig årsak er forekomsten av takeffekter. I følge en oversiktsartikkel av PAM understrekes det at pasienter med lavest skår ved baseline har en tendens til å øke sin skår mest (45). Takeffekter ved baseline kan derfor ha svekket muligheten for øking av skår etter endt kurs i flere av spørsmålene og resultert i en statistisk ikke-signifikant endring (45). PAM har også en lav SRM på like over 0.2 (0.22). I denne studien er det tre av de fire måleinstrumentene med lavest forekomst av takeffekter som også har statistisk signifikante endringer og høyest SRM tilsvarende "liten effekt" ved analyse av før- og etter-skår. Dette gjelder EC-17, heiQ "emosjonelt velvære", "tilegning av ferdigheter og teknikker" og "orientering i helsetjenesten". Dette kan antyde en sammenheng mellom forekomst av takeffekter og evne til å fange opp endringer. På den andre siden er endringene små og det kan tilfældigheter som gjør at akkurat de instrumentene som har statistisk signifikante endringer også har lavest forekomst av takeffekter.

I tillegg til at måleinstrumenter og kursinnhold må samsvare, er det flere andre faktorer som kan påvirke endringsskår og true den indre validiteten. En av faktorene er instrumentering. Alle måleinstrumentene er validert i flere ulike land med gode resultater, noe som reduserer sjansen for instrumentering. Det er likevel funnet store prosentandeler takeffekter i PAM som kan være et tegn på instrumentering. En annen faktor er "historie". Historie er hendelser som kan oppstå i respondentens liv fra baseline til tre måneder etter kurs og påvirke effekten (62). Man kan redusere denne trusselen ved å inkludere et "anker" i spørreskjemaet. Dette kan være et spørsmål som gir et subjektivt svar på kursets betydning for respondenten. Spørreskjemaene inneholder ikke et slikt anker, og man kan ikke avskrive at historie kan ha påvirket resultatene. Fordi respondentene ikke ble tilfeldig fordelt over sA og sB kan historie ha hatt ulik påvirkning på gruppene og ført til ulike forutsetninger for før- og etterskår. Jo mer ulike gruppene er før kurs, desto større sannsynlighet for at historie kan forekomme (62). Det er noen prosentvise forskjeller på kjønn, alder og utdanningsnivå i sA og sB, noe som kan ha innvirket på utbyttet av kursene. Dette betyr at forutsetningene for å finne endringer kan ha vært ulik i de to gruppene. Dette er vanskelig å undersøke blant annet fordi sA og sB inneholder ulike måleinstrumenter, med unntak av ASES. Det er utført analyser av ASES med respondenter fra sA og sB isolert fra hverandre. Dette resulterte i ikke-signifikante funn i begge grupper. Dette trenger ikke nødvendigvis å ha noen betydning da det, som nevnt

ovenfor, kan være andre årsaker til ikke-signifikante funn i ASES. Det er i tillegg mer frafall i sB etter tre måneder. Mens 51 respondenter (81%) besvarte sA etter tre måneder, var det bare 27 (66%) i sB. SB har dermed både færre respondenter og større frafall. Dette reduserer statistisk styrke, svekker intern validitet og øker sjansen for type 2-feil. Denne mistanken styrkes når man ser på endringskår i heiQ "egenovervåking og innsikt" og "sosial integrasjon og støtte". Begge undergruppene har ikke-signifikante endringer men SRM tilsvarende liten effekt (0.39 og 0.38). Når man studerer p-verdien nærmer den seg nivå av signifikans og er på henholdsvis 0.058 og 0.06.

6.2 Metodiske styrker og svakheter

Formålet med denne studien har vært å undersøke anvendeligheten til ASES smerter og andre symptomer, EC-17, heiQ og PAM for bruk til evaluering av pasientopplæringskurs i en fremtidig hovedstudie. Til dette formålet er anvendelighetsstudie et egnet design. Resultatene i denne studien er i hovedsak gjort med deskriptive analyser, noe som er vanlig i anvendelighetsstudier. Fordi en av problemstillingene ser på før- og etter-skår har det vært nødvendig med et før- og etter design. En begrensning ved anvendelighetsstudier er at det ikke utføres styrkeberegninger, og det er heller ikke gjort i denne studien. Dette kan redusere validiteten til resultatene av før- og etter-skår da det ikke kan utelukkes at utvalget er for lite til å oppdage endring før og etter kurs (type-2 feil). Anvendelighetsstudier er likevel betydningsfulle da de bidrar til å sikre at hovedstudien er praktisk gjennomførbar og reduserer trusler mot gyldigheten av hovedstudiens resultater (21).

Ekstern validitet gjelder generaliserbarhet av kausale slutninger og angår i hovedsak analysen av før- og etter-skår. For å kunne generalisere resultatene må utvalget representere den aktuelle populasjonen (65), som i denne studien er pasienter med inflammatoriske revmatiske sykdommer som deltar på pasientopplæringskurs. Vi kunne ikke randomisere deltakerne, og det var heller ikke ønskelig i dette designet. Vi kunne derimot fordelt sA og sB tilfeldig på kursene ved å "trekke" hvilket kurs som skulle få hvilket skjema. På denne måten kunne vi ha sikret et jevnere antall respondenter i hver gruppe og redusert sjansen for skjevhet mellom gruppene. Utvalget i studien er hentet fra seks ulike sykehus. Sykehusene er fordelt utover

Norge og representerer både storby og småby, lokalsykehus og universitetssykehus. Dette mangfoldet bidrar til å styrke studiens eksterne validitet.

Seleksjonen av pasienter var basert på frivillighet. Deltakelse i studien kunne for eksempel oppfattes som en mulighet til å bidra til forskning og utvikling av dagens pasientopplæringskurs, eller det kunne sees på som lite interessant og unødvendig bruk av tid. Det er i forskermiljøet anerkjent at personer som frivillig deltar i forskningsstudier kan være forskjellig fra de som ikke gjør det. Spesielt gjelder dette utdanningsnivå der deltakere i studier ofte har høyere utdanning enn de som ikke deltar (66). Lav studiedeltakelse kan medføre systematisk skjevhet og svekke den eksterne validiteten. Fordi vi ikke har informasjon om ikke-deltakende pasienter, kan ikke systematisk over- eller underrepresentasjon av bakgrunnsvariablene undersøkes. Responsstatistikken viser at 69% (n=104) av de 151 pasientene som deltok på pasientopplæringskurs ble inkludert i studien. En svarprosent over 60 regnes som god (67). Utvalgets demografiske variabler er også sammenliknbare med andre studier av samme populasjon. Studiens utvalg er veldig lik i alder, kjønnsfordeling og utdanningsnivå sammenliknet med en annen norsk studie av pasientundervisning (68). Dette styrker den eksterne validiteten. Oppsummert betyr dette at pasientutvalget representerer en selektert gruppe pasienter med inflammatoriske revmatiske sykdommer og resultatene kan derfor ikke generaliseres til utenfor studiens utvalg. Det kan heller ikke trekkes kausale slutninger da dette er en før- og etterstudie og ikke en randomisert kontrollert studie. På den andre side påvirker dette studien i mindre grad da interessen ligger i anvendeligheten til måleinstrumentene og ikke effekten av pasientopplæringskursene.

6.3 Teoretiske implikasjoner av funn

Denne studien bidrar til å belyse anvendeligheten til ASES smerter og andre symptomer, EC-17, heiQ og PAM. Styrker og begrensninger ved måleinstrumentene er identifisert, og resultatene kan bidra til et mer evidensbasert valg av måleinstrument til en effektstudie av pasientopplæringskurs.

7 Konklusjon og anbefaling av instrumentvalg

De pasientrapporterte spørreskjemaene ASES smerter og andre symptomer, EC-17, heiQ og PAM har nå gjennomgått en anvendelighetsstudie der manglende data, gulv- og takeffekter, intern validitet og før- og etterskår er undersøkt. Resultatene danner grunnlaget for anbefalinger om valg av måleinstrument(er) for en effektstudie av pasientopplæringskurs for pasienter med inflammatoriske revmatiske sykdommer. Alle måleinstrumentene hadde lav forekomst av manglende data under 5%. EC-17 og heiQ undergruppene "emosjonelt velvære", "tilegning av ferdigheter og teknikker" og "orientering i helsetjenesten" fremtrer som de mest anvendelige måleinstrumentene i forhold til gulv- og takeffekter, intern konsistens og før- og etterskår. På bakgrunn av dette anbefales disse måleinstrumentene til bruk i en hovedstudie. EC-17 og heiQ undergruppene "tilegning av ferdigheter og teknikker" og "orientering i helsetjenesten" har flere likheter i innhold. Det kan derfor være aktuelt å velge enten EC-17 eller de to undergruppene i heiQ. For å dekke flest mulig av aspektene identifisert i POPOP revma (figur 1), vil EC-17 være et naturlig valg da det dekker fire aspekter. EC-17 er det måleinstrumentet som er benyttet i færrest studier og derav minst undersøkt med hensyn til validitet og reliabilitet. Behovet for videre validering av instrumentet må derfor vurderes. Hvis det er ønskelig å dekke alle de seks aspektene, bør ASES smerter og andre symptomer også vurderes som et aktuelt måleinstrument, fordi det er det eneste instrumentet som måler aspektet "mestringstro". ASES viser tilfredsstillende resultater i analyser av manglende data, takeffekter og Cronbachs alfa i begge undergrupper, men resulterte i ikke-signifikante endringer og lave SRM-verdier. Det kan være ulike årsaker til dette, men en av mulighetene er at kursenes innhold ikke reflekterer innholdet i ASES i stor nok grad. Det er viktig at valg av måleinstrumenter vurderes i forhold til innholdet i kurset/kursene som skal forskes på. Til slutt poengteres det at forskjeller i sA og sB kan ha medført ulike forutsetninger for funn av endringer, og at utvalget i hver av gruppene kan ha vært for lite til å fange opp små endringer. Studiens eksterne validitet medfører også at resultatene ikke kan generaliseres, men resultatene danner et grunnlag til å teste instrumentene videre i en randomisert kontrollert studie.

Litteraturliste

1. Combe B, Landewe R, Lukas C, Bolosiu HD, Breedveld F, Dougados M, et al. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis.* 2007;66(1):34-45.
2. Zangi HA, Ndosu M, Adams J, Andersen L, Bode C, Bostrom C, et al. EULAR recommendations for patient education for people with inflammatory arthritis. *Ann Rheum Dis.* 2015;74(6):954-62.
3. Zochling J, van der Heijde D, Burgos-Vargas R, Collantes E, Davis JC, Jr., Dijkmans B, et al. ASAS/EULAR recommendations for the management of ankylosing spondylitis. *Ann Rheum Dis.* 2006;65(4):442-52.
4. Niedermann K, Fransen J, Knols R, Uebelhart D. Gap between short- and long-term effects of patient education in rheumatoid arthritis patients: a systematic review. *Arthritis Rheum.* 2004;51(3):388-98.
5. Riemsma RP, Taal E, Kirwan JR, Rasker JJ. Systematic review of rheumatoid arthritis patient education. *Arthritis Rheum.* 2004;51(6):1045-59.
6. Christie A, Jamtvedt G, Dahm KT, Moe RH, Haavardsholm EA, Hagen KB. Effectiveness of nonpharmacological and nonsurgical interventions for patients with rheumatoid arthritis: an overview of systematic reviews. *Phys Ther.* 2007;87(12):1697-715.
7. Zangi HA. Prosjektbeskrivelse Pasient- og pårørende-opplæring i revmatologi (POPOPP revma). Prosjektskisse. Upublisert: NKRR Diakonhjemmet sykehus; 2014.
8. Helse- og omsorgsdepartementet. Ot.prp. nr. 10 (1998-99). Om lov om spesialisthelsetjenesten m. m. Oslo: Helse og omsorgsdepartementet; 1998.
9. Gran JT. Kompendium for assistentleger i revmatologi. 8. ed. Oslo: Revmatologisk avdeling, Rikshospitalet; 2009.
10. Brage S, Ihlebæk C, Natvig B, Bruusgaard D. Muskel- og skjelettlidelser som årsak til sykefravær og uføreytelser. *Tidsskr Nor Lægeforen.* 2010;23(130):2369-70.
11. Spesialisthelsetjenesteloven. Lov om spesialisthelsetjenesten m.m. 1999-07-02-61.
12. Helsedirektoratet. Spesialisthelsetjenesteloven med kommentarer. Rundskriv IS-5/2013. 2013.

13. NOU 2005:5. Fra stykkevis til helt. En sammenhengende helsetjeneste. Oslo: Statens forvaltningstjeneste; 2005.
14. Lorig K. Patient education: a practical approach. 3rd ed. Thousand Oaks, Calif.: Sage; 2001. 246 p.
15. Helse Sør-Øst. Regional strategi for pasient- og pårørendeopplæring 2013-2016. 2013.
16. Zangi HA, Hauge MI, Steen E, Finset A, Hagen KB. "I am not only a disease, I am so much more". Patients with rheumatic diseases' experiences of an emotion-focused group intervention Patient Educ Couns. 2011;85(3):419-24.
17. Hamnes B, Hauge MI, Kjekken I, Hagen KB. 'I have come here to learn how to cope with my illness, not to be cured': a qualitative study of patient expectations prior to a one-week self-management programme. Musculoskeletal Care. 2011;9(4):200-10.
18. Austvoll-Dahlgren A, Nøstberg AM, Steinsbekk A. Effekt av gruppeundervisning i pasient- og pårørendeopplæring. Oslo: Nasjonalt kunnskapssenter for helsetjenesten; 2011. Report No.: 575.
19. Saunes I, Deilkås E, Kristoffersen M, Lyngstad M, Nyen B. Formaliserte konsensusprosesser, veiledning ved vanskelige valg? Oslo: Nasjonalt kunnskapssenter for helsetjenesten; 2009. Report No.: 912.
20. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol. 2010;10:67.
21. Tickle-Degnen L. Nuts and bolts of conducting feasibility studies. Am J Occup Ther. 2013;67(2):171-6.
22. National Institute for Health Research. Feasibility and Pilot studies Storbritannia: National Institute for Health Research (GB); [Available from: https://www.nihr.ac.uk/funding-and-support/documents/funding-for-research-studies/research-programmes/RfPB/FAQs/Feasibility_and_pilot_studies.pdf]
23. Lorig K, Chastain RL, Ung E, S. S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis Rheum. 1989;32(1):37-44.
24. Garratt AM, Lochting I, Smedslund G, Hagen KB. Measurement properties of instruments assessing self-efficacy in patients with rheumatic diseases. Rheumatology (Oxford). 2014;53(7):1161-71.
25. Garratt AM, Klokke M, Lochting I, Hagen KB. Rasch analysis of the Norwegian version of the Arthritis Self-Efficacy Scale (ASES). Scand J Rheumatol. 2017;46(1):33-9.

26. Lomi C. Evaluation of a Swedish version of the arthritis self-efficacy for individuals with arthritis. *Scand J Caring Sci.* 1992;20:177-87.
27. Barlow JH, Williams B, Wright CC. The reliability and validity of the arthritis self-efficacy scale in a UK context. *Psychol Health Med.* 1997;2(1):3-17.
28. Lomi C, Burckhardt C, Nordholm L, Bjelle A, Ekdahl C. Evaluation of a Swedish version of the arthritis self-efficacy scales in people with fibromyalgia. *Scand J Rheum.* 1995;24(5):282-87.
29. Kristjansson E, Tugwell PS, Wilson AJ, Brooks PM, Driedger SM, Gallois C, et al. Development of the Effective Musculoskeletal Consumer Scale. *J Rheumatol.* 2007;34(6):1392-400.
30. Hamnes B, Garratt A, Kjekken I, Kristjansson E, Hagen KB. Translation, data quality, reliability, validity and responsiveness of the Norwegian version of the Effective Musculoskeletal Consumer Scale (EC-17). *BMC Musculoskelet Disord.* 2010;11:21.
31. Santesso N, Rader T, Wells GA, O'Connor AM, Brooks PM, Driedger M, et al. Responsiveness of the Effective Consumer Scale (EC-17). *J Rheumatol.* 2009;36(9):2087-91.
32. Bremander A, Wikstrom I, Larsson I, Bengtsson M, Hagel S, Strombeck B. Cultural adaptation, validity, reliability and responsiveness of the Swedish version of the effective musculoskeletal consumer scale (EC-17). *Musculoskeletal Care.* 2012;10(1):43-50.
33. ten Klooster PM, Taal E, Siemons L, Oostveen JC, Harmsen EJ, Tugwell PS, et al. Translation and validation of the Dutch version of the Effective Consumer Scale (EC-17). *Qual Life Res.* 2013;22(2):423-9.
34. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. *Patient Educ Couns.* 2007;66(2):192-201.
35. Wahl AK, Osborne RH, Langeland E, Wentzel-Larsen T, Mengshoel AM, Ribu L, et al. Making robust decisions about the impact of health education programs: Psychometric evaluation of the Health Education Impact Questionnaire (heiQ) in diverse patient groups in Norway. *Patient Educ Couns.* 2016;99(10):1733-8.
36. Deakin University. Scoring the Health Education Impact Questionnaire (heiQ) – Version 3 (upublisert).
37. Ammerlaan JW, van Os-Medendorp H, Sont JK, Elsworth GR, Osborne RH. Validation of the dutch version of the health education impact questionnaire (HEIQ) and comparison of the

Dutch translation with the English, German and French HEIQ. *Health Qual Life Outcomes*. 2017;15(1):28.

38. Belanger A, Hudon C, Fortin M, Amirall J, Bouhali T, Chouinard MC. Validation of a French-language version of the health education impact Questionnaire (heiQ) among chronic disease patients seen in primary care: a cross-sectional study. *Health Qual Life Outcomes*. 2015;13:64.

39. Schuler M, Musekamp G, Faller H, Ehlebrach-König I, Gutenbrunner C, Kirchhof R, et al. Assessment of proximal outcomes of self-management programs: translation and psychometric evaluation of a German version of the Health Education Impact Questionnaire (heiQ™). *Qual Life Res*. 2013;22:1391-403.

40. Elsworth GR, Nolte S, Osborne RH. Factor structure and measurement invariance of the Health Education Impact Questionnaire: Does the subjectivity of the response perspective threaten the contextual validity of inferences? *SAGE Open Med*. 2015;3:1-13.

41. Larsen MH, Krogstad AL, Aas E, Moum T, Wahl AK. A telephone-based motivational interviewing intervention has positive effects on psoriasis severity and self-management: a randomized controlled trial. *Br J Dermatol*. 2014;171(6):1458–69.

42. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM)- Conceptualizing and Measuring Activation in Patients and Consumers. *Health Serv Res*. 2004;39(4 Pt 1):1005-26.

43. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res*. 2005;40(6 Pt 1):1918-30.

44. Insignia Health Inc. Patient Activation Measure (PAM) 13™ License Materials. 2010.

45. Hibbard JH, Gilbert H. Supporting people to manage their health. An introduction to patient activation. London; 2014.

46. Steinsbekk A. Måling av effekt av pasientopplæring. *Tidsskr Nor Lægeforen*. 2008;128:2316-318.

47. Graffigna G, Barello S, Bonanomi A, Lozza E, Hibbard J. Measuring patient activation in Italy: Translation, adaptation and validation of the Italian version of the patient activation measure 13 (PAM13-I). *BMC Med Inform Decis Mak*. 2015;15:109.

48. 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.

49. Rademakers J, Nijman J, van der Hoek L, Heijmans M, Rijken M. Measuring patient activation in the Netherlands: translation and validation of the American short form Patient Activation Measure (PAM13). *BMC Public Health*. 2012;12(1).
50. Moljord IE, Lara-Cabrera ML, Perestelo-Perez L, Rivero-Santana A, Eriksen L, Linaker OM. Psychometric properties of the Patient Activation Measure-13 among out-patients waiting for mental health treatment: A validation study in Norway. *Patient Educ Couns*. 2015;98(11):1410-7.
51. Zill JM, Dwinger S, Kriston L, Rohenkohl A, Härter M, Dirmaier J. Psychometric evaluation of the German version of the patient activation measure (PAM13). *BMC Public Health*. 2013;13(1).
52. Rademakers J, Maindal HT, Steinsbekk A, Gensichen J, Brenk-Franz K, Hendriks M. Patient activation in Europe: an international comparison of psychometric properties and patients' scores on the short form Patient Activation Measure (PAM-13). *BMC Health Serv Res*. 2016;16(1):570.
53. Grønning K, Skomsvoll JF, Rannestad T, Steinsbekk A. The effect of an educational programme consisting of group and individual arthritis education for patients with polyarthritis - A randomised controlled trial. *Patient Educ Couns*. 2012;88(1):113-20.
54. Svensson E, Hjartåker A, Laake P. Hva skal måles og hvordan? In: Laake P, Hjartåker A, Thelle DS, Veierød MB, editors. *Epidemiologiske og kliniske forskningsmetoder*. 1. ed. Oslo: Gyldendal Norsk Forlag AS; 2007.
55. Dong Y, Peng CJP. *Principled missing data methods for researchers*. SpringerPlus. 2013;2(222):1-17.
56. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
57. Clausen TH, Johansen V. Chronbachs alfa. In: Eikemo TA, Clausen TH, editors. *Kvantitativ analyse med SPSS En praktisk innføring i kvantitative analyseteknikker*. 2. ed. Trondheim: Tapir Akademisk Forlag; 2007.
58. Palliant J. *SPSS survival manual: a step by step guide to data analysis using SPSS*. XIV ed. Maidenhead: McGraw-Hill; 2010.
59. Middel B, Stewart R, Bouma J, van Sonderen E, van den Heucel WJA. How to validate clinically important change in health-related functional status. Is the magnitude of the effect

size consistently related to magnitude of change as indicated by a global question rating? *J Eval Clin Pract.* 2001;7(4):399-410.

60. Cohen J. *Statistical power analysis for the behavioral sciences* 2nd. ed. ed. Hillsdale, N. J. : Lawrence Erlbaum 1988.

61. Bartman BA. Health Status Measurement, Floor and Ceiling Effects. In: Kattan MW, editor. *Encyclopedia of Medical Decision Making*. Thousand Oaks: SAGE Publications, Inc.; 2009.

62. Lund T. Metodologiske prinsipper og referanserammer. In: Eikemo TA, Clausen TH, editors. *Kvantitativ analyse med SPSS En praktisk innføring i kvantitative analyseteknikker*. 2. ed. Trondheim: Tapir Akademisk Forlag; 2007.

63. Hammond A, Bryan J, Hardy A. Effects of a modular behavioural arthritis education programme: a pragmatic parallel-group randomized controlled trial. *Rheumatology (Oxford)*. 2008;47(11):1712-8.

64. Bandura A. *Self-efficacy : the exercise of control*. New York, NY: Freeman; 1997.

65. Lund T. Generaliseringsproblematikk. In: Eikemo TA, Clausen TH, editors. *Kvantitativ analyse med SPSS En praktisk innføring i kvantitative analyseteknikker*. 2. ed. Trondheim: Tapir Akademisk Forlag; 2007.

66. Rosenthal R, Rosnow RI. *The volunteer subject*. New York: Wiley & Sons 1975.

67. Jacobsen DI. *Hvordan gjennomføre undersøkelser? Innføring i vitenskapelig metode*. 3 ed. Oslo: Cappelen Damm AS; 2015.

68. Grønning K, Rannestad T, Skomsvoll JF, Rygg LO, Steinsbekk A. Long-term effects of a nurse-led group and individual patient education programme for patients with chronic inflammatory polyarthritis - a randomised controlled trial. *J Clin Nurs.* 2014;23(7-8):1005-17.

Vedlegg

Vedlegg 1: Poster fremvist på EULAR-kongressen juni 2017

Vedlegg 2: De ni måleinstrumentene som ble inkludert i Delphi-prosessen

Vedlegg 3: Bakgrunnsvariabler

Vedlegg 4: Arthritis Self-Efficacy Scale Andre symptomer

Vedlegg 5: Arthritis Self-Efficacy Scale Smerter

Vedlegg 6: Effective Consumer Scale

Vedlegg 7: The Health Education Impact Questionnaire

Vedlegg 8: Patient Activation Measure

Vedlegg 9: Informasjonsskriv

Vedlegg 10: Samtykkeerklæring

Vedlegg 11: Svar på fremleggsvurdering fra Regional etisk komité

Vedlegg 12: Svar på meldeskjema fra Personvernombudet