

Vejledning for implementering og anvendelse af patient rapporterede outcome data i kliniske kvalitetsdatabaser

Juni 2015. Opdateret September 2015

Indhold

1. Baggrund for projektet	4
2. Formål med projektet	5
3. Projektgruppen.....	6
4. Definition og anvendelse af patient rapporterede outcome data.....	7
4.1. Definition af patient rapporterede outcome data.....	7
4.2. Anvendelse af PRO-data	8
4.3. Er der evidens for at brug af PRO-data medfører bedre kvalitet?.....	11
4.4. Definition af PRO-instrumenter.....	11
4.4. Definition af PRO-baserede kvalitetsindikatorer	13
5. Datatilladelser og datasikkerhed.....	14
6. Erfaringer med PRO-data i de kliniske kvalitetsdatabaser	15
6.1. Danmark	15
6.2. Udland	18
7. Hvilken type PRO-instrument databasen skal vælge	20
7.1. Generiske PRO-instrumenter	21
7.2. Domænespecifikke PRO-instrumenter	22
7.3. Sygdomsspecifikke PRO-instrumenter	23
7.4. Ad hoc spørgsmål.....	23
8. Valg af eksisterende eller udvikling af nye PRO-instrumenter	25
8.1. Eksisterende PRO-instrumenter	25
8.2. Oversættelse af eksisterende PRO-instrumenter	25
8.3. Udvikling af nye PRO-instrumenter.....	26
9. Krav til egenskaber ved PRO-instrument.....	28
9.1. Validitet	28
9.2. Reliabilitet (pålidelighed)	29
9.3. Responsiveness (følsomhed).....	29
9.4. Floor-/ceiling effect (gulv- og lofteffekt)	30
9.5. Acceptability.....	30
10. Definition af inklusions-/eksklusionskriterier før indsamling af PRO-data	31
11. Layout for spørgeskema og generel vejledning til patienterne	32
12. Pilotfase (field test)	33
13. Hvornår skal PRO indsamles?	33
14. Kontakt- og besvarelsesmetoder	36

14.1 PRO-besvarelsesmetoder	36
14.2 Kontaktmetode	39
14.3. Rykkerprocedurer	41
14.4. Responstrater ved internet- og papirbaseret besvarelsesmetode	43
15. Infrastruktur for dataindsamling	45
16. Analyse og fortolkning af PRO-data.....	46
16.1 I tilfælde hvor kun en måling af PRO-data er tilgængelig.....	47
16.2. I tilfælde hvor to eller flere målinger af PRO-data er tilgængelige.....	47
16.3. Valg af referencegruppe	48
16.4. Justering for case-mix.....	48
16.5. Mangelfulde (missing) PRO-data.....	49
16.6. “Minimal clinically important improvement” (MCII).....	51
17. Præsentation af PRO-data	52
18. Dokumentation	59
19. Konklusioner og anbefalinger.....	60
20. Referencer	63

1. Baggrund for projektet

Der foregår indenfor sundhedsvæsenet en stor og mangestrengt indsats for at monitorere og forbedre den sundhedsfaglige kvalitet (1). Et meget centralt element i denne indsats i Danmark såvel som internationalt har været systematisk monitorering af behandlingskvaliteten for udvalgte patientgrupper via kvalitetsindikatorer.

Kvalitetsindikatorer afspejler enten den behandling som patienter modtager (procesindikatorer) eller udfaldet af deres sygdomsforløb (resultatindikatorer). Traditionelt har resultatindikatorer som patientoverlevelse, reoperation, komplikationer eller genindlæggelse spillet en meget stor rolle, særligt inden for det kirurgiske område. Der er dog to grundlæggende problemer med disse resultatindikatorer:

For det første er indikatorerne i vid udstrækning baseret på indberetninger fra sundhedsprofessionelle. Dette kan være problematisk, idet der ikke nødvendigvis er overensstemmelse mellem sundhedspersonalets og patientens vurdering af kvaliteten. For eksempel kan en operation, hvor der bliver indsat en kunstig hofteprotese, af kirurgen blive vurderet som en succes, hvis patienten ikke efterfølgende får udført en reoperation i den samme hofte. Patienten kan dog på samme tid alligevel godt have vedvarende smerter og nedsat gangfunktion efter operationen. Fra kirurgens synsvinkel kan operationen således umiddelbart fremstå som vellykket, mens patienten muligvis vil have en anden opfattelse. Der kan således være uoverensstemmelse mellem patientens og sundhedspersonalets præferencer i forhold til at vurdere sygdomsudfald.

Patienternes præferencer kan endvidere også adskille sig fra baggrundsbefolkningens præferencer. Således vil alvorligt syge patienter i højere grad være villige til at acceptere f.eks. kemoterapi og medfølgende bivirkninger sammenlignet med deres pårørende eller personer fra baggrundsbefolkningen (2).

For det andet, er resultatindikatorer som f.eks. overlevelse ofte ikke særligt sensitive kvalitetsmål og derfor vanskelige at anvende til at vurdere kvaliteten af behandlingen både for den enkelte patient og for en hel afdeling. Således kan afdeling A og B have lige høj overlevelse efter en indsættelse af knæprotese, men patienter fra afdeling A er mere tilfredse med behandlingen, de har færre smerter efter operation, de har bedre funktionsniveau, de har bedre livskvalitet, og de kommer hurtigere og i højere grad tilbage på arbejdsmarkedet end patienter fra afdeling B. I dette tilfælde vil fokus alene på patientoverlevelse derfor ikke give et retvisende billede af behandlingskvaliteten på de to afdelinger.

Det er derfor vigtigt at patientperspektivet bliver inddraget i forbindelse med vurdering og udvikling af behandlingskvalitet. Patient rapporterede outcome data, kan bruges som supplement til kliniske informationer om patienten.

Der er stigende interesse for patient rapporterede outcome data fra en bred vifte af aktører:

- Patienter og patientorganisationer, der ønsker at træffe informerede valg om deres egen behandling.
- Ledelse og myndigheder, der ønsker at træffe beslutninger om sundhedspolitik og ressourceallokering.
- Forskere, klinikere og fagpersoner, som ønsker at yde den bedste behandling til den individuelle patient.

På trods heraf inddrages patient rapporterede outcome data aktuelt endnu kun i meget begrænset omfang systematisk i tilrettelæggelsen af sundhedsvæsenets ydelser og i monitoreringen af kvaliteten. Den afgørende barriere i forhold til inddragelse af patient rapporterede outcome data i klinisk praksis og i de landsdækkende kvalitetsdatabaser er de mange praktiske udfordringer (vedr. indsamling af data og betydning af data for de enkelte sundhedsfaglige personer) og metodologiske udfordringer (vedr. analyse og fortolkning af data samt anvendelse af data til at ændre behandling af patienter) (3). Disse udfordringer er emnet for denne rapport.

Målgruppen for denne rapport er primært sundhedsprofessionelle inklusiv styregruppen for Regionernes Kliniske Kvalitetsudviklingsprogram (RKKP), medlemmer af de kliniske kvalitetsdatabasers styregrupper samt andre med interesse for patient rapporterede outcome data.

2. Formål med projektet

Formålet med projektet er at beskrive et fælles grundlag for implementering af patient rapporterede outcome-dataindsamling og -analyse i de nationale kliniske kvalitetsdatabaser. Herunder vil gruppen beskrive de forskellige forudsætninger, barrierer og udfordringer som kliniske kvalitetsdatabaser står overfor i forhold til indsamling og anvendelse af patient rapporterede outcome data.

Projektet søger at beskrive og besvare følgende specifikke spørgsmål og problemstillinger:

1. Beskrive nuværende evidens for anvendelse af patient rapporterede outcome data, især i forhold til kvalitetsarbejde
2. Definere patient rapporterede outcome data, patient rapporterede outcome-instrumenter og patient rapporterede outcome-baserede kvalitetsindikatorer
3. Beskrive nationale og internationale erfaringer vedr. implementering af patient rapporterede outcome data i de kliniske kvalitetsdatabaser, især som resultatmål
4. Datatilladelse og datasikkerhed
5. Databasens valg af patient rapporterede outcome-instrumenter
6. Valg af eksisterende eller udvikling af nye patient rapporterede outcome-instrumenter
7. Oversættelse af eksisterende patient rapporterede outcome-instrumenter
8. Krav til egenskaber ved patient rapporterede outcome-instrumenter
9. Layout for spørgeskema
10. Pilotfase

11. Hvordan skal patient rapporterede outcome data indsamles?
12. Hvornår skal patient rapporterede outcome data indsamles?
13. Kontakt- og besvarelsesmetoder
14. Infrastruktur for dataindsamling
15. Hvad er en acceptabel svarprocent?
16. Håndtering af missing patient rapporterede outcome data
17. Hvordan skal patient rapporterede outcome data analyseres for at sammenligne patienter og afdelinger?
18. Udvikling af case-mix justeringsmodeller til sammenligning af afdelinger
19. Præsentation af patient rapporterede outcome data

3. Projektgruppen

Rapporten er udarbejdet i et samarbejde mellem følgende institutioner:

- Kompetencecenter for Epidemiologi og Biostatistik (KCEB)-Nord, Klinisk Epidemiologisk Afdeling, Aarhus Universitetshospital
- VestKronik, Arbejdsmedicinsk Klinik, Hospitalsenhed Vest, Regionshospitalet Herning
- Kompetencecentre for Klinisk Kvalitet & Sundhedsinformatik (KCKS)-Vest

Projektgruppens medlemmer:

Afdelingslæge, ph.d. Alma B. Pedersen (projektleder), seniorforsker, MPH, ph.d. Ellen M. Mikkelsen
forskningsoverlæge, ph.d. Søren P. Johnsen, overlæge, ph.d. Niels H. Hjøllund, cand.scient.san Anne Jessen,
kvalitetskonsulent, cand.scient.san. Anne Nakano.

4. Definition og anvendelse af patient rapporterede outcome data

4.1. Definition af patient rapporterede outcome data

Patient rapporterede data kan omhandle mange aspekter og forhold, der er relevante for såvel patienter som personale i sundhedssektoren bredt set. Patient rapporterede data dækker som minimum over:

- Oplevelser og tilfredshed: Mål for patienters oplevelser og tilfredshed med sundhedsvæsenet, kaldet PREM (*Patient Reported Experience Measures*), som det fx ses i den Landsdækkende Undersøgelse af Patientoplevelser (LUP). (Enheden for Evalueringen og Brugerinddragelse, 2015)
- Sundhedsadfærd: Patienters oplysninger om deres sundhedsadfærd, som det f.eks. ses i de Nationale Sundhedsprofiler (Sundhedsstyrelsen, 2014)
- Patient Reported Outcome (PRO), som omfatter patient rapporterede data til klinisk anvendelse og til kvalitetsudvikling.

The US Food and Drug Administration (FDA) definerer PRO som:

”any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else” (4)

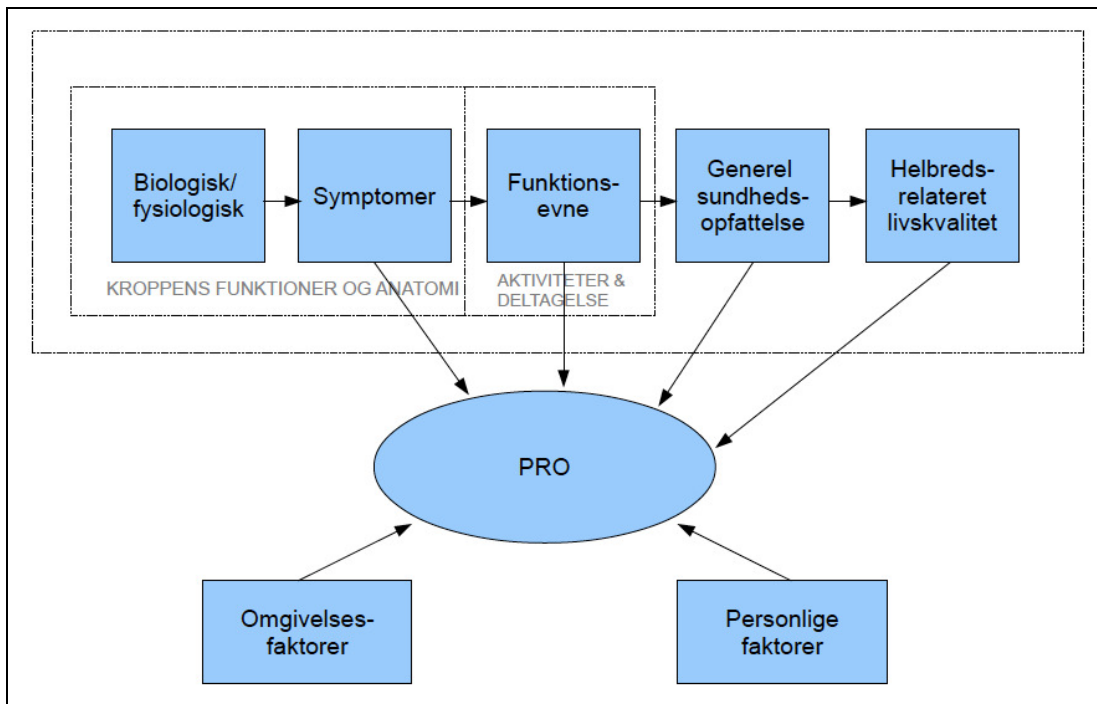
PRO-begrebet blev udviklet til brug ved lægemiddelforskning i USA, men indholdsmæssigt er konceptet ikke nyt og stammer tilbage til forskningstraditionen i helbredsrelateret livskvalitet (H-QoL), som blev udviklet i 1980-erne og efterhånden anvendt i forskning indenfor en række specialer. FDA har gjort det obligatorisk at rapportere PRO-data som resultatmål i forbindelse med medicinalforsøg og godkendelse af nye produkter. FDA’s retningslinjer har haft stor betydning for standardisering og udbredelse af PRO-data (4;5).

PRO kan anvendes til at afdække sundhedsrelaterede områder som symptomer, livskvalitet, funktionsniveau, compliance og andre forhold af betydning for vurdering af helbredstilstand og behandling (6;7). De instrumenter, der anvendes til at indsamle PRO-data på systematisk vis, benævnes ofte PROM (Patient Reported Outcome Measures). Det skal bemærkes at der aktuelt pågår en proces, hvor den aktuelle nomenklatur er under overvejelse, og det er således muligt at den nuværende nomenklatur vil blive afløst af en ny i de kommende år.

Valderas og Alonso har opstillet en model til klassificering af PRO-instrumenter og deres indholdselementer (Figur 1). Modellen bygger dels på Wilson og Cleary’s model for kategorisering af ”patient-outcomes” og disses indbyrdes forhold (8), og dels på WHO’s ICF-klassifikation, der kan betegnes som en begrebsramme for klassificering af funktionsevne (9). Modellen indeholder desuden biologiske og fysiologiske domæner, men selvom disse data kan rapporteres som PRO (f.eks. vægt), vil dette relativt sjældent være tilfældet (10). De øvrige områder er derimod velegnede til PRO-indsamling:

- Symptomer: Patientens oplevelse af fysisk, psykiske og kognitive problemer.
- Funktionsevne: Patientens evne til at udføre opgaver og dagligdagsaktiviteter.
- Generel sundhedsopfattelse: Patientens subjektive sammenfatning af alle informationer om symptomstatus og funktionsevne.
- Helbredsrelateret livskvalitet: Livskvalitetsaspekter, der relaterer sig til patientens helbred.

Figur 1. En integreret model for helbredsoutcome. Modificeret fra Valderas & Alonso (10).



Modellen giver en ramme for overvejelser omkring, hvilke PRO-instrumenter, der er relevante at anvende, og til kategorisering af hvilke domæner de enkelte instrumenter indeholder.

Konklusion:

PRO-data kommer direkte fra patienten uden fortolkning fra fx sundhedsfagligt personale.

4.2. Anvendelse af PRO-data

Anvendelsesmuligheder for PRO-data kan beskrives ud fra en opdeling efter hvorvidt indsamlingen sker som en driftsopgave eller som enkeltstående projekt, samt om formålet retter sig mod gruppeniveau eller individuelt patientniveau, jf. nedenstående figur.

Tabel 1. Fire grundperspektiver for anvendelsen af PRO-data

	Projekt	Drift
Gruppeniveau	Forskning	Kliniske kvalitetsdatabaser
Individniveau	Prognostiske modeller	Klinisk anvendelse af PRO

PRO kan indsamles i tidsafgrænsede projekter eller som en kontinuerlig driftsopgave.

Projektmæssig indsamling og anvendelse ses i forskningsprojekter med analytiske og deskriptive formål på gruppeniveau. Projekter kan også have til formål at opstille prognostiske modeller for PRO-baserede outcomes baseret på de indsamlede data med henblik på at kunne beskrive prognose for den enkelte patient, ud fra en række indsamlede parametre.

PRO indsamlet som driftsopgave kan også opdeles i anvendelse på gruppe- og individniveau.

PRO-data kan anvendes driftsmæssigt i behandlingen af den enkelte patient, som det f.eks. sker på det reumatologiske område i DanBio (danbio-online.dk). I AmbuFlex (ambuflex.dk) indsamles og anvendes PRO i den ambulante opfølgning af patienter med epilepsi, søvnapnø, lungecancer, prostatacancer, nyreinsufficiens og astma. Oversigtsartikler har fundet positive effekter af anvendelse af PRO på patientniveau, herunder bedre læge-patient kommunikation, øget patientinddragelse samt bedre identifikation af symptomer og psykosociale problemer. En national analyse af klinisk anvendelse af telePRO iværksat som en del af regeringsudspillet "Digitalisering med effekt" har i foråret 2015 på baggrund af erfaringer fra AmbuFlex konkluderet, at der er netto sundhedsøkonomiske og kvalitative gevinster og anbefaler en national strategi for udbredelse af telePRO (11).

I de kliniske kvalitetsdatabaser anvendes PRO-data driftsmæssigt på gruppeniveau til vurdering af behandlingskvaliteten for en gruppe af patienter, der har gennemgået en specifik behandling eller forløb. PRO-data indsamlet på gruppeniveau kan ligeledes anvendes til forskningsprojekter, uanset om data er indsamlet i de kliniske kvalitetsdatabaser.

Som det fremgår af tabel 1, deler PRO til klinisk brug og til kvalitetsudvikling den essentielle egenskab, at de indgår som en kontinuerlig driftsopgave. Det vil sige, der skal være sat en infrastruktur op, som koster både i etablering og drift, og som har afledte omkostninger i form af tidsforbrug i klinikken. Det er derfor naturligt at overveje, om der kan udvikles en model, som muliggør fælles anvendelse af data. Udover en åbenlys økonomisk fordel er der også en række andre argumenter herfor:

Set fra patientens synspunkt er det ønskeligt at undgå besvarelse af flere spørgeskemaer, som omhandler samme temaer. Danske erfaringer med PRO-dataindsamling til kliniske databaser har vist, at det undertiden

kan være problematisk at indsamle PRO på gruppeniveau, hvis patienten ved besvarelsen forventer en personlig reaktion. Nogle patienter opfatter, uanset tydelige formuleringer i følgebrevet om det modsatte, at hospitalet læser besvarelsen og reagerer, hvis der findes behov herfor. Der vedlægges endog undertiden håndskrevne breve. Hvis PRO primært anvendes til klinisk opfølgning på den enkelte patient og sekundært som kvalitetsdata, vil dette kunne løse problemet (12).

Set fra hospitalets side er der også fordele. På samme måde som kliniske data fra journalsystemer i fremtiden ønskes overført direkte til kvalitetsdatabaser, kunne man tænke sig det samme med PRO-data. Dette ville samtidig løse de juridiske problemer, da PRO-data indsamlet til klinisk anvendelse, efterfølgende må indgå og bruges på gruppeniveau i godkendte kliniske kvalitetsdatabaser, mens det omvendt ikke er tilladt at anvende data fra kvalitetsdatabaser på individuelt patientniveau.

Men det er ikke nogen simpel sag, og der er kun få erfaringer at trække på. Den nationale analyse af telePRO omhandler udelukkende PRO på patientniveau. Med hensyn til PRO på gruppeniveau er der erfaring med at indsamle og anvende data både i Danmark og ikke mindst i udlandet. Der er dog uhyre få erfaringer med at anvende samme data til begge formål. Der er endvidere to problemer at forholde sig til:

a) Det er ikke helt samme type data der efterspørges

Som beskrevet i den nationale analyse af telePRO, er PRO til anvendelse på patientniveau tilpasset kliniske problemstillinger og beslutningssituationer såsom ”skal patienten ses eller ej?”, mens PRO til kvalitetsudvikling skal indeholde variabler/indikatorer som beskriver outcomes, som kan tænkes påvirkelig af behandlingskvalitet. Der kan være en stor fællesmængde, men de to typer spørgeskemaer vil ofte ikke være identiske.

b) Timing er forskellig

PRO til klinisk brug indhentes ift. kontrolintervaller, mens PRO til kvalitetssikring skal indhentes på fikserede tidspunkter: fx 3 måneder efter operation, 12 måneder efter diagnose. I kvalitetsdatabaserne vil man lægge stor vægt på, at disse tidsterminer overholdes mhp. validitet og sammenlignelighed.

I øjeblikket pågår et forsøg, hvor PRO indsamlet til rutinemæssige kontrol af lungecancerpatienter (PROLUC-2), indgår i et parallelprojekt, hvor PRO indsamles nationalt til kvalitetsbrug (PROLUC-1). Det sker ved, at en patient ikke bliver kontaktet i PROLUC-2, hvis vedkommende har afleveret data i PROLUC-1. Erfaringer fra dette projekt vil kunne anvendes ved fortsatte overvejelser om fælles brug af data i stor skala.

Konklusion:

PRO-data kan anvendes på flere niveauer og i forskellige sammenhænge. I de kliniske kvalitetsdatabaser anvendes data på gruppeniveau, men det kan overvejes, om data også bør anvendes på individniveau.

4.3. Er der evidens for at brug af PRO-data medfører bedre kvalitet?

Inddragelse af PRO-data på individniveau i klinisk praksis er undersøgt i en række studier, og resultaterne er blevet sammenfattet i systematiske oversigtsartikler. Brug af PRO-data kan forbedre kommunikationen mellem patienten og klinikerne og medføre øget patienttilfredshed (13;14). Der er endvidere stigende evidens for at PRO-data kan forbedre monitorering af behandling og identificere komplikationer i patientforløbet (14). PRO-data indhentet i perioden før behandlingsstart og ændringer i PRO-status under behandlingen, har en stærk sammenhæng med det efterfølgende sygdomsudfald, inklusiv overlevelse hos cancerpatienter (15) eller risiko for genindlæggelse, reoperation, og komplikationer hos ortopædiske patienter (16;17). Generelt er der dog kun meget begrænset evidens for, at brug af PRO-data på individniveau direkte medfører en ændret patientbehandling og et forbedret sygdomsudfald (13).

Når det gælder anvendelsen af PRO-data på gruppeniveau i form af kvalitetsindikatorer, er den eksisterende videnskabelig evidens endnu mere sparsom. Der foreligger endnu kun få studier på området (14), og ingen af disse har påvist, at brug af PRO-data medfører en bedre behandlingskvalitet eller et bedre sygdomsudfald for patienterne. I England har kontinuerlig indsamling og afrapportering af PRO-data på gruppeniveau på fire kirurgiske områder i perioden fra 2009-2012 tilsyneladende kun været associeret med minimale ændringer i variationen i behandlingspraksis og sygdomsudfald (18;19). Man mener dog, at disse data er misbrugt og fejlfortolket, idet de mange metodologiske problemstillinger ikke var håndteret på bedste vis (20). Tilsvarende kunne der i et studie fra USA ikke påvises nogen effekt af afrapportering af data vedr. patienternes funktionsniveau (21).

Konklusion:

Der er svag evidens for at anvendelsen af PRO-data på individniveau i klinisk praksis kan medføre et bedre patientforløb inkl. sygdomsudfald. Der er endnu ingen evidens for at brug af PRO-data på gruppeniveau, herunder brug af PRO-baserede kvalitetsindikatorer, medfører en bedre behandlingskvalitet og bedre sygdomsudfald. Der er dog tale om et umodent forskningsfelt med ganske få studier, som er behæftet med væsentlige metodologiske begrænsninger.

4.4. Definition af PRO-instrumenter

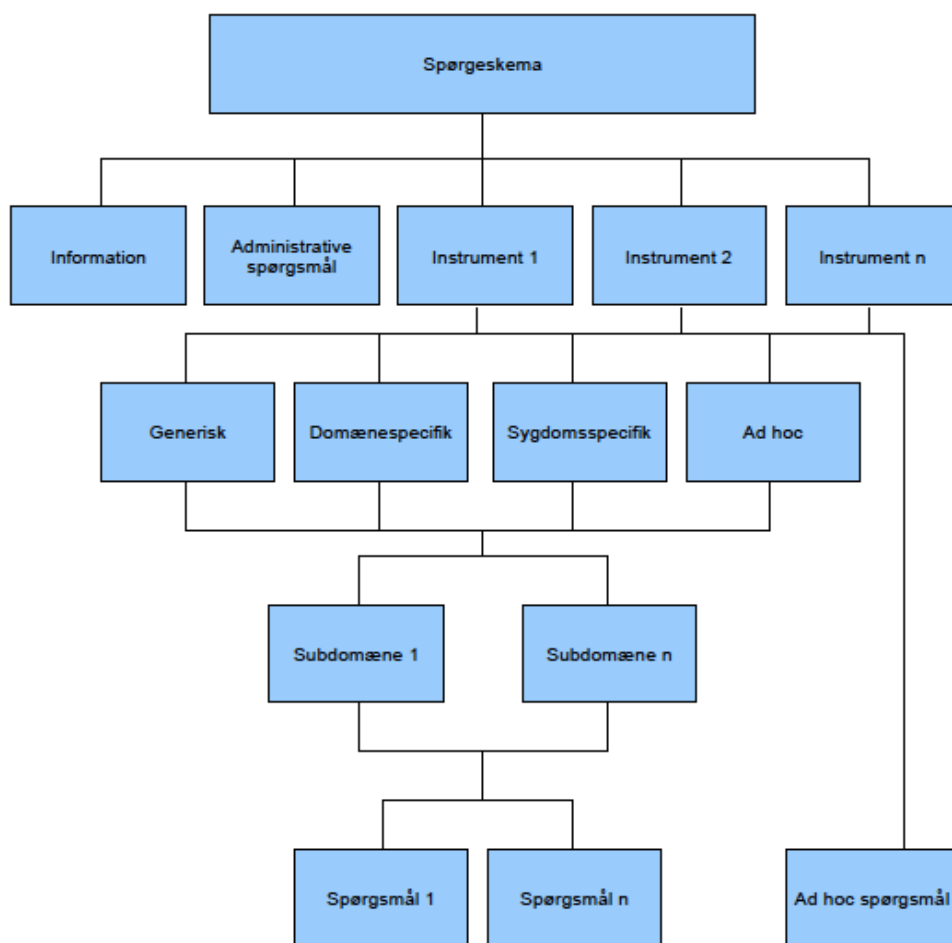
Spørgeskemaet som patienten præsenteres for, består af en overordnet vejledning og information til patienten om instrumentet, samt en række andre elementer (tabel 2).

De forskellige begrebs indbyrdes relation fremgår af nedenstående tabel 2 og figur 2. Et spørgeskema kan indeholde et eller flere PRO-instrumenter og et PRO-instrument kan bestå af flere skalaer eller af flere enkelte spørgsmål.

Tabel 2. Spørgeskematerminologi (modificeret fra (4)).

Spørgeskema	Den samlede enhed bestående af en overordnet vejledning samt et eller flere instrumenter. Eksempel: ”Hvordan har du det”-undersøgelsen.
Instrument	Et navngivet sæt spørgsmål som vedrører et eller flere domæner inklusiv information, som understøtter brugen af instrumentet, fx instruktioner og vejledning. Hvert svar tilknyttes en numerisk værdi, hvorfra der ved hjælp af en algoritme kan udregnes en eller flere scores. Et instrument kan dække et eller flere domæner. Eksempel: SF-36, HADS
Domæne	En begrebsmæssigt overordnet beskrivelse af det som søges målt. Et domæne kan evt. indeles i subdomæner. Eksempel: kognitiv funktion, depression, rygsmerter.
Skala	Metode til kvantitativ måling af tilstedeværelse af træk (traits) knyttet til domænet/subdomænet. Eksempel: SF-36 består af skalaer knyttet til to overordnede domæner: fysisk og psykisk livskvalitet. Derudover 8 skalaer knyttet til en række subdomæner: fysisk funktion, rollefunktion (fysisk), smerte, overordnet helbred, vitalitet, social funktion, rollefunktion (psykisk) og psykisk velbefindende. Bemærk at betegnelsen ”skala” også kan bruges om svarmulighederne for de enkelte spørgsmål.
Spørgsmål	Et spørgsmål eller udsagn med tilhørende svarmuligheder (på engelsk: item). Svarmuligheder kan være et tekstfelt, en skala (f.eks. en VAS skala) eller være opdelt i kategorier. Svarkategorisering af Likert-typen er hyppigt anvendt. En Likert-skala er oprindeligt en skala som måler graden af enighed/uenighed i forhold til et konkret udsagn, men bruges undertiden mere bredt om skalaer som måler grader i 4 eller flere udtømmende kategorier, f.eks. hyppigheder. Eksempel: ”Hvordan synes du dit helbred er alt i alt?” og svarkategorierne: Fremragende/ vældig godt/ godt/ mindre godt/ dårligt.

Figur 2. Spørgeskemaelementer



4.4. Definition af PRO-baserede kvalitetsindikatorer

PRO-data kan inkluderes i den systematiske monitorering og vurdering af behandlingskvaliteten og for at sammenligne afdelinger / indberettende enheder. PRO kan først og fremmest anvendes som resultatindikator, dvs. til måling af resultatet af en operation eller behandlingsforløb, og som procesindikator, fx andel af patienter, der har besvaret PRO-instrumentet. Derudover kan PRO benyttes som prognostisk faktor, dvs. som prædikator for godt eller dårligt udfald af behandlingen eller som confounder-kontrol i statistiske analyser, dvs. som forklarende variable i associationen mellem en bestemt eksponering og udfald.

Kvalitetsindikatorer baseret på PRO-data vil ligesom andre indikatorer, belyse i hvilken grad fastlagte kvalitetsmål er nået, og bør opfylde samme krav om relevans, videnskabeligt grundlag og anvendelighed (22).

Tabel 3. Eksempler på PRO-data, PRO-instrumenter og PRO-baserede kvalitetsindikatorer

Begreber	Patienter som gennemgik skulderalloplastik	Patienter som gennemgik fedmekirurgi
PRO-data	Fysiske symptomer, sport/fritid./arbejde, livsstil og følelser	Livskvalitet
PRO-instrument	Western Ontario Osteoarthritis of the Shoulder instrument (WOOS) (23)	Moorehead (24)
PRO-baseret kvalitetsindikator	Andel af patienter med WOOS score >40 point hos patienter som har fået skulderalloplastik pga. frisk fraktur	Andel af patienter med en uændret eller forbedret livskvalitet 1 år efter operationen.

Konklusion:

Det er væsentligt at skelne mellem PRO-data, -instrumenter og – kvalitetsindikatorer, selvom der er en tæt sammenhæng mellem begreberne. PRO-kvalitetsindikatorer bygger på PRO-data, som indsamles gennem PRO-instrumenter.

5. Datatilladelser og datasikkerhed

Bekendtgørelse nr. 459 af 16/05/2006 (<https://www.retsinformation.dk/Forms/R0710.aspx?id=10263>) om godkendelse af landsdækkende og regionale kliniske kvalitetsdatabaser og Bekendtgørelse nr. 1725 af 21.12.2006 (<https://www.retsinformation.dk/Forms/R0710.aspx?id=11046>) om indberetning af oplysninger til kliniske kvalitetsdatabaser beskriver de grundlæggende principper vedr. indsamling af data til kliniske databaser. Indsamling af PRO-data til de godkendte kliniske kvalitetsdatabaser bør foregå efter de principper, som er beskrevet i de to bekendtgørelser.

Derfor bør man sikre sig at indsamling af PRO-data og deres formål er velbeskrevet i databasen og i ansøgningen om godkendelse som klinisk kvalitetsdatabase.

Ifølge persondataloven (Lov nr. 429 af 31/05/2000

<https://www.retsinformation.dk/Forms/R0710.aspx?id=828>) skal indsamling, behandling og opbevaring af personoplysninger som hovedregel anmeldes til Datatilsynet. Endvidere, ifølge persondataloven, må indsamling og behandling af data kun ske efter databasens formål, hvilket også gælder for PRO-data.

Af persondataloven fremgår også, at der skal træffes de fornødne tekniske og organisatoriske sikkerhedsforanstaltninger mod, at oplysninger hændeligt eller ulovligt tilintetgøres, fortabes eller forringes, samt mod, at de kommer til uvedkommendes kendskab, misbruges eller i øvrigt behandles i strid med loven.

Endvidere beskriver persondataloven, hvilke sikkerhedsforanstaltninger der skal foretages i forbindelse med overførsel og udveksling/videregivelse af data, hvilket også gælder for PRO-data.

Den generelle standard for datasikkerhed har hidtil været beskrevet i DS 484, men denne er nu under udfasning til fordel for ISO 27001 (25), som indebærer krav vedrørende implementering og vedligeholdelse af et informationssikkerhedsledelsessystem, som i så fald også skal omfatte PRO-data.

Hvis der er det mindste tvivl omkring gældende juridiske forhold i relation til PRO-data, anbefales det at tage kontakt til Databasernes Fællessekretariat, Regionernes Kliniske Kvalitetsudviklingsprogram eller Datatilsynet i den region, som er databasens dataansvarlige myndighed.

I planlægningsfasen, hvor man overvejer forskellige PRO-instrumenter, skal databasen undersøge om instrumentet er gratis eller om der skal betales licens for at bruge instrumentet. For eksempel kræver anvendelse af EuroQol-instrumentet (EQ-5D) licens. Beløbet fastsættes af EuroQol-forretningsudvalget, på baggrund af de oplysninger man har angivet i ansøgningsskemaet. Beløbet afhænger af hvilken type undersøgelse / forsøg / projekt, finansieringskilde, stikprøvestørrelse og anmodede antal sprog.

Konklusion:

Indsamling og anvendelse af PRO-data til de godkendte kliniske kvalitetsdatabaser skal foregå efter de samme principper, som gælder for alle andre data i de kliniske databaser, jf. de to bekendtgørelser (Bekendtgørelse nr. 459 af 16/05/2006 og Bekendtgørelse nr. 1725 af 21.12.2006) for kliniske kvalitetsdatabaser samt persondataloven.

6. Erfaringer med PRO-data i de kliniske kvalitetsdatabaser

6.1. Danmark

I 7 ud af 60 danske kliniske kvalitetsdatabaser er der implementeret PRO-data på nuværende tidspunkt.

PRO-data har været en del af **Dansk Skulderalloplastik Register (DSR)** siden etablering af DSR i 2005. DSR anvender Western Ontario Osteoarthritis of the Shoulder (WOOS) spørgeskemaet, som er udviklet i England 2001 (23), specifikt til at evaluere resultatet efter skulderproteseoperation, og er valideret og oversat til dansk (26). WOOS indeholder 4 subdomæner: fysiske symptomer (6 spørgsmål), sport/fritid/arbejde (5 spørgsmål), livsstil (5 spørgsmål) og følelser (3 spørgsmål). Skemaet udsendes i papirform med forudbetalt svarkuvert til alle ca. 1000 patienter om året, 1 år efter operation. De returnerede skemaer indtastes manuelt i DSR via det internetbaserede Klinisk Måle System (KMS). Databasen sender ingen rykkerbreve til patienterne. Responsraten ligger på ca. 70%. WOOS data anvendes i DSR som resultatindikator med henblik på at sammenligne operationskvaliteten på tværs af afdelingerne: Andel af patienter med WOOS score >50

point for opererede patienter hvor skulderprotese er indsat under indikationen artrit/ artrose. Grænsen på 50 point er fastsat ud fra konsensus i styregruppen. Der findes således ikke en international standard, som kan anvendes som grænse for optimal postoperative WOOS-totalscore. Der indsamles ikke PRO-data før operationen, hvilket gør det besværligt at sammenligne på tværs af afdelingerne, idet det ikke er muligt at tage højde for eventuelle forskelle i sværhedsgraden af skulderproblemer ved baseline.

PRO-data har også været en del af **Dansk Korsbånds Rekonstruktions Register (DKRR)** siden etablering i 2005. Cirka 2500 patienter er blevet opereret og registreret i DKRR. DKRR anvender "Knee injury and Osteoarthritis Outcome Score" (KOOS) (27), som er et selvevalueringsinstrument vedrørende patienternes subjektive knæfunktion. Patienterne indberetter elektronisk både før og 1 år efter korsbåndsoperation. Ved et ambulant besøg får patienterne påmindelse om at udfylde skemaet. Databasen sender ingen rykkerbreve eller rykkermails til patienterne, men sidstnævnte er ved at blive etableret. PRO-data kobles ikke automatisk til primær operation, men efterfølgende i Kompetencecenter for Epidemiologi og Biostatistik. Responsraten ligger på 30% for KOOS før operationen og 35% efter operationen, hvilket vanskeliggør fortolkningen af data. KOOS-data anvendes derfor ikke som resultatindikator i DKRR.

Dansk Palliativ Database har siden 2010 årligt indsamlet PRO-data på ca. 6000 patienter med palliative behov. European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 spørgeskemaet består af 30 spørgsmål om livskvalitet, som er generelt for alle cancerdiagnosegrupper, og er blevet et af de mest anvendte spørgeskemaer i internationale undersøgelser af kræftpatienter. Det er oversat til over 80 sprog. En version med 15 supplerende spørgsmål (28) er blevet udviklet specielt til patienter, der modtager en palliativ indsats (Palliative Care questionnaire (QLQ-C15-PAL)). PRO-data indsamles i papirform indenfor de seneste tre dage før eller på dagen for første palliative kontakt, og indtastes manuelt i databasen via KMS. Responsraten ligger på ca. 50%. Det er uklart, om der sendes rykkerbrev til patienterne, og hvordan patienterne kontaktes. PRO-data anvendes som procesindikator i databasen, i form af andelen af patienter, hvor skemaet bliver udfyldt.

Dansk Fedmekirurgi Register (DFR) indsamler PRO-data på ca. 700 patienter både før og 1 år efter operation. Moorehead-skemaet er et livskvalitetsinstrument specifikt udviklet til at afspejle livskvalitet hos patienter, som har undergået et fedmekirurgisk indgreb (24). Der anvendes en dansk version af skemaet i DFR, men der er ikke foretaget validering af den danske version. Data indsamles i papirform, og patienterne svarer kun på PRO-skemaet i forbindelse med fysisk fremmøde i ambulatoriet før og 1 år efter operation. Der er ingen rykkerprocedure. Responsraten ligger på ca. 50%. PRO-data anvendes som resultatindikator i DFR, hvor forbedringen i livskvaliteten fra før til 1 år efter operation beregnes.

Dansk Urogynækologisk Database (DUGABase) har siden 2006 indsamlet PRO-data i papirform både før og efter operation. Subjektiv patientbedømmelse efter inkontinensoperation (eller efter prolapsoperation) måles med ”Patient Global Index of Improvement scale” (PGI-I) (29). Det er en selvrapporteret single-item-score, udviklet til kvinder med stress-inkontinens. Scoren afdækker ændringer i selvoplevet stress-inkontinens, dvs. med et enkelt spørgsmål bliver patienten bedt om at bedømme deres urinvejes tilstand nu, i forhold til før behandlingsstart på en skala fra 1 til 7. Skemaet kan også anvendes til kvinder med genital prolaps eller overaktiv blære, og til mænd med vandladningsproblemer relateret til forstørret prostata. Ca. 5000 patienter inkluderes årligt i databasen, hvoraf ca. 90% udfylder skemaet. PRO-data anvendes som resultatindikator, idet forbedring fra før til efter behandling bliver beregnet.

Dansk Reumatologisk Database (DANBIO) har i mange år indsamlet PRO-data ved start af behandling med antireumatisk lægemiddel og 2 år og flere år efter behandlingsstart. Der bliver indsamlet data fra patienten vedr. livskvalitet (VAS- global skema), smerter (VAS - smerte skema), træthed (VAS – træthed skema) og funktionsniveau (HAQ skema). Patienten indtaster de første besvarelser efter ankomst til ambulatoriet på en trykfølsom skærm (patient-kiosk). Efterfølgende indsamling af data foregår online. Databasen har formået at opnå en meget høj svarprocent på over 90% på nævnte PRO-instrumenter (30). PRO-data anvendes som resultatindikatorer i databasen.

PRO-data indsamles også i **Dansk Prostatacancerdatabase (DaProCa)** siden 2009. Der inkluderes ca. 4000 nydiagnosticerede prostatacancerpatienter om året i databasen. Der indsamles data på diagnosetidspunkt for prostatakræft og 12 måneder efter. Det overvejes aktuelt, at supplere med yderligere dataindsamling efter 36 måneders follow-up. Første spørgeskema udleveres på afdelingen, hvor patienten får stillet diagnosen. Patienten svarer på spørgsmål vedr. rygning, alkohol, fysisk form, selvhjulpethed, vandladning, afføring og seksuel funktion, samt om hestigninger, brystspændinger/hævelser, depression, manglende energi og ændret legemsvægt. 12-måneders skemaer bliver tilsendt patienten fra VestKronik i Herning. Besvarelse sker i papirform, dog kan patienterne vælge at svare på opfølgningsskemaerne online, hvis de ønsker. Der er to rykkerprocedurer og en svarprocent på 91.5% på follow-up skemaet. PRO-data anvendes for nuværende ikke som indikator i databasen.

Konklusion:

PRO-data indsamles og anvendes kun i relativt få nationale kliniske databaser i Danmark. Databaserne har typisk anvendt allerede eksisterende PRO-instrumenter, som regel efter en officiel validering og oversættelse til dansk. De fleste databaser forsøger at indsamle PRO-data både før og efter operation/ behandling med henblik på at kunne måle ændringer efter intervention og tage højde for forskelle ved baseline. Data anvendes primært som resultatindikatorer. Lave besvarelsesprocenter har i nogle databaser givet

udfordringer i forhold til fortolkning af resultater og sammenligninger på tværs af afdelinger. Ingen af databaserne har dog i analysefasen taget stilling til håndtering af lave besvarelsesprocenter, kun en måling af PRO-data, eller anvendt PRO-data til justering for case-mix. Elektronisk indsamling af PRO-data er mere undtagelsen end reglen, og de indsamlede PRO-data er derfor ikke en del af den elektroniske patientjournal.

6.2. Udland

PRO-data indsamles til kvalitetsudvikling i en række lande. Det er uden for denne vejlednings rammer at give et fyldestgørende indblik i de udenlandske erfaringer med anvendelse af PRO i kvalitetssammenhænge, men herunder følger eksempler på nationale PRO-initiativer i England, Sverige, New Zealand og USA.

I **England** har man siden 2009 rutinemæssigt indsamlet PRO i forbindelse med fire elektive indgreb: hoftealloplastik, knæalloplastik, operation for lyskebrok og operation for åreknuder.

Til alle fire grupper anvendes det generiske instrument EQ-5D samt et sygdomsspecifikt instrument til hofte, knæ og åreknuder. Dette suppleres med spørgsmål om boforhold, funktionsnedsættelse og hvorvidt patienten selv har udfyldt skemaet.

Data indsamles præoperativt, i intervallet mellem at patienten er henvist til operationen og selve operationsdagen. Postoperativt indsamles data 6 måneder efter hofte- og knæalloplastik, og 3 måneder efter lyskebrok- og åreknudeoperation. I det postoperative spørgeskema er der tilføjet spørgsmål om komplikationer, selv vurderet resultat af behandling og generelt helbred. PRO anvendes som resultatindikator, og der beregnes, på baggrund af præ- og postoperative data, forbedringer i generiske og sygdomsspecifikke scores.

Der anvendes papirskemaer, som scannes og kobles på patientniveau til data fra hospitalsregister mv., når dette er muligt. Responsraten i det postoperative skema ligger på ca. 85% for hofte- og knæalloplastik, 73% for lyskeoperation og 65% for åreknudeoperation (31). Der er planer om at udvide indsamlingen af PRO til andre indgreb og på sigt også kroniske sygdomme (32;33).

Initiativet til indsamlingen er kommet fra regeringens side, og implementering af PRO på hospitalerne har været styret ovenfra. De enkelte hospitaler har haft et stærkt økonomisk incitament til at indsamle data. Der udarbejdes månedlige kvalitetsrapporter, som muliggør sammenligning af hospitaler og de enkelte kirurger (20).

I **Sverige** har man i en årrække, overvejende på initiativ fra klinisk side, indsamlet PRO-data til nationale diagnosespecifikke kvalitetsdatabaser. PRO indgår nu i en lang række af de svenske kvalitetsdatabaser. For en gennemgang af alle de svenske databaser, henvises til et publiceret oversigtsartikel fra 2011 (34), samt en oversigt over kvalitetsdatabaser som anvender PRO fra 2015, udgivet af det svenske PROM-center (35).

Flere af databaserne er for nyligt startet PRO-indsamling, og der foreligger derfor ikke rapporter på disse data endnu.

Her følger blot to eksempler fra Sverige:

Det svenske hofteregister (www.shpr.se/en/) inkluderer PRO-instrumenterne Charnley classification, smerte-VAS-score og EQ-5D. Patienterne svarer på et spørgeskema præoperativt via en berøringfølsom skærm eller i papirformat på hospitalet. Postoperativt besvarer patienterne spørgeskemaet efter 1 år, og på nogle afdelinger igen efter 6 år. Responsraten er ca. 85% på det præoperative skema, mens den er næsten 90% ved et års follow up (36;37). PRO anvendes som resultatindikator, og de årlige rapporter fra det svenske hofteregister beskriver bl.a. den aktuelle gennemsnitsscore for de enkelte hospitaler og den forventede score (36).

Det svenske gastro- og oesophagal cancer register har siden 2009 indsamlet PRO-data om helbredsrelateret livskvalitet. De anvender validerede instrumenter i form af EQ-5D og EORTC QLQ-C30, samt tilhørende diagnosespecifikke spørgsmål (QLQ-C25). Spørgeskemaet sendes ud til patienter et år efter behandling. Svarkategorierne i EORTC-instrumentet anvendes dikotomiseret, således at andel af patienter, der rapporterer enten "En hel del" eller "Meget" i forhold symptomer og funktionsproblemer, kategoriseres som havende symptomer. Andel af patienter med symptomer beskrives i årlige kvalitetsrapporter. Responsraten er 58,3%, og der er tale om en lille patientgruppe, hvilket betyder at der er usikkerhed forbundet med resultaterne (38).

I **New Zealand** har New Zealand Joint Registry (www.nzoa.org.nz/nz-joint-registry) anvendt de sygdomsspecifikke PRO-instrumenter Oxford Hip Score og Oxford Knee Score siden 1998. Herudover indgår spørgsmål om bl.a. infektion.

PRO indsamles præoperativt og postoperativt efter 6 måneder og efter 5 år. På grund af den tiltagende patientvolumen og dermed den øgede logistiske og økonomiske byrde ved dataindsamlingen, indsamles nu kun PRO i et sample bestående af 20% af det totale antal patienter, der har fået udført hofte- eller knækirurgi (baseret på tilfældig udtræk af patienter). Ti procent af patienterne benytter sig af muligheden for at svare på internettet, og responsraten ved det første skema er ca. 70-75% (39).

PRO-resultater indgår i registrets årsrapport, hvor Oxford-score anvendes som mulig prædikator for revision inden for 2 år. Dette mhp. på sigt at anvende Oxford-score som screeningsværktøj, så kun patienter som scorer dårligt på Oxford-score, og dermed en større risiko for revision indenfor 2 år, vil blive indkaldt til klinisk kontrol.

Også i **USA** anvendes PRO til vurdering af behandlingskvalitet, dog ikke på nationalt niveau endnu. I 2013 udkom The National Quality Forums hvidbog, der beskriver hvordan PRO udvikles til anvendelse i "Performance Measurement" - PRO-PM (6). På baggrund af dette, er der blevet nedsat en arbejdsgruppe,

som er i gang med at udvikle PRO-PM specifikt for onkologien, som kan indgå i kvalitetsprogrammet. Gruppen besluttede at fokusere på PRO som resultatindikator, men ønsker på sigt også inddragelse som procesindikator. I første omgang prioriteres områderne smerte og kvalme. Der er udvalgt spørgsmål om kvalme fra PRO-CTCAE, som er PRO-versionen af Common Terminology Criteria for Adverse Events (PRO-CTCAE), der er en amerikansk udviklet skala til beskrivelse af bivirkninger ved kemoterapi. Der spørges til hyppighed og sværhedsgrad af kvalme. Til måling af smerte anvendes et spørgsmål fra Brief Pain Inventory omkring værste smerte indenfor de sidste 24 timer på en skala fra 1-10. De udvalgte PRO afprøves nu i pilottest (40).

Der findes derudover en lang række lokale initiativer. I Minnesota f.eks., indsamles PRO-data i primærsektoren (DIAMOND) i form af The Patient Health Questionnaire (PHQ-9), der er et valideret spørgeskema til screening for depression. Skemaet bruges som resultatindikator til at beregne andel patienter med depression, der er i bedring efter 6 måneder, defineret som andel af patienter der ved diagnosetidspunkt havde PHQ-9 score over 9, som efter 6 måneder har en score under 5. Derudover anvendes PRO-skemaet som procesindikator, idet andel af patienter, der besvarer PHQ-9 instrumentet beregnes (41).

I en rapport fra det amerikanske Patient-Centered Outcomes Research Institute (PCORI) findes beskrivelser af en række eksempler på anvendelse af PRO, som er integreret i den elektroniske patientjournal. Herunder findes information om de enkelte eksemplers anvendelse af data til kvalitetsudvikling (42).

Konklusion:

Nogle lande har adskillige års erfaringer med anvendelse af PRO-data i kvalitetssammenhæng. Generelt er det først inden for de seneste 10-15 år, at udviklingen har taget fart. De beskrevne eksempler på udenlandske databaser anvender oftest validerede instrumenter, typisk en kombination af et generisk og et sygdomsspecifikt. I nogle tilfælde har man tilføjet ad hoc-spørgsmål. Databaserne anvender primært PRO som resultatindikator. Der må forventes en stigning i publikationer på området i de kommende år, i takt med at flere databaser implementerer PRO.

7. Hvilken type PRO-instrument databasen skal vælge

Typen af PRO-instrument hænger nøje sammen med formålet for PRO-indsamlingen. Det skal overvejes på hvilke områder/ domæner de eksisterende variabler findes utilstrækkelige, hvad man gerne vil måle og hvorvidt det lader sig gøre at måle dette i form af PRO. Vil et patientperspektiv på behandlingseffekten bidrage med en mere retvisende vurdering af effekten af pleje og behandling?

I diskussionen af formålet med at indsamle PRO må det også sikres, at data ikke kan fremskaffes ad andre veje, således at patienterne skånes for overflødige spørgsmål. Det bør medtænkes, hvad der er bydende nødvendigt at vide (need-to-know), kontra hvad der blot ville være ønskeligt (nice-to-know).

Konklusion:

Hvis man overvejer at indsamle PRO-data i en klinisk kvalitetsdatabase, er det relevant først at tage stilling til følgende spørgsmål:

- Findes der aspekter (domæner) som ikke kan måles med eksisterende outcomes/indikatorer?
- Er PRO-data egnede til at måle dette aspekt (domæne)?
- Hvilke PRO-data er nødvendige for at måle dette?
- Skal PRO-data også anvendes på individniveau?

Når formålet er fastsat, bliver det næste skridt, at vælge et eller flere egnede PRO-instrumenter.

7.1. Generiske PRO-instrumenter

Generiske instrumenter (5;43) kan med fordel anvendes, fordi de kan bruges på tværs af patientgrupper. Data fra patienterne kan derfor sammenlignes med data fra andre patientgrupper eller evt. med referencemateriale fra den generelle befolkning. Generiske instrumenter er ofte, qua deres udbredelse, velbeskrevet med hensyn til validitet, reliabilitet og responsrate. En ulempe kan være, at et generisk instrument kun omfatter domæner med bred relevans. Desuden skal skalaer dække et bredt range af sværhedsgrader, og er dermed mindre følsomme overfor niveauet af patientens helbredstilstand. Dette bevirker, at der kan opstå floor eller ceiling effect, som udover tab af præcision, kan medføre at instrumentet ikke opleves som relevant af patienten.

Tabel 4. Eksempler på hyppigt anvendte generiske PRO-instrumenter

SF-12, SF-36	Et hyppigt anvendt instrument til vurdering af generel helbredsstatus og livskvalitet. 36 spørgsmål inden for 8 domæner, som kan sammenfattes til en fysisk og en psykisk sumscore. Findes også i en kortere version, SF-12. Det oprindelige SF-36 er public domain men anvendes sjældent i dag. Den mest anvendte version (SF-12 v2) er belagt med copyright og ganske bekostelig at anvende. www.sf-36.org
EQ-5D	EuroQol-instrumentet anvendes til vurdering af generel helbredsstatus. 5 spørgsmål samt en VAS-skala, som afdækker fysisk, psykisk og social funktion. Instrumentet kan desuden anvendes til cost-utility analyser. Skemaet kan kun anvendes efter tilladelse, og prisen fastsættes ved det enkelte projekt. www.euroqol.org
EORTC QLQ-C30	Instrumentet er udviklet af the European Organisation for Research and Treatment of Cancer til at måle livskvalitet hos patienter med kræft. Består af 30 generiske kerne-spørgsmål, som kan anvendes sammen med sygdomsspecifikke moduler. Organisationen er i gang med at udvikle en QLQ-C30 baseret på Computer Adaptive Testing.

	EORTC har copyright og der skal ansøges om tilladelse til brug. http://groups.eortc.be/qol/eortc-qlq-c30
WHO-5	Fem spørgsmål om generelt helbred i de seneste to uger. WHO-5 måler graden af positive oplevelser, og kan anvendes som et mål for personers generelle trivsel eller velbefindende. Der beregnes én samlet score. Er gratis at anvende. www.who-5.org

7.2. Domænespecifikke PRO-instrumenter

Domænespecifikke instrumenter måler forekomsten af et bestemt fænomen f.eks. angst, depression, træthed, kvalme eller smerter. Disse instrumenter kan anvendes på enhver diagnosegruppe, hvor fænomenet er relevant. Ligesom ved de generiske instrumenter foreligger der ofte fyldig dokumentation af psykometriske egenskaber (5;43). Sygdomsspecifikke instrumenter indeholder ofte domænespecifikke elementer. Selvom det f.eks. giver begrænset mening at tale om en diagnosespecifik form for træthed, kan der optælles 252 instrumenter til måling af træthed (44), heraf de fleste ad hoc-udviklede, på trods af at der forefindes flere veldokumenterede domænespecifikke instrumenter af varierende længder.

Domænespecifikke instrumenter (43) kan være hensigtsmæssige at anvende, hvis et bestemt domæne er særligt relevant for den givne patientgruppe, og der ønskes en grundig og detaljeret beskrivelse af dette domæne. Domænespecifikke instrumenter kan anvendes i flere patientgrupper, og giver dermed mulighed for at sammenligne på tværs af grupper.

Tabel 5. Eksempler på domænespecifikke PRO-instrumenter

HADS	Hospital Anxiety and Depression Scale består af 14 spørgsmål, 7 udsagn relateret til generaliseret angst og 7 til depression. Instrumentet kan anvendes efter tilladelse. Der opkræves gebyr for anvendelse. www.proqolid.org
MFI-20	Multidimensional Fatigue Inventory er et instrument med 20 spørgsmål udviklet til at belyse træthed. Træthed afdækkes indenfor fem dimensioner: generel træthed, fysisk træthed, mental træthed, nedsat motivation og nedsat aktivitet. Instrumentet har copyright, men kan efter tilladelse anvendes uden gebyr. www.proqolid.org

7.3. Sygdomsspecifikke PRO-instrumenter

Sygdomsspecifikke instrumenter fokuserer på domæner, der anses som specifikke for patienter med en bestemt sygdom. De giver mere detaljeret information om særligt relevante problemstillinger hos en given patientgruppe (5). En del sygdomsspecifikke instrumenter inddrager imidlertid også generiske elementer for symptomer og funktion.

Sygdomsspecifikke instrumenter er hensigtsmæssige at vælge, hvis man ønsker at måle specifikke symptomer eller begrænsninger, som ikke indfanges i et generisk eller domænespecifikt instrument. De sygdomsspecifikke instrumenter vil ofte opleves som relevante af patienterne (5), da de er udarbejdet præcist til den givne patientgruppe. De sygdomsspecifikke instrumenter er ofte tilstrækkeligt sensitive til at måle selv små ændringer i et specifikt sygdomsproblem (5;43). Problemet ved at anvende et sygdomsspecifikt instrument kan være, at sammenligningsmulighederne begrænses.

Tabel 6. Eksempler på sygdomsspecifikke PRO-instrumenter

OHS	Oxford Hip Score består af 12 spørgsmål om smerter og funktionsbegrænsninger som følge af hofteproblemer. Instrumentet er udviklet til patienter som skal have udskiftet en hofte. http://www.bjj.boneandjoint.org.uk/content/78-B/2/185.full.pdf
CAT	COOPD Assessment Test er et instrument med 8 spørgsmål specifikt rettet mod KOL-patienter. Der spørges til symptomer, funktionsniveau og mental tilstand. Instrumentet har copyright, men kan muligvis efter tilladelse anvendes uden gebyr. www.catestonline.org
EORTC QLQ-LC13	Tillægsmodul til EORTC QLQ-C30 udviklet specifikt til patienter med lungecancer, som kan tilføjes det generiske kernemodul. Det består af 13 spørgsmål som retter sig mod symptomer hos patienter med lungekræft. Der er udviklet moduler til en række cancerformer. http://groups.eortc.be/qol/why-do-we-need-modules

7.4. Ad hoc spørgsmål

Ad hoc-udviklede instrumenter er instrumenter, der udvikles til det aktuelle formål. Dette forudsætter dog et omfangsrigt metodearbejde, hvis der tale om egentlige skalaer.

Ad hoc-spørgsmål kan være relevante som et supplement til et eksisterende instrument, f.eks. hvis de valgte instrumenter ikke indfanger et specifikt og betydningsfuldt fænomen eller symptom. Ad hoc-spørgsmål bør placeres umiddelbart efter det instrument, det er relateret til, eller eventuelt i slutningen af spørgeskemaet, for ikke at påvirke besvarelsen af de øvrige instrumenter (5).

Det kan overordnet konstateres, at nogle PRO-instrumenter er omfattende med mange spørgsmål, og kan anses som et forsøg på en detaljeret PRO-baseret karakteristik af patientens tilstand i forhold til det pågældende subdomæne, og er undertiden primært egnet til forskningsbrug, mens andre instrumenter belyser samme domæne meget kortfattet med blot enkelte spørgsmål, og derfor må formodes at have lavere præcision.

En nyere type instrumenter er de dynamisk individuelt tilpassede. De benytter Item Response Theory (45) og Computerized Adaptive Testing (46) med henblik på kun at stille patienten de relevante og nødvendige spørgsmål, men med den samme præcision som et langt instrument. Spørgsmål udvælges fra en "itembank" på baggrund af patientens forudgående svar ud fra en algoritme. Det amerikanske PROMIS-initiativ er et af de mest systematiske initiativer på dette område, (nihpromis.org). Der er på nuværende tidspunkt kun oversat enkelte itembanker til dansk.

I overvejelserne omkring hvilke instrumenter, der skal vælges til en kliniske kvalitetsdatabase, anbefales det som udgangspunkt at anvende en grundstamme af et generisk eller et domænespecifikt instrument, der kombineres med et sygdomsspecifikt instrument og/eller ad hoc-spørgsmål.

Når man opbygger et spørgeskema af allerede etablerede instrumenter, bør man aldrig på egen hånd 'rette' i et instrumentets formuleringer eller svarkategorier, idet man derved mister referencen til både valideringsstudier og referencegrupper.

Konklusion

PRO-instrumenter kan overordnet kategoriseres efter, hvilken gruppe af patienter de er anvendelige på og hvad de måler. Generiske instrumenter kan anvendes bredt på tværs af sygdomsgrupper. Domænespecifikke instrumenter måler et bestemt domæne detaljeret, f.eks. smerte eller træthed, og kan anvendes bredt på grupper af patienter, hvor domænet er relevant. Sygdomsspecifikke instrumenter er udviklet til bestemte patientgrupper, og måler på forhold af særlig relevans for den givne gruppe. Ad hoc-instrumenter er udviklet til det aktuelle formål.

Hvis man skal vælge et PRO-instrument til brug i en klinisk kvalitetsdatabase, vil det være relevant at medtænke følgende:

- Generiske instrumenter giver mulighed for sammenligning på tværs af patientgrupper.
- Sygdomsspecifikke instrumenter vil oftest opleves meget relevante af patienterne.
- Det anbefales at bruge en grundstamme af generiske og/eller domænespecifikke instrumenter, som kan kombineres med et sygdomsspecifikt instrument eller ad hoc-spørgsmål.
- Enkeltstående supplerende ad hoc-spørgsmål kan anvendes, mens ad hoc-instrumenter med sumscoringalgoritmer kun bør anvendes efter en detaljeret og specifik valideringsproces.

8. Valg af eksisterende eller udvikling af nye PRO-instrumenter

8.1. Eksisterende PRO-instrumenter

Der findes talrige PRO-instrumenter indenfor ethvert domæne, og det vil derfor være relevant at starte med at undersøge, om der allerede er udviklet et instrument, som kan anvendes i databasen. Det første step vil derfor være en litteratursøgning med henblik på at afdække relevante PRO-instrumenter på området.

Søgningen bør omfatte relevante databaser som fx PubMed, EMBASE, PsycINFO. Her vil det ofte være muligt at finde artikler, der dokumenterer valideringsprocesser for de enkelte instrumenter og oversigtsartikler på specifikke patientgrupper (5).

Derudover findes der forskellige hjemmesider, som forsøger at give et overblik over eksisterende PRO-instrumenter. Det gælder fx www.proqolid.org, www.OLGA-qol.com og www.pfizerpatientreportedoutcomes.com, som alle er på engelsk. (45). På dansk findes en oversigt på <http://fysio.dk/fafo/Maleredskaber/Om-Maleredskaber/>.

En fordel ved at anvende et eksisterende PRO-instrument er, at det omfattende arbejde med at udvikle og validere instrumentet allerede er foretaget. Der vil sandsynligvis findes beskrivelse af instrumentets psykometriske egenskaber og referencedata.

På den anden side bør man overveje om et aktuelt instrumentet er anvendeligt i forhold til målgruppen. Er instrumentet tidligere anvendt på en målgruppe svarende til databasens? Det skal desuden undersøges, om instrumentet er beskyttet med copyright, og det kan i nogle tilfælde være bekosteligt at få tilladelse til at anvende instrumentet. Endeligt skal det undersøges, om instrumentet findes i en dansk valideret version.

8.2. Oversættelse af eksisterende PRO-instrumenter

Findes der et relevant PRO-instrument, som ikke er oversat til dansk, foreligger der et betydeligt oversættelsesarbejde. Målet er at skabe en dansk version, som er lige så klar og præcis som den oprindelige version og med de samme psykometriske egenskaber. Fokus bør være på forståelse og ikke på ord-for-ord oversættelse. Oversættelsen skal fungere i den kulturelle kontekst, den skal anvendes i (5;47). Opgaven kan inddeles i følgende trin:

1. Oversættelse (forward translation)

En oversætter, der kender terminologien på området, oversætter instrumentet fra engelsk til dansk. Oversætteren bør have dansk som modersmål, men bør samtidig have kendskab til den engelsktalende kultur.

2. Ekspertpanel

Et panel af eksperter, der taler både dansk og engelsk, gennemgår skemaet i forhold til ord og vendinger, og sammenholder det med den oprindelige version. Der foretages nødvendige rettelser i oversættelsen.

3. Tilbage-oversættelse (backward translation)

En af projektet uafhængig translatør oversætter den danske version af skemaet tilbage til engelsk. Oversætteren bør have engelsk som modersmål, og bør ikke have kendskab til instrumentet.

4. Sammenligning

En anden uafhængig person skal nu sammenligne spørgsmål og svarmuligheder i hvert spørgsmål i originalversionen med oversættelsen tilbage til engelsk og skrive alle forskelle ned.

5. Pilottest og kognitive interviews

Det danske instrument skal nu testes i målpopulationen. Formålet er at undersøge patienterne forståelse, og om der er forvirrende spørgsmål eller svarmuligheder, spørgsmål der er svære at forstå, tvetydige, stødende eller irriterende.

6. Endelig version

Efter tilretning, ud fra patient-feedback, har man nu et færdigt dansk instrument, som bør udsættes for psykometrisk testning, mere herom senere.

8.3. **Udvikling af nye PRO-instrumenter**

Udvikling af et nyt PRO-instrument er et omfattende arbejde, der kræver viden og ressourcer, hvis man påtænker at konstruere et instrument bestående af flere spørgsmål og en samlet score. Det forudsætter et stort metodearbejde, og falder uden for formålet for nærværende oversigt.

Det kan dog, som nævnt, være relevant at udvikle og tilføje et eller flere enkelt-spørgsmål som et supplement til eksisterende instrumenter (ad hoc-spørgsmål). På den måde kan man på en overkommelig måde opbygge et spørgeskema, som dækker de ønskede domæner, selvom der ikke kan beregnes scores, som inddrager de nyudviklede ad hoc-spørgsmål (5;45).

Udviklingsprocessen for nye spørgsmål kan opstilles i følgende trin:

1. Overordnede overvejelser og identifikation af potentielle domæner ved inddragelse af kliniske eksperter og patienter.

2. Litteratursøgning

Der foretages indledningsvis en litteratursøgning for at afdække om der findes relevante instrumenter og for at identificere studier om patienternes oplevelser med at leve med sygdommen. Disse kan give idéer til, hvad de nye spørgsmål bør indeholde. Ved ønske om belysning af et enkelt symptom (f.eks. kvalme) kan man gennemgå systematiske symptomspørgeskemaer som f.eks. Symptom Check List (SCL-92).

3. Inddragelse af kliniske eksperter

Herefter kan klinikere inddrages, og give deres bud på, hvilke relevante domæner de nye spørgsmål bør dække. De kan begynde at udforme punkter eller spørgsmål, som ønskes inddraget.

4. Inddragelse af patienter

Patienterne inddrages for at bidrage med deres mening om, hvad der bør indgå, og eventuelt vurdering af klinikernes udkast. Dette punkt kan gennemføres som individuelle interviews eller som fokusgruppeinterview.

5. Udformning af spørgsmål

Herefter udformes udkast til spørgsmål og svarkategorier. Spørgsmålene bør være korte, klart formulerede, lette at forstå, utvetydige og nemme at besvare. Hvert spørgsmål må kun indeholde én dimension. Det bør overvejes hvilke svarkategorier, der er relevante og det bør sikres, at svarkategorierne er udtømmende. Man bør ikke indbygge kausale formodninger i spørgeskemaet.

Der kan hentes inspiration i allerede udviklede spørgsmål, både hvad angår spørgsmålets formulering og svarkategorier. Hvis spørgsmålet omhandler symptomer kan man med fordel anvende samme svarmuligheder som Symptom Check List (SCL-92): ”Slet ikke”, ”Lidt”, ”Noget”, ”En hel del”, ”Særdeles meget”.

6. Pilottest og kognitive interviews

Det færdige udkast bør herefter testes på en bred gruppe af patienter med forskellige karakteristika ift. køn, alder, uddannelsesmæssig baggrund, sværhedsgrad af sygdom/ symptomer mv.

Pilottesten kan foretages ved at bede patienten tænke højt, mens patienten udfylder instrumentet, og samtidig spørge til forståelse mv. undervejs (kognitivt interview). Der kan afslutningsvis spørges til svære spørgsmål, relevans og manglende emner.

7. Endelig version

De forståelsesproblemer som pilottesten har afdækket, bør rettes til i spørgeskemaet. Der kan være behov for at omformulere, slette eller tilføje spørgsmål, inden der foreligger en endelig version.

Konklusion:

Det anbefales at starte med at undersøge muligheden for at anvende et eksisterende instrument, da der altid vil være fordele ved at anvende et instrument, der er afprøvet, og hvor der evt. findes referencedata.

Udvikling af nye instrumenter kan kun anbefales på enkelt-spørgsmålsniveau.

9. Krav til egenskaber ved PRO-instrument

Det er nødvendigt med kendskab til de centrale begreber til beskrivelse af et instruments psykometriske egenskaber, for at kunne bedømme dets anvendelighed. Begreberne er kort gennemgået i det følgende. For en grundigere gennemgang henvises til litteraturen (4;5;45;48;49)

9.1. Validitet

Begrebet validitet (5;49) omhandler i hvilken grad et instrument måler det fænomen (f.eks. livskvalitet eller funktionsniveau), som det foregiver at måle. Validitet kan inddeles i en række underbegreber, som hver især beskriver forskellige aspekter af et instruments validitet.

For de anerkendte PRO-instrumenter foreligger der typisk en række valideringsundersøgelser, som det er hensigtsmæssigt at tage i betragtning, når man skal vælge hvilket instrument, man ønsker at bruge. Ved vurdering af disse undersøgelser skal det medtænkes, hvorvidt den foretagne validering kan overføres til den kontekst, som man ønsker at anvende instrumentet i. Mulighed for overførsel kan f.eks. være lav, hvis instrumentet er valideret på en ældre patientpopulation, der har levet længe med en sygdom, og man ønsker at anvende instrumentet på en ung population, der netop er blevet diagnosticeret. Validitet er sjældent et spørgsmål om ja eller nej, men en afvejning af fordele og ulemper.

Content validity (indholdsvaliditet)

Content validity (4;5;45) omhandler, hvorvidt instrumentet måler alle relevante aspekter af det fænomen, som det skal måle. Instrumentets spørgsmål skal således dække alle relevante aspekter (se også face validity).

PRO-instrumenter skal reflektere det, som er væsentligt for patienterne. Content validity undersøges derfor bedst med kvalitative undersøgelser, f.eks. interviews med patienter og eksperter på området, i forbindelse med udvikling af instrumentet.

Face validity (overfladevaliditet)

Face validity (5;45) er i praksis en del af content validity og refererer til, om instrumentet umiddelbart ser ud til at handle om de ønskede emner på en klar og utvetydig måde. Ser spørgsmålene ud til at spørge om det, der ønskes undersøgt, giver spørgsmålene mening, er de forståelige og relevante og er svarkategorierne dækkende og gensidigt udelukkende? Face validity har derfor stor betydning for respondenternes oplevelse af relevans, og kan dermed få betydning for responsraten. Face validity er som nævnt, tæt knyttet til content validity, men hvor begrebet content validity primært anvendes ved udvikling og dokumentation af instrumentet, anvendes begrebet face validity primært ved vurdering i forhold til det konkrete valg af et instrument.

Construct validity (begrebsvaliditet)

Mange af de fænomener, som undersøges med PRO-instrumenter, er latente strukturer (f.eks. angst og træthed), som ikke er direkte observerbare og målelige, da det netop handler om patientens oplevelse. Construct validity (5;45) handler om at undersøge, om de resultater instrumentet giver, repræsenterer den underliggende latente variabel, f.eks. angst eller træthed. Dette kan undersøges ved at opstille en teoretisk model/ hypoteser på baggrund af eksisterende teorier. Hypoteserne skal beskrive de aspekter, som PRO-instrumentet har til hensigt at måle, og disse aspekters indbyrdes forhold. Herefter indsamles data mhp. en vurdering af, om disse hypoteser kan bekræftes. Hypoteser kan testes med korrelationsanalyser (test af om 2 fænomener, der ifølge teorien bør være korrelerede, også er det), ændring over tid (f.eks. skal fænomener, der ifølge teorien ændres over tid, kunne måles, hvis instrumentet anvendes på den samme population med et passende tidsinterval) og forskelle mellem kendte grupper, også benævnt discriminative validity (der skal kunne måles en forskel, hvis man f.eks. anvender instrumentet på en gruppe med og uden erkendt angst).

Criterion validity (kriterievaliditet)

Dette involverer en vurdering af instrumentet i forhold til en "guldstandard", der kan opfattes som den sande værdi (5;45). Det vil typisk indebære, at instrumentet holdes op imod et eksisterende og velafprøvet instrument indenfor samme område eller, især på det psykiatriske område, holdes op imod resultatet af et struktureret psykiatrisk interview. Hvis man skal vælge mellem to instrumenter, der måler det samme – f.eks. det lange og komplekse SF36 og den kortere version SF12, er det relevant at se på studier af criterion validity. Til dette vil man typisk have anvendt de to skemaer på samme patientpopulation og opgjort korrelationen mellem resultatet i det nyudviklede og det eksisterende instrument, eller opgjort det nye instruments sensitivitet og specificitet i forhold til guldstandard (det eksisterende skema).

9.2. Reliabilitet (pålidelighed)

Reliabilitet eller pålidelighed (5;45;49) handler om variation og præcision i målinger. Vurdering af reliabilitet består derfor i at vurdere, om instrumentet giver reproducerbare og stabile resultater. Instrumentet bør give reproducerbare eller ens værdier, hvis det anvendes flere gange på samme patient, under forudsætning af at patientens tilstand er uforandret. Dette kaldes test-retest reliabilitet. Det er vigtigt at vide, om et instrument giver præcise værdier for alle typer af patienter, både dem der scorer højt og lavt på et instrument, for at vurdere om det er anvendeligt i den aktuelle patientpopulation.

9.3. Responsiveness (følsomhed)

Responsiveness (5) beskriver instrumentets evne til at påvise ændringer, når patientens tilstand forbedres eller forværres. Ofte vil man være interesseret i at se effekten af en given behandling ved at lade patienten svare på PRO-instrumentet før og efter behandlingen. Instrumentet skal være tilstrækkeligt følsomt til at

kunne opfange og vise en ændring, når der faktisk er sket en ændring i patientens tilstand.

Flere faktorer kan have betydning for instrumentets responsivens. Et instrument med subdomæner, der er mindre relevante for den aktuelle patientgruppe, vil ikke ændre sig over tid som en effekt af behandlingen. Et meget begrænset antal svarkategorier kan også medføre, at det kan være vanskeligt at påvise mindre ændringer i patientens tilstand. Der bør desuden være opmærksomhed på hvorvidt det tidsinterval, der spørges til, giver mulighed for at detektere en ændring (hvis der spørges til tilstanden det sidste år, vil man ikke kunne påvise ændringer med kortere interval).

9.4. Floor-/ceiling effect (gulv- og lofteffekt)

Floor-/ceiling effect (5) skal ses i sammenhæng med responsivens. Ceiling effect henviser til, at en stor andel af det samlede antal besvarelser opnår den maksimale score, eller har sat kryds i den højest mulige svarkategori. Modsat refererer floor effect til, at en stor andel af patienterne opnår den minimale score. Begrebet er derfor snævert knyttet til den aktuelle population. Hvis en stor andel af patienterne svarer i de ekstreme kategorier, vil det ikke være muligt at påvise yderligere bedring eller forværring. Floor-/ceiling effect medfører nedsat kontrast og reducerer responsivens. Det må dog medtænkes, at nogle spørgsmål ikke nødvendigvis bør ændres, selvom der er påvist floor-/ceiling effect. Eksempelvis kan det være relevant at spørge til et sjældent, men potentielt alvorligt symptom, selvom det vil udvise floor effect.

9.5. Acceptability

Acceptability (48;49) refererer til, hvorvidt patienterne finder det acceptabelt at besvare instrumentet. I praksis kan en række forhold betyde, at patienterne ikke besvarer instrumentet som ønsket. Der kan være tale om forhold ved selve instrumentet, ved patienterne som kan være for syge, eller ved omstændighederne i øvrigt. Problemerne kan bl.a. vise sig ved en lav responsrate og manglende svar på enkelte spørgsmål. Følgende kan eksempelvis påvirke et instruments - og dermed hele spørgeskemaets - acceptability:

Lav umiddelbar relevans af spørgsmål og svarkategorier (face validity)

- Forståelsesproblemer i spørgsmålene
- Lange recall perioder
- Stort tidsforbrug ved besvarelse
- Følsomme spørgsmål placeret i begyndelsen af spørgeskemaet

Konklusion:

Ved valg af PRO-instrument skal der indhentes dokumentation for validitet og reliabilitet. Det skal medtænkes i hvilken kontekst instrumentet er testet, og om disse forhold er sammenlignelige med de omstændigheder instrumentet skal anvendes under i forbindelse med kvalitetsovervågning. Samtidig skal det overvejes, om de instrumenter der vælges, kan identificere ændringer i patienternes tilstand og dermed dokumentere, hvorvidt der sker en forbedring/forværring i patientens tilstand over tid.

10. Definition af inklusions-/eksklusionskriterier før indsamling af PRO-data

Inklusions- og eksklusionskriterier skal være klart definerede, og metoden til identifikation af patienter skal være præcist beskrevet. Identifikation af patienter kan ske i klinikken, hvor inklusionskriterier kan være defineret ift. bestemte procedurer, fx en henvisning til operation eller inklusion kan ske på baggrund af registerbaserede oplysninger, evt. fra selve kvalitetsdatabasen. Det sidste er normalt at foretrække, men forudsætter at patienterne er registreret i kvalitetsdatabasen på tidspunktet, hvor PRO ønskes indhentet. Endvidere skal data som anvendes til identifikation af patienten være valide. Det kan have store konsekvenser at kontakte patienter, som er fejlregistrerede med en given diagnose. Validiteten af inklusionsdata bør derfor være grundigt undersøgt, inden man kontakter større patientgrupper på dette grundlag.

De kliniske kvalitetsdatabaser omfatter normalt alle patienter i den pågældende diagnosegruppe, blandt andet fordi det kan være logistisk vanskeligt, at udpege hvilke patienter der skal indsamles kliniske data på. Det kan overvejes, om det for større afdelinger vil være tilstrækkeligt at indsamle PRO-data på en stikprøve, hvor stikprøvens størrelse er fastlagt på baggrund af statistiske styrkeberegninger. Det er muligt at anvende forskellige samplingmetoder, f.eks. tilfældigt udtrækning af en stikprøve eller stratificeret sampling, med højere sandsynlighed for udtrækning af patienter med særlige karakteristika, såfremt dette karakteristikum er registreret i databasen.

Konklusion:

Der skal tages stilling til, om alle patienter i den pågældende diagnosegruppe skal inkluderes eller der skal udvælges en stikprøve.

11. Layout for spørgeskema og generel vejledning til patienterne

Uanset om et spørgeskema er papirbaseret eller webbaseret er det vigtigt, at patienten orienteres om formålet med spørgeskemaet, hvad data anvendes til, hvem der har adgang til data, at persondataloven overholdes og hvordan datasikkerheden opretholdes. Patienten kan gøres opmærksom på PRO-indsamlingen allerede i indkaldelsesbrevet eller ved udskrivelse fra afdelingen.

Der bør være en let forståelig instruktion i, hvordan skemaet skal udfyldes. Anvendes der papirbaserede skemaer, bør der desuden være eksempler der viser, hvordan man skal markere sit svar. Det er hensigtsmæssigt, at det beskrives hvor lang tid det ca. tager at besvare spørgeskemaet, hvad man gør, hvis man ikke ønsker eller ikke kan svare, og hvordan man afleverer et besvaret skema. Patienter gør sig ofte overvejelser om, hvorvidt de tilhører målgruppen, og patienter med få symptomer er undertiden tilbøjelige til at undlade besvarelse, fordi de fejlagtigt mener, at dette vil medvirke til at give et misvisende billede. Det samme er tilfældet for patienter, som har en anden lidelse, som i højere grad influerer på deres almene tilstand. Man kan søge at imødegå disse kilder til selektionsbias ved at understrege at alle patienters besvarelser er lige vigtige.

Der findes studier, der har undersøgt, hvilken effekt længden på spørgeskemaet har for responsraten. Resultaterne er dog ikke entydige (50). Nilsson har eksempelvis undersøgt patienters præferencer for EQ-5D i forhold til SF-36. Selvom EQ-5D er væsentligt kortere (5 spørgsmål) end SF-36 (36 spørgsmål) fandt de fleste patienter begge instrumenter lige anvendelige, mens 25% foretrak SF-36 mod 8% der foretrak EQ-5D (51).

Villigheden til at besvare et spørgeskema afhænger ikke kun af længden, men synes primært at afhænge af, om patienterne oplever at spørgsmålene er forståelige og relevante i forhold til deres aktuelle situation. Det bør overvejes om patienterne kan få adgang til egne data via sundhed.dk, og om de kan få en automatisk tilbagemelding på deres besvarelse kombineret med referenceværdier. En sådan tilbagemelding vil øge patientens udbytte af besvarelsen og formentlig forbedre besvarelsesprocenten. Et indbydende layout med en skriftstørrelse og illustrationer, der er tilpasset målgruppen, kan have positiv indflydelse på besvarelsesprocenten. Hvis PRO besvares på hospitalet, bør det sikres at patienten har tilstrækkelig tid, og der bør være et afskærmet område til rådighed, hvor patienten kan udfylde spørgeskemaet (7).

12. Pilotfase (field test)

Hvis der ikke foreligger danske erfaringer med de konkrete PRO-instrumenter i den pågældende patientgruppe, anbefales det at indlægge en pilotfase, inden der iværksættes indsamling af PRO på alle patienter, der skal kvalitetsmonitoreres. I en pilotfase kan både patienternes villighed til at besvare spørgeskemaet og logistiske problemer afdækkes. Pilotfasen vil give indikationer på, hvordan administration af skemaer, datahåndtering og arbejdsgange fungerer og hvilke faldgruber der måtte være (5). Derudover vil man få mulighed for at vurdere de valgte PRO-instrumenters egenskaber i den specifikke kontekst.

Det vil først være muligt at få indikationer på, om patienterne vil besvare spørgeskemaet, når man når til pilotfasen. En pilotfases størrelse vil normalt være begrænset, men med kvalitative og semi-kvantitative metoder vil man kunne få værdifuld viden om nedenstående temaer.

Pilotfasen kan således afdække følgende perspektiver:

- Patient:
 - Kan patienterne finde ud af at besvare spørgeskemaet i den valgte administrationsform?
 - Vil patienterne besvare spørgeskemaet?
 - Er der spørgsmål som mange patienter ikke besvarer?
- Logistik:
 - Identificeres de rette patienter, som instrumentet skal sendes til?
 - Hvordan fungerer det med administration af spørgeskema og evt. rykkerprocedure?
 - IT infrastruktur
 - Arbejdsgange på afdelingerne

Konklusion:

En pilotfase er tilrådelig.

13. Hvornår skal PRO indsamles?

Såvel formålet med indhentning af PRO-data, som en række ydre omstændigheder, har betydning for hvornår PRO-data kan indsamles.

Foreligger der en naturlig baselinetilstand, fx før en operation, vil man ofte være interesseret i at indsamle PRO både før og efter en given behandling med henblik på longitudinelt at kunne beskrive intraindividuelle før-efter forskelle. Praktiske forhold omkring tidsmæssige muligheder og mulighed for identifikation af patienterne skal overvejes. Den konkrete tidslinje for follow-up vil afhænge af sygdommen, og det kan,

afhængigt af lidelsens forløb, være relevant med mere end et follow-up tidspunkt, f.eks. hvis tidspunktet for forventet steady state er variabelt.

Kriterierne for indhentning af PRO-data skal beskrives nøje. Disse kriterier danner baggrund for overvejelser om, hvornår og med hvilke metoder PRO-data kan indsamles, hvordan patienten kan kontaktes mhp. at besvare PRO, og hvordan eventuelle rykkerprocedurer kan organiseres.

Kriterierne for indhentning af PRO er afgørende for, om PRO bør indsamles decentralt på den enkelte afdeling (klinikbaseret PRO-indsamling) eller med centraliseret PRO-indsamling, hvor patienten kontaktes i hjemmet.

Klinikbaseret PRO-indsamling

Kriteriet for indhentning af PRO kan tage udgangspunkt i en procedure i klinikken, fx at patienten møder op i et ambulatorium, at patienten får stillet en bestemt diagnose eller patienten bliver henvist til en bestemt undersøgelse eller operation. I disse tilfælde vil den eneste mulighed oftest være umiddelbart at bede patienten besvare PRO på hospitalet.

I nogle kliniske databaser forsøger man at udlevere PRO-spørgeskemaer på hospitalet i forbindelse med den første kontakt, med henblik på at opnå en baselinemåling inden behandlingsstart. Det viser sig i praksis vanskeligt at få dette set-up til at fungere i hverdagen. Hvis PRO-data indsamles på hospitalet, og kun anvendes på gruppeniveau, pålægges personalet en ekstra opgave, som forudsætter både ansvar og vedvarende engagement.

Det forholder sig anderledes, hvis PRO-data samtidigt anvendes på individniveau. PRO bliver herved en del af patienthåndteringen, og der er erfaringer fra Danmark (DanBIO, AmbuFlex), som viser, at det er muligt at indsamle PRO med høje responsrater i ambulatoriet med brug af tablets eller berøringsfølsomme skærme. Det er dog en forudsætning, at der er personale tilstede til at opfordre patienterne til at svare og vejlede dem ved tvivlsspørgsmål.

Alternativt kan man forsøge at få patienten til at besvare et skema forud for ankomsten, f.eks. i forbindelse med modtagelsen af indkaldelsen. Der kan dog være behov for at personalet individuelt undersøger, om patienten har besvaret spørgeskemaet med henblik på at anmode nonrespondenter om at besvare på afdelingen.

Det er i modsætning til centraliseret PRO-indsamling (se nedenfor) i praksis ikke muligt at gennemføre egentlige rykkerprocedurer ved klinikbaseret dataindsamling.

Centraliseret PRO-indsamling

Kriteriet for indhentning af PRO kan bestå af en algoritme på baggrund af registerbaserede data om patienten. Det kan fx være, at PRO ønskes indsamles et antal dage/uger/måneder efter en registreret diagnosedato eller behandlingsdato, fx follow-up 12 måneder efter en operationsdato. I disse tilfælde er PRO-indsamling i hjemmet oftest reelt eneste mulighed.

Centraliseret PRO-indsamling giver muligheder for høje responsrater og stordriftsfordele, samt mulighed for at sikre at PRO-indsamlingsprotokollens fastsatte tidslinje overholdes nøje, herunder en konsekvent løbende rykkerprocedure.

Det er et særskilt problem at anmode samme patient om at udfylde flere relativt ens spørgeskemaer til klinisk brug og til kvalitetsdatabaser. Dette kan medføre utilfredshed og reducere responsraten. Det er derfor vigtigt at overveje håndtering af dette problem. Indsamles PRO ved ambulante kontrolbesøg, bør man derfor overveje, hvorvidt man kan anvende disse PRO-data også til kvalitetsformål, og kun kontakte de patienter, som af forskellige årsager ikke møder til kontrol. Dette forsøges i øjeblikket for patienter med lungekræft (PROLUC 1+2). Det forudsætter dog, at man kan identificere disse patienter, og det vil være vanskeligt at følge den samme tidslinje som ved konsekvent indhentning af PRO i hjemmet.

Konklusion:

Ved indsamling af PRO-data skal der tages stilling til:

- Hvor skal PRO-data indsamles - på hospitalet eller hjemme hos patienten?
- Hvornår skal PRO måles?

14. Kontakt- og besvarelsesmetoder

Beslutninger om kontakt- og besvarelsesmetoder til PRO-dataindsamlingen har vital betydning for kvaliteten af data, organisering af dataindsamlingen og for økonomien.

Da den valgte besvarelsesmetode er afgørende for kontaktmetoden, beskrives denne først.

14.1 PRO-besvarelsesmetoder

Indsamling af PRO-data kan opdeles i følgende metoder:

1. Papirbaseret
2. Internetbaseret
 - a. Webspørgeskema
 - b. Tablet/patient-kiosk
 - c. App downloadet af patienten
3. Interviewbaseret
 - a. Telefoninterview
 - b. Face-to-face interview
 - c. Automatiseret
4. Mixed-mode besvarelsesmetoder, som kombinerer to eller flere af ovenstående

Papirbaseret PRO-besvarelse

Historisk er dataindsamling med papirspørgeskemaer den mest anvendte metode. Det er en metode, som patienten kender, og kan anvendes i alle aldersgrupper. Ulemper er den økonomiske byrde til dækning af udgifter til trykning, adressering og kuvertering samt modtagelse og håndtering af spørgeskemaerne, hvor data efterfølgende skal bringes i en elektronisk database. Sidstnævnte kan ske manuelt, som i så fald bør ske med dobbelt indtastning med henblik på at reducere simple tastefejl (49). Ved højt volumen og kontinuerlig dataindsamling er dette uhensigtsmæssigt og tidskrævende. Alternativet er scanning med efterfølgende OCR (Optical Character Recognition). Dette forudsætter et kompliceret set-up med hensyn til software, konfiguration, verifikation af tvivlagte afkrydsninger og løbende kvalitetskontrol (7).

Tekstfelter er særligt problematisk ved papirskemaer, da OCR-software ikke håndterer håndskrevne bogstaver med tilfredsstillende resultat, hvilket nødvendiggør en løbende verifikationsproces.

I et papirskema er der ikke noget, der hindrer patienten i at sætte kryds mellem rubrikker eller skrive kommentarer i margin. Dette kan være problematisk, hvis patienten forventer at sundhedspersonale følger op på den specifikke besvarelse, f.eks. ved at indkalde patienten til en ny undersøgelse.

Internetbaseret PRO-besvarelse

Internetbaseret PRO-indsamling kan ske via et website, som patienten tilgår direkte, via en webportal som f.eks. sundhed.dk, på berøringfølsomme skærme på hospitalet (patient-kiosk) eller som en mobil app (i realiteten oftest blot et website). Fejl ved dataindtastning er her minimeret (7), og patienten kan ikke sætte krydser udenfor felterne, og kan endvidere mindes om at besvare alle spørgsmål, inden der fortsættes til næste side i spørgeskemaet. Patienten kan dog med forsæt have undladt at besvare et bestemt spørgsmål, og man bør ikke ubetinget forsøge at tvinge patienten til at svare. Dette kan medføre, at patienten afbryder besvarelsen, og man reducerer desuden sammenligneligheden med PRO-data indsamlet med papirskema. Webaserede skemaer kan gøres brugervenlige ved at tilføje illustrationer, hjælpeknapper med ekstra information samt tilpasse antallet af spørgsmål til den enkelte patient. F.eks. vil en patient, som har angivet aldrig at have røget, ikke blive præsenteret for yderligere spørgsmål om rygning. Computerized Adaptive Testing (CAT) (se tidligere) er en systematisk metode til anvendelse af individuelt tilpassede spørgeskemaer. Denne metode kan kun anvendes ved internetbaserede indsamlingsmetoder.

Det er nemt at tilføje ekstra spørgsmål eller et ekstra instrument til et allerede etableret internetbaseret dataindsamling. Via elektronisk dataindsamling opnår man samtidig præcise data om svartidspunkt, og data kan gøres tilgængelig med det samme (49).

Ved nyreinsufficiens samt prostata- og lungecancer og på det reumatologiske område viser danske erfaringer, at det er muligt at få patienter til at besvare et webaseret spørgeskema i klinikken på tablets eller fastmonterede ”patient-kiosks” i ambulatoriet (52). Denne procedure forventes udbredt, og det kan være en fremtidig kanal til indsamling af PRO-data fra andre patientgrupper inden behandlingsstart. Dette forudsætter dog, at juridiske spørgsmål om anvendelse af kvalitetsdata på patientniveau tages i betragtning, dels at det bliver praktisk muligt at udveksle data mellem databasen i opsamlingssystemet og den kliniske database.

Ved internetbaseret PRO-indsamling viser erfaringer, at man skal være forberedt på praktiske it-mæssige udfordringer på de enkelte afdelinger, bl.a. i form af banale problemer som internetadgang. Denne type problemer er ofte vanskelige at løse på afstand. Alternativt kan man overveje en mere robust, men også dyrere løsning, såsom trådløst internetforbindelse udenom de lokale netværk. Driftsudgiften ved selve dataindsamlingen er relativt lav, men der må påregnes personaleressourcer til instruktion af patienter og besvarelse af tvivlsspørgsmål.

Interviewbaseret PRO-besvarelse

En fordel ved interviewbaseret PRO-dataindsamling er, at man kan være sikker på, hvem der svarer på spørgsmålene, og at der kan opnås en høj responsrate. Desuden er det en fordel for patienter, som har

sproglige problemer. Der bør stilles krav til at patientinterviewerne skal være oplærte, bl.a. med henblik på at imødegå interviewer-bias (7).

Interviewbaseret PRO-dataindsamling er tidskrævende og omkostningstungt. I praksis vil man sjældent vælge interviewbaseret dataindsamling som den primære metode i en kontinuerlig driftsopgave, som det er tilfældet med en klinisk kvalitetsdatabase. I en mixed-mode strategi (se senere) kan telefoninterviewbaseret dataindsamling dog overvejes som en ultimativ sidste rykkerprocedure, og ved centraliseret drift kan udgiften reduceres, og besvarelsen kan indtastes i en eksisterende webformular.

Automatiserede computerstyrede 'interviews', f.eks. Interactive Voice Response (IVR), hvor patienten ringes op og besvarer indtalte spørgsmål ved tale eller tastetryk, er udbredt i bl.a. USA, men er så vidt vides ikke forsøgt anvendt i Danmark.

Mixed-mode PRO-besvarelsesmetoder

Når der lægges vægt på at opnå en høj responsrate, og man samtidig ønsker at udnytte fordelene ved internetbaseret dataindsamling, kan det være hensigtsmæssigt at kombinere dataindsamlingsmetoderne. Dette kaldes mixed-mode, og kan enten foregå som et fast set-up f.eks. primært webspørgeskema med et tilsendt papirbaseret skema ved rykkerprocedure, eller hvor patienterne selv frit vælger, om de ønsker at svare på papir eller internet. Mixed-mode omfatter både besvarelsesmetode og kontaktmetode.

Det har været diskuteret, hvorvidt instrumenter, der er udviklet til papiradministration, bør gennemgå ny valideringsproces, inden skemaet anvendes elektronisk. Et oversigtsartikel fra 2014 har gennemgået litteraturen på området, og finder ikke evidens for dette, såfremt opsætningen er så ens som muligt i de to metoder, og med mindre særlige forhold indikerer andet (53;54). Som udgangspunkt kan man derfor blande disse to PRO-dataindsamlingsmetoder vilkårligt. Hvis mixed-mode inkluderer interview, for eksempel som sidste rykkerprocedure, er der ikke nogen generel dokumentation for sammenligneligheden. Her bør eventuel dokumentation for sammenlignelighed opsøges specifikt for de enkelte redskaber, og en mulig reduceret sammenlignelighed skal holdes op mod fordele ved den højere responsrate, man måtte opnå herved.

14.2 Kontaktmetode

Første skridt ved indsamling af PRO-data er at kontakte de relevante patienter. Kontaktmetoder omfatter:

1. Personlig henvendelse
2. Elektronisk ankomstregistrering
3. Postomdelt brev
4. Digital post / eBoks
5. E-mail
6. SMS
7. App-notifikation

Kontaktmetoderne er forskellige med hensyn til ressourceforbrug, forudsætninger, logistik og økonomi.

Personlig kontakt og elektronisk ankomstregistrering

Personlig kontakt ved patientens fremmøde i ambulatoriet omfatter opfordring og instruktion i udfyldelse af PRO-spørgeskemaet. Metoden forudsætter, at det er gjort til en del af afdelingens rutine, og patienterne opfordres og instrueres deri. I disse år sker der en overgang til elektronisk ankomstregistrering, hvor patienten scanner sit sundhedskort. Såfremt nogle indprogrammerede betingelser er opfyldt (fx en henvisningsdiagnose), vil patienten automatisk kunne modtage en instruks på skærmen om at udfylde PRO-spørgeskema.

Når kriteriet for indsamling af PRO omfatter en hændelse, som indtræffer efter ankomsten, fx en henvisning eller en diagnose, f.eks. med henblik på at opnå baselinedata, er det vanskeligere at indpasse PRO-indsamlingen i afdelingens rutiner. Erfaringen med indsamling af PRO-data på denne måde er ikke gode, bortset fra tidsafgrænsede forskningsprojekter. Der er ofte lav responsrate med geografiske variationer, hvilket reducerer anvendelsesmulighederne væsentligt.

Postomdelt brev

Besvarelse i hjemmet forudsætter en kanal til at transmittere opfordringen til patienten om at besvare PRO. Dette kan ske ved indkaldelsen, som en del af indkaldelsesproceduren, eller som en selvstændig henvendelse ved follow-up. Kontakt til patienten med postomdelt brev er den metode, der er længst erfaringer med.

Metoden er sikker og effektiv og er eneste mulighed, når der anvendes papirbaseret PRO. De eneste oplysninger som kræves, er et CPR-nummer samt en dato for hvornår udsendelsen skal ske.

Adresseoplysninger kan herefter skaffes fra CPR-registeret. Ressourceforbruget ved anvendelse af post er mindst, når brevet blot indeholder en instruktion i at besvare spørgeskemaet på et website. Indeholder brevet derimod både følgebrev, papirskema og returkuvert, er der udgifter forbundet med returnering og håndtering af papirskemaet. Papirbrevet udfases nu i de offentlige kommunikation med borgerne. Det må dog vurderes

at en stor og vigtig restgruppe fortsat kun kan nås med traditionel post. Postomdelte breve vil også have en central rolle i et mixed-mode set-up.

Digital post/eBoks

Digital post er en ny kanal til at nå patienterne, og som ved postomdelt brev er patientens CPR-nummer eneste nødvendige identifikation. Afsendelse af digital post er i modsætning til e-mail ikke gratis, men er dog prissat betydeligt under den tilsvarende brevporto.

Digital post anvendes nu som eneste kommunikation i forbindelse med indkaldelse i sundhedsvæsnet, såfremt borgeren ikke har fået dispensation til fortsat at modtage papirindkaldelser. Webbaseret besvarelse er den oplagte besvarelsesmetode ved digital post, hvor der kan sættes et klikbart link til websitet ind i det elektroniske brev. Metoden er anvendelig som førstevalg ved PRO-indsamling, og også ved follow-up.

Det er grundlæggende frivilligt om patienten vil besvare PRO. Det forhold at patienter reagerer på en indkaldelse til et hospital via digital post, betyder ikke nødvendigvis, at patienter også besvarer et spørgeskema leveret ad samme kanal. Der mangler konkrete erfaringer hermed.

Udsendelse af digital post sker i nogle set-ups ved en manuel procedure til en patient ad gangen, hvilket kan fungere ved indkaldelse på patientniveau. Løbende udsendelser fra en klinisk kvalitetsdatabase på baggrund af registerbaserede data, skal dog kunne ske automatiseret, hvilket stiller teknologiske krav. Ved anvendelse af digital post kan man i nogle tilfælde ikke frit bestemme afsenderinformationen, hvilket kan medføre forvirring for patienten, hvis denne ikke umiddelbart kan associere afsenderen til sit aktuelle hospitalsforløb. Man skal endvidere sikre de nødvendige tilladelser fra den afsendende myndighed.

E-mail

En e-mail kan indeholde et link til webbesvarelse og kan også bruges ved rykkerprocedurer. Metoden forudsætter at e-mailadressen er kendt. Ved PRO-indsamling som indbefatter flere runder, kan denne information indhentes i første skema. I papirskemaer kan e-mailadresser oftest ikke aflæses korrekt, hverken maskinelt eller ved manuel indtastning. Metodens fordele ved lav pris og hurtighed, modsvares delvist af at e-mailadresser skifter jævnligt, og er følsomme for opsætning af lokale pc, spamfiltre, nedbrud og internetforbindelsesproblemer hos patienten.

SMS

Kontakt via SMS forudsætter kendskab til mobiltelefonnummer samt en adgang til afsendelse. Metoden kan næppe anvendes som primær kontaktmetode, men eventuelt som et element i rykkerprocedurer.

App-notifikation

App-notifikation er medtaget for fuldstændighedens skyld, og har mest relevans såfremt dataindsamlingen foregår via samme metode, hvilket sjældent vil være tilfældet ved PRO-indsamling til kvalitetsdatabaser.

Mixed-mode kontaktmetode

De beskrevne kontaktmetoder kan blandes, således at den hurtigste og/eller mindst ressourcekrævende metode bruges primært, og den dyreste og/eller langsomste ved den sidste rykkerprocedure.

14.3. Rykkerprocedurer

Rykkerprocedurer er essentielle for at nå en tilfredsstillende responsrate. Erfaringer fra flere områder viser, at der kan opnås tilfredsstillende resultater, også selvom data ikke anvendes i forhold til den enkelte patient. Danske erfaringer er opstillet i tabel 7. Der er en vis variation mellem diagnosegrupperne, men alle på nær apopleksi, ligger over 85%. Disse responsrater er højere end hvad man normalt ser i spørgeskemaundersøgelser i både generelle - og patientpopulationer.

Tabel 7. Responsrate ved PRO-dataindsamling til kvalitetsudviklingsformål

Diagnosegruppe	Tidspunkt	Periode	Patienter	Rykker procedurer	Responsrate
Prostatacancer ^a	12 mdr efter diagnose	2011-	9574	2	90.7% ^a
Hofte- og knæ-alloplastik ^b	Post-operativt	2011-12	722	2	87.2%
Nyrecancer	12 mdr efter diagnose	2012	816	2	88.4%
Apopleksi ^c	3 mdr post-stroke	2012-14	3887	2	73.6%
Lungecancer	3mdr efter terapi-start	2014-15	3997	2	85.4% ^a

a) Foreløbigt estimat, da PRO-dataindsamling fortsat pågår.

b) Udvalgte afdelinger. Tidspunkt for follow-up varierede ift. afdelingernes praksis

c) Kun Region Midtjylland.

Note. Eksemplerne stammer fra VestKronik. I alle tilfælde var den primære metode et postomdelt brev med papirskema med mulighed for webbesvarelse, første rykkerprocedure et brev og anden et brev med et nyt papirskema.

Mixed-mode hører naturligt sammen med rykkerprocedurer, hvor man kan optrappe med mere ressourcekrævende metoder med antallet af rykkerprocedurer. Med introduktionen af digital post er det, som nævnt ovenfor, nu blevet muligt at anvende mixed-mode allerede ved første kontakt. Mixed-mode med hensyn til både kontaktmetode og besvarelsesmetode giver en mulighed for at opnå fordele ved f.eks. følgende kombination:

- 1) Primær henvendelse: Digital post med link til internetbesvarelse
- 2) Første rykkerprocedure: Postomdelt brev med opfordring til internetbesvarelse

3) Anden rykkerprocedure: Postomdelt brev med papirbaseret skema

Der er planlagt forskningsprojekter med henblik på belysninger af denne model, men der foreligger ingen resultater for nuværende.

Hvis responsraten ønskes yderligere løftet i hele eller dele af populationen, kan man overveje at anvende telefoninterview ved en afsluttende tredje rykkerprocedure (se dog tidligere vedrørende validitet).

Ved longitudinel PRO-indsamling kan information om mobiltelefonnummer eller e-mail indhentes i første PRO-skema, og efterfølgende anvendes ved udsendelse af SMS/ e-mail ved senere rykkerprocedurer.

I nedenstående tabel 8 er de mulige kontaktmetoders og PRO-dataindsamlingsmetoders egnethed vurderet for henholdsvis klinikbaseret og centraliseret PRO-dataindsamling.

Tabel 8. Mulige kontakt- og besvarelsesmetoder i forhold til klinikbaseret og centraliseret PRO-dataindsamling.

	Klinikbaseret PRO-indsamling			Centraliseret PRO-indsamling		
	Ved indkaldelse	Ved ankomst	Ved diagnose, henvisning OP, etc	Ved tid= t_1 Follow-up 1	Ved tid= t_n Follow-up n	Ved tid= t_r Rykkerprocedurer
Kontaktmetode						
Personlig kontakt (skranke)		●	○			
Elektronisk ankomstregistrering		●				
Postomdelt brev	●			●	●	●
Digital post/ eBoks	●			●	●	○
e-mail					●	●
SMS					●	○
App-notifikation					○	○
Besvarelsesmetode						
Papirskema	●	○	○	●	●	●
Webskema og portal (fx sundhed.dk)	●		○	●	●	●
Tablet/patientkiosk		●	○			
App (downloadet af patient)				○	○	○
Telefoninterview	○		○			○

●: egnet ○: evt egnet

Note: Såfremt der sker en tidstro registrering af f.eks. diagnoser vil en centraliseret dataindsamling også være mulig.

14.4. Responsrater ved internet- og papirbaseret besvarelsesmetode

Viden om hvilken responsrate man kan forvente ved de forskellige metoder og kombinationer i mixed-mode, er central ved beslutninger om planlægning af PRO-dataindsamling. I betragtning af hvor central en rolle indsamling af spørgeskemadata har inden for epidemiologisk forskning, findes overraskende få veldefinerede studier på området. Der findes en række studier i populationer, som er selekteret på web-brug, f.eks. hvor rekruttering af deltagere er sket på internettet. Responsraten ved webbesvarelse i disse studier kan ikke generaliseres til generelle patientpopulationer.

De foreliggende nyere randomiserede studier i veldefinerede populationer er opsummeret i tabel 9. Der forefindes endnu ikke systematisk opgjorte erfaringer ved indsamling af PRO ved hjælp af digital post/eBoks.

For single-mode er der betydelig variation i de absolutte tal, hvilket delvist kan tilskrives forskellige protokoller, f.eks. tidsinterval til første rykkerprocedure. I alle studier findes den højeste responsrate ved papirbesvarelsesmetoden (single-mode). I studiet af Hohwü et al, hvor forskellen er marginal, findes lav responsrate i begge modes. Studiet er et generelt survey, og papirversionen omfattede 28 sider. Forfatterne foreslår selv længden af spørgeskemaet som mulig årsag til den lave responsrate. Da omfanget er åbenlyst ved papirversionen, mens længden af et webskema vanskeligere lader sig bedømme umiddelbart, er det tænkeligt, at modtagere af denne version i højere grad straks har fravalgt at deltage. Studiet af Van den Berg et al fremhæves nogle steder som et af få studier, som finder højere responsrate i webgruppen. Imidlertid er 54 af de indgåede 83 besvarelser i webgruppen indgået efter rykkerprocedurer som papirbaserede besvarelser (83,3%). Studiet er således i højere grad et eksempel på et godt resultat af mixed-mode.

Responsraten efter mixed-mode varierer i mindre grad mellem studierne, og anvendelse af mixed-mode ser ikke ud til at reducere responsraten sammenlignet med en single-mode papirbaseret indsamlingsmetode.

Studierne har også analyseret determinanter for høj responsrate, men der tegner sig ikke et konsistent billede. Dette kan delvis tilskrives forskellige geografiske og kulturelle forhold, forskelle i populationer samt forskellige protokoller. Det bemærkes, at der ikke findes konsistent sammenhæng mellem alder og responsrate på webbaserede skemaer. Det har været antaget, at tiden arbejder for webbaserede svarmetoder, i takt med en øget udbredelse af internetforbindelser, samt at de ældre aldersgrupper erstattes af mere internetvante yngre grupper. Der er dog ikke nogen konsistent aldersgradient i forhold til responsrate ved internetbesvarelser, og da langt hovedparten af befolkning i dag har adgang til internet, må det konkluderes, at tilskyndelsen til webbesvarelse afhænger af andet end adgangen til internet. Foreslåede, men

udokumenterede forklaringer er en vis mætning i forhold til webbaserede spørgeskemaer generelt, og det forhold at det med tiden bliver eksklusivt at modtage et brev med posten.

Ved gentagende (kontinuerlig) dataindsamling er der erfaring for, at patienter efter start med papirskema efterfølgende kan bringes over på en internetbaseret metode, når de har fået kendskab til PRO-indsamlingen (52).

Tabel 9. Randomiserede studier af responsrate ved single-mode (web- eller papirskema som besvarelsesmetode) og mixed-mode (den alternative mode ved første rykkerprocedure)

Studie	Land	Population	N	Primær mode	Responsrate	
					Efter single-mode	Efter mixed-mode
Leece 2004 ^a (55)	Canada	Kirurger (medlemsregister)	442	Web	34,8%	63,8%
				Papir	51,1%	63,3%
Bälter 2005 ^b (56)	Sverige	Nationalt populations sample	875	Web	51,0%	n.a.
				Papir	64,3%	n.a.
Kongsved 2007 ^c (57)	Danmark	Kvinder henvist til mammografi	533	Web	17,9%	62,2%
				Papir	73,2%	76,5%
Zuidgeest 2011(58)	Holland	Kvinder med brystkræft	800	Web	32,0%	60,5%
				Papir	64,0%	n.a.
Van den Berg 2011 ^d (59)	Holland	Kvindelige overlevende efter børnecancer	277	Web	20,7%	59,3%
				Valgfrit	n.a.	65,7%
Hohwü 2013 ^e (60)	Danmark	Befolkningssample (forældre)	863	Web	53,4%	n.a.
				Papir	56,2%	n.a.

Note: Kun studier i populationer som ikke er selekteret ift mode-præferencer er medtaget.

a) Tre rykkerprocedurer. De to første rykkerprocedurer til web-gruppen bestod af e-mails.

b) Anden rykkerprocedure bestod af en telefonisk opringning.

c) Besvarelser modtaget efter dato for mammografi var ekskluderet. Kun tid til én rykkerprocedure.

d) Telefon som 2. rykkerprocedure. I arm med valgfrihed valgte 83% papir og 17% web-besvarelsesmetoden. I artiklen medregnes nægttere som respondenter, hvilket giver misvisende høje tal i begge grupper.

e) Ikke egentlig mixed-mode, men webgruppen fik ved to rykkerprocedurer mulighed for at bestille et papirskema. Det er ikke oplyst hvor mange, der benyttede denne mulighed.

Ved besvarelse af PRO på en afdeling må responsraten antages at være mindre afhængig af svarmetoden, og primært afhænge af hvorvidt udfyldelse af PRO er en del af afdelingens rutine. Dermed er internetbaserede metoder (tablets eller patient-kiosk) som regel at foretrække ved PRO-indsamling i klinikken.

Konklusion:

- Der skal tages stilling til om PRO-data skal indsamles papirbaseret, internetbaseret eller ved mixed-mode.
- Det skal afgøres, hvordan patienten kontaktes: personligt, via e-mail, digital post, brev eller sms.

15. Infrastruktur for dataindsamling

Kravene til infrastruktur afhænger af volumen samt den protokol man vil indsamle PRO efter. I den simpleste udgave kræves blot et dataindtastningsprogram eller et simpelt opsat webbaseret skema med en tilhørende webbaseret database. Begge dele kan anskaffes gratis, og administration kan foretages manuelt. Denne løsning kan være velegnet ved tidsafgrænsede projekter som f.eks. ved pilotafprøvninger af skemaer. Ved langvarig dataindsamling med blot et moderat volumen, er det ganske urealistisk at indsamle PRO-data med en acceptabel kvalitet og responsrate på denne måde. Endvidere stiller løbende PRO-indsamling både logistiske, teknologiske og datasikkerhedsmæssige krav.

CPR-registret er løbende opdateret med oplysning om evt. dødsdato. Forud for hver kontakt til patienten bør der ske opdatering af information om adresser og vitalstatus, med henblik på at undgå at sende skemaer til afdøde. Et fuldt integreret PRO-system kan håndtere protokoller med kontinuerlig mixed-mode dataindsamling på papir og internet, og kan kommunikere med patienter via e-mail, digital post, sms og papirbreve samt håndtere dynamisk skift mellem modes for både kontaktmetode og besvarelsesmetode. Anvendelse af papirskemaer forudsætter integrering af OCR-scanningssystem, som nævnt ovenfor, og der er relativt store etableringsomkostninger forbundet hermed. Uanset om PRO indsamles via papir, internet eller mixed-mode, skal systemet varetage en løbende automatisk proces, styret af den protokolfastlagte tidslinje, i forhold til f.eks. en diagnose- eller operationsdato samt rykkerprocedurer. Desuden skal det sikres, at der gennemføres en eller flere rykkerprocedurer ift. den patientspecifikke tidslinje.

16. Analyse og fortolkning af PRO-data

I dette afsnit beskrives de statistiske analyser, som kan være passende at bruge ved sammenligning af grupper af patienter (dvs. de indberettende enheder), når PRO-data er indhentet på samme tidspunkt for de pågældende patienter. Samtidig kommer vi ind på, hvordan resultaterne kan fortolkes i klinikken.

PRO-instrumenter består som regel af en række spørgsmål (variable), hvor der til hvert svar tilhører et tal eller en værdi. Svarene kan anvendes alene eller kombineres i flere subdomæner eller i én total score, afhængig af beregningsreglerne for det pågældende PRO-instrument.

Alt afhængig af hvordan svarene er konstrueret, inddeles måleskalaer generelt i 3 typer:

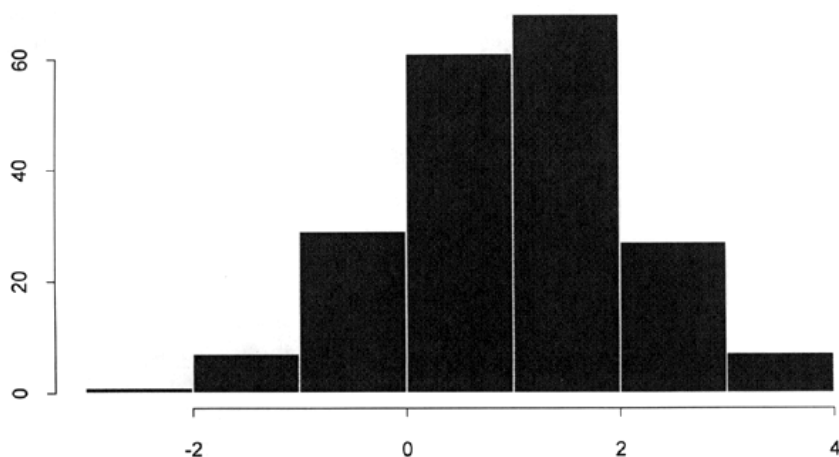
- Nominal-skalaer, hvor svarene ikke har en naturlig rækkefølge, f.eks. ja og nej.
- Ordinal-skalaer, hvor svarene kan rangordnes naturligt f.eks. ingen, lette, moderate og svære smerter.
- Kontinuerlige skalaer, hvor svarene kan antage alle værdier i et interval (f.eks. højde, vægt).

Valget af statistiske metoder afhænger af typen af svar i måleskalaen.

Data skal også håndteres forskelligt afhængig af, om de er normalt eller ikke normalt fordelt, hvilket bør undersøges før yderligere analyser foretages. Ved normalfordeling af data forstås en situation, hvor flest observationer ligger inde mod midten, og så fordeler de sig ellers symmetrisk ud til begge sider.

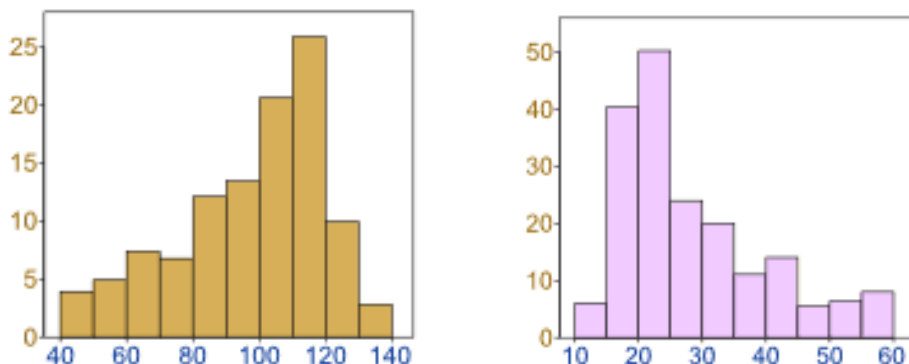
Normalfordeling kan bestemmes f.eks. ved at tegne et histogram (se eksempel i figur 3).

Figur 3. Normalfordeling af data



Ud over normalfordelingen findes en lang række fordelinger, som er relevante i forskellige sammenhænge, hvor data er f.eks. typisk ”højreskæv” eller ”venstreskæv” fordelt. En fordeling med en middelværdi, der ligger tydeligt over medianen, kaldes højreskæv, mens en fordeling med en middelværdi, der ligger tydeligt under medianen, kaldes venstreskæv. Se eksempler med ikke normalfordeling af data i figur 4.

Figur 4. Ikke normalfordeling af data



16.1 I tilfælde hvor kun en måling af PRO-data er tilgængelig

Hvis data er normalfordelt, kan der beregnes en gennemsnitsværdi og en standard deviation af de indsamlede PRO-data. To enheder kan sammenlignes ved at beregne forskellen mellem gennemsnitsværdier i de to grupper med standard error som mål for statistisk usikkerhed. Desuden kan man undersøge, hvorvidt gennemsnitsværdierne i to grupper kan antages at være identiske ved at lave en statistisk hypotesetest, t-test. Hvis flere grupper af patienter skal sammenlignes, kan ANOVA-metoden (på engelsk: Analysis of variance) anvendes (61).

Ofte er PRO-data ikke normalfordelt, og bør beskrives med medianværdier i stedet for gennemsnitværdier. Desuden skal man bruge ikke-parametriske metoder, som f.eks. Wilcoxon rank-sum test (Mann-Whitney test) for at sammenligne to uparrede grupper af patienter, eller Kruskal-Wallis test for at sammenligne tre eller flere grupper af patienter (61).

16.2. I tilfælde hvor to eller flere målinger af PRO-data er tilgængelige

Flere målinger på den samme patient over tid er vigtig, idet det giver mulighed for at analysere udviklingen/ændringer over tid.

Ændringer over tid kan beregnes for hver enkelt patient ved hjælp af følgende ligning:

$$\text{Ændring over tid} = \text{PRO-score ved follow up} - \text{PRO-score ved baseline}$$

F.eks. ændring i PRO-score efter hofteoperation = Postoperativ PRO-score – Præoperativ PRO-score.

Ændringer i PRO-score beregnet på samtlige patienter behandlet på en afdeling, giver mulighed for at beregne en gennemsnitsværdi af PRO-ændringer på afdelingsniveau.

Bagefter kan man undersøge, om der er signifikant forskel mellem gruppernes (afdelingernes) gennemsnitsændringer i PRO-score over tid. Til det formål vil t-test for to grupper eller variansanalyse for flere grupper kunne anvendes (ANOVA). Selvom denne metode er nem at udføre, er den ikke så nemt at tolke, idet justering for case-mix ikke er mulig.

16.3. Valg af referencegruppe

For at sammenligne afdelinger skal man vælge en referencegruppe.

- Man kan vælge en bestemt afdeling som reference.
- En anden og mere relevant mulighed er at sammenligne afdelinger med normdata eller internationale standarddata, hvis sådanne data findes for det valgte PRO-instrument. F.eks. findes der danske normdata for PRO-instrumentet SF-36 både for mænd og kvinder, og for forskellige aldersgrupper, se dansk manual til SF-36 (62). En kvalitetsindikator kunne f.eks. være: ”Andelen af patienter, hvis PRO-værdi ligger under eller over normværdien.”
- Den tredje mulighed er at oprette en dikotomisk kvalitetsindikator baseret på indsamlede PRO-data. F.eks. i DSR har man besluttet, at resultatet af en skulderalloplastik-operation blandt osteoatrose patienter er tilfredsstillende, hvis WOOS-score for den enkelte patient ligger på >50 point. Kvalitetsindikatoren hedder: ”Andel af patienter med WOOS-score >50 point for patienter med skulderalloplastik indsat under indikationen osteoatrose.” Hvis man ikke kan vælge et bestemt cut-off niveau, så kan man vælge at bruge kvartiler eller percentiler. Kvartiler og percentiler er de observationer, som – når datasættet er sorteret – deler data i henholdsvis fire eller ti lige store mængder. Eksempelvis er der tale om et utilfredsstillende resultat af en skulderalloplastik, hvis WOOS ligger i første kvartil eller nedre kvartil. Kvalitetsindikatoren vil således hedde: ”Andel af patienter med WOOS-score, som ligger i den nedre kvartil.” En kvalitetsindikator, der indbefatter en ”andel af patienter,” opgøres som proportion eller procentdel med 95% konfidensinterval. Chi²-test kan bruges til at analysere forskel mellem proportioner (61).
- Man kan også vælge at beregne andelen af patienter, som har oplevet ”minimal clinically important improvement”, som er beskrevet i afsnittet ”Minimal clinically important improvement”.

16.4. Justering for case-mix

I de fleste tilfælde vil PRO-målinger, ligesom mange andre resultatmål, afhænge af patientkarakteristika som alder, køn, komorbiditet, og socioøkonomiske faktorer. Derfor er det vigtigt, når man sammenligner afdelinger, at man tager højde for disse faktorer.

Lineære- eller logistiske regressionsmetoder anvendes for at undersøge en sammenhæng mellem afdeling og a priori-specificeret PRO-score ved at tage højde for forskellige baselinefaktorer. I regressionsanalyser kan der også tage højde for baseline PRO-score. Se eksempel i afsnit 17 ”Præsentation af PRO-data”.

16.5. Mangelfulde (missing) PRO-data

Ved indsamling af PRO-data, ligesom ved indsamling af andre data, er det svært at undgå mangelfulde eller manglende data, især hvis data indsamles i flere opfølgingsperioder. Missing PRO kan opstå af flere forskellige årsager. Nogle patienter svarer ikke, fordi de dør eller fordi de er så raske, at de ikke finder det relevant at svare. Andre er for syge til at svare, og der kan være logistiske eller praktiske årsager, som gør at afdelinger glemmer at sende PRO-skema til patienten eller at skemaet forsvinder i posten.

Missing data udgør et problem både i de statistiske analyser og ved fortolkning af resultaterne. Hvis datasættet er for lille, vil det mangle statistisk styrke til at identificere forskel mellem grupperne, og missing data øger risikoen for, at den forskel man har identificeret er biased.

Der eksisterer ikke faste grænser for, hvor mange missing data, der er acceptable. Det vil være en vurderingssag i hvert enkelt tilfælde. Den vigtigste overvejelse er, om missing data er relateret til outcome (resultatet af PRO-målingen), hvilket øger sandsynligheden for uafhængig af antallet af missing data.

Hvis andelen er meget lille (<5%) vil det formentlig ikke have betydning, mens missing på mellem 5% og 20% kan have betydning for potentielle bias og styrke i studiet. Andele over 20% må anses som potentielt havende betydning for resultaterne og medføre begrænsninger for fortolkning af resultaterne (9).

En metode til at afklare disse forhold er at foretage en pilottest, inden den egentlige dataindsamling igangsættes.

Missing data kan forekomme på forskellige måder.

- 1) **Missing completely at random (MCAR)**, dvs. missing PRO-data er total uafhængige af patientkarakteristika eller behandling. MCAR forekommer f.eks. når en afdeling har glemt at sende PRO-skema til patienten, eller hvis patienten er flyttet til et andet land, hvilket ikke er relateret til patientens heldbred.
- 2) **Missing at random (MAR)**, dvs. missing PRO-data er afhængige af patienternes tidligere svar på PRO-skema. MAR forekommer f.eks. hvis en patient, som har dårlig livskvaliteten ved baseline, er mindre tilbøjelige til at besvare PRO-skema ved den næste opfølgning.
- 3) **Missing not at random (MNAR)**, dvs. missing PRO-data er afhængig af en bestemt ukendt /ikke observeret årsag. MNAR forekommer f.eks. hvis en patient som har seksuelle problemer, bevidst springer over spørgsmål omkring ”seksuel interesse”, og dermed at missing svar kan betyde ”rigtig mange seksuelle problemer”.

Vi kan ikke sige ud fra de indsamlede PRO-data, om de manglende observationer er MCAR, MAR eller MNAR. Der findes dog enkelte analyser, som kan hjælpe med at afsløre om data er MCAR eller MAR (45).

Påvisning af mekanismer for forekomst af manglende data, vil kunne bruges til at forudsige eventuel retning og størrelse af bias, når vi udtaler os om forskelle mellem afdelinger. Det er dog ikke så nemt at gennemføre i praksis. Statistisk testning, korrelationsanalyse eller logistik regression kan måske afsløre om de patienter, som har besvaret PRO-skemaet (responders) adskiller sig væsentligt og systematisk fra de patienter, som ikke har besvaret PRO-skemaet (non-responders). Disse analyser kan sige noget om, hvorvidt de indsamlede PRO-data er repræsentative for den populationen, der undersøges.

Der er flere typer af mangelfulde PRO-data, som skal håndteres i analyserne, nemlig hvis der mangler et helt PRO-skema, eller et subdomæne eller hvis PRO-skemaet er indsamlet, men er mangelfuldt udfyldt (dvs. der mangler svar på et eller flere spørgsmål). I alle tilfælde kan fejltolkning af resultaterne forekomme.

Håndtering af missing PRO-data

Indsamlingsperioden

I indsamlingsperioden er det vigtigt, at et dedikeret personale deltager og koordinerer indsamlingen af data. Kontinuerlig undervisning af patienter og personale om betydningen af missing data for dataanalyserne, forekomst af troværdige resultater og fortolkning af resultaterne kan nedsætte andelen af missing data. Desuden kan regelmæssig fokus på kompletthed af PRO-data nedsætte andelen af missing data. Andre aspekter spiller yderligere en rolle i forhold til besvarelse af PRO-data, som eksempelvis indsamlingsmåden, indsamlingstidspunktet, relevans af spørgsmål og udsendelse af reminder.

I indsamlingsperioden kan man dokumentere årsager til missing data for senere at bruge det i analysefasen og ved fortolkning af resultater. F.eks. kan der tilføjes et ekstra spørgsmål i starten af PRO-skemaet, som ”Patienten ønsker ikke at besvare PRO-skemaet pga. manglende tid, dårligt helbred, godt helbred, problemer med at forstå det danske sprog, eller noget andet end helbred.”

Analysefasen

Ved manglende svar på et eller flere spørgsmål i PRO-skemaet, er det nemmeste at betragte disse PRO-skemaer som totalt missing og ekskludere dem fra analyserne. Dette vil dog nedsætte den statistiske styrke. I andre tilfælde bliver missing data ikke taget med i beregningen af total score. Det vil sige, at gennemsnittet af alle svarværdierne divideres med antal besvarede spørgsmål i skemaet, og ikke med det samlede antal mulige svar i spørgeskemaet, se f.eks. SF-36 beregningsmanual (62). Til sidst, kan man antage at missing værdier er de samme som gennemsnitværdien af de besvarede spørgsmål (såkaldt gennemsnitsværdi-imputation). ”Multiple imputation” kan også bruges til at forudsige værdien af missing data, baseret på værdierne af de besvarede spørgsmål (63).

Manglende besvarelser af hele PRO-skemaer er en lidt mere alvorlig situation, og kan håndteres på flere måder:

- 1) I en komplet case-analyse ekskluderes patienter med missing PRO-skemaer fra analysen, og analysen baseres kun på de komplet udfyldte PRO-skemaer. Sample size og statistisk styrke kan blive alvorligt påvirket ved at bruge denne metode.
- 2) Værdien fra det sidst besvarede PRO-skema overføres til det næste PRO-skema, som er missing.
- 3) Data fra baseline PRO-skema bliver overført til et missing opfølgings-PRO-skema.
- 4) Sample gennemsnits-imputation er en metode, hvor der først beregnes en gennemsnitsværdi af PRO-data på en afdeling, baseret på de patienter, som har besvaret PRO-skemaet. Bagefter imputeres denne gennemsnitsværdi for de patienter, der har et missing PRO-skema.
- 5) "Multiple regression model imputation" er metode, som beregner PRO-totalscore for patienter med missing skema baseret på alle de patientrelaterede data, der er tilgængelige (63).

Generelt afhænger pålideligheden af de forskellige metoder til håndtering af missing data af årsagerne til at missing data er opstået. Det er dog ikke altid nemt at afgøre, hvorfor data er missing. Derfor er det god videnskabelig praksis at undersøge robustheden af disse metoder ved at lave sensitivitetsanalyser. Formålet med sensitivitetsanalyser er at sammenligne resultaterne på tværs af de forskellige metoder, og for at se hvordan og hvor meget de forskellige metoder påvirker resultaterne. Et eksempel på sensitivitetsanalyse er at erstatte missing PRO-værdier med ekstrem høj eller ekstrem lav værdi og gentage analyserne (såkaldt "worse and best case scenario"). Hvis resultaterne er konsekvent de samme, er konklusionerne valide. Hvis resultaterne ikke er de samme eller modsatte, er vi nødt til yderligere at undersøge mulige årsager til denne uoverensstemmelse.

16.6. "Minimal clinically important improvement" (MCII)

Statistisk signifikant forskel er ikke det samme som klinisk signifikant forskel. MCII er den mindste forskel (forbedring) i PRO-score, som patienten selv har vurderet til at være gavnlig, og som kan få sundhedspersonalet til at ændre behandling af patienten (64;65).

For at bestemme MCII kan der anvendes en ekstern måling (ankerbaseret måling) relateret til det bestemte PRO-instrument (66). En sådan ekstern måling kan komme fra patientens journal, laboratorietest, anden klinisk undersøgelse, eller helst fra patienten selv. F.eks. kan patienten, ud over PRO-skemaet, få tilsendt spørgsmål omkring forbedring af heldbred. For at finde MCII blandt hoftealloplastikpatienter i Danmark, lavede Paulsen et al. (67) et studie, hvor patienterne blev spurgt både før og efter operation: "Hvordan synes du, dit heldbred er nu?". Patienterne kunne svare følgende: "Fremragende", "Vældig godt", "Godt", "Nogenlunde" og "Dårligt". Patienten havde oplevet minimal forbedring, hvis han/hun svarede minimum 1 kategori bedre efter operation i forhold til før operation (66). Man kunne bagefter beregne MCII for

gennemsnitsændring i PRO-scoren over tid, baseret på minimal forbedring besvaret af patienten. Paulsen et al (67) beregnede også at en forbedring på 24 point i smerte, målt på subdomæne i HOOS-skemaet, svarede til en minimal klinisk relevant forbedring vurderet af patienten.

MCII kan være forskellig fra PRO-skema til PRO-skema, men kan også være forskellig for forskellige subdomæner i det samme PRO-skema. Desuden er MCII cut-off point meget afhængig af den kontekst PRO-skemaet bliver anvendt i, dvs. afhængig af patientgruppen, patientens alder, køn, diagnose for behandling/operation, type af anker (ekstern) måling, og tidspunkt for besvarelse af PRO-skemaet.

MCII kan bruges i de kliniske kvalitetsdatabaser for at fremme udarbejdelsen af relevante PRO-baserede resultatmål ved at udtrykke dem som andel af forbedrede patienter.

Konklusion:

- Databasen skal på forhånd overveje hvilke statistiske metoder, der skal anvendes, afhængigt af om der er 1, 2 eller flere PRO-målinger tilgængelige.
- Der skal tages stilling til valg af referencegruppe før dataindsamlingen påbegyndes.
- Missing PRO-data er vanskelige at undgå. Procedurer for at begrænse missing data samt procedurer for håndtering af missing data i analysefasen skal besluttes før indsamlingsfasen begynder.
- Det skal besluttes, hvordan der tages højde for forskelle i patientsammensætning fra afdeling til afdeling.

17. Præsentation af PRO-data

Afrapportering i tabelform er den mest basale afrapporteringsmetode, og kan indeholde detaljerede oplysninger om antal besvarede skemaer, antal uoplyste (ikke besvarede skemaer), middel- og medianværdier, proportioner (andelen der har scoret over / under en fastsat norm), differenser samt relative estimater, der er justeret for f.eks. køn, alder og komorbiditet. PRO-data kan også illustreres ved forskellige typer af grafer, som typisk giver et bedre visuelt udtryk på bekostning af færre detaljer. Principielt er kvalitetskravene til tabeller og figurer de samme, uanset om man afrapporterer PRO-data eller andre data.

Tabel 10 er et eksempel på afrapportering af livskvalitetsdata målt med SF36, hos patienter, der har fået foretaget en gastric banding operation pga. fedme. Livskvalitet er målt både før og 3 år efter operation, og der afrapporteres middelværdier og standarddeviationer for begge måletidspunkter samt middeldifferensen for de to målinger. Tabellen inkluderer data for to overordnede subskalaer PCS og MCS (mental og fysisk) samt for 8 underordnede subskalaer, og giver dermed et billede af den oplevede livskvalitet. Desuden viser tabellen den forbedring, der er sket i livskvaliteten svarende til en positiv middeldifferens varierende for de

enkelte subskalaer fra 9 til 44 point. Tabellen inkluderer kun data fra én enhed, og den kan blive noget uoverskuelig, hvis den skal dække mange afdelinger, hvilket ofte er tilfældet i kvalitetsovervågning.

Tabel 10. Forandringer i helbredsrelateret livskvalitet efter fedmeoperation målt prospektivt med SF36
Robert M et al. *Obes Surgery*, 2013 (68)

SF-36 assessment	Baseline	Follow up ^a	Mean difference (95%CI)	P value ^b
PCS score, mean ± SD	39.2±8.5	53.7±5.4	14.4 [12.9; 15.9]	<0.001
MCS score, mean ± SD	39.8±10.2	49.0±8.4	9.2 [7.2; 11.1]	<0.001
Physical functioning score, mean ± SD	50.7±24.2	88.9±15.9	38.9 [34.8; 42.9]	<0.001
Role physical score, mean ± SD	48.7±38.7	93.5±17.1	44.2 [37.8; 50.6]	<0.001
Bodily pain score, mean ± SD	48.5±23.8	76.5±16.4	33.9 [29.4; 38.5]	<0.001
General health score, mean ± SD	50.7±19.0	82.9±20.0	25.8 [22.3; 29.3]	<0.001
Vitality score, mean ± SD	39.2±18.7	65.1±18.8	25.9 [21.9; 29.8]	<0.001
Social functioning score, mean ± SD	54.2±24.3	83.3±20.2	28.8 [24.0; 33.5]	<0.001
Role emotional score, mean ± SD	53.0±40.4	88.7±25.0	35.0 [27.9; 42.1]	<0.001
Mental health score, mean ± SD	50.6±19.4	69.7±16.8	19.1 [15.7; 22.6]	<0.001

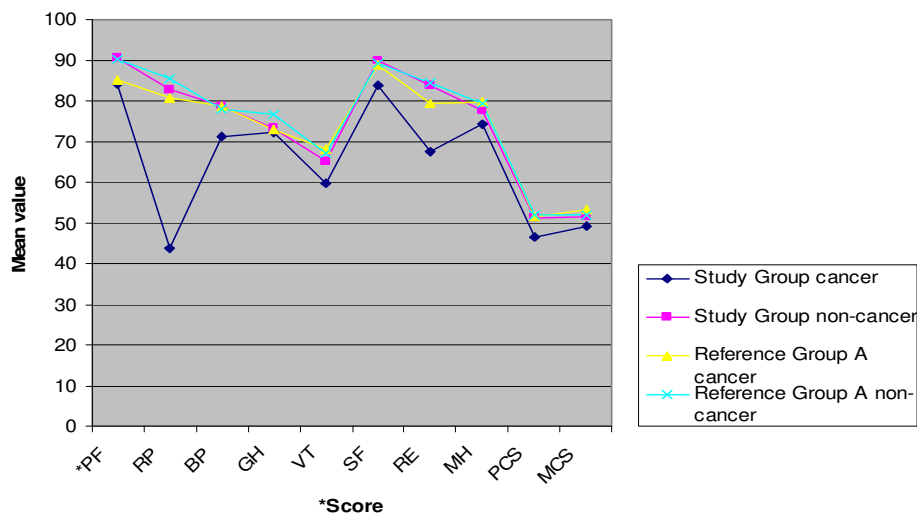
PCS physical component summary scale, MCS mental component scale

^aNumber of patients with last follow-up data recorded at: 3 years=83, 21/2 years=31, 2 years=19, 1,5 years=12 and 12 months=19

^bMc Nemars's test or Wilcoxon's signed rank test

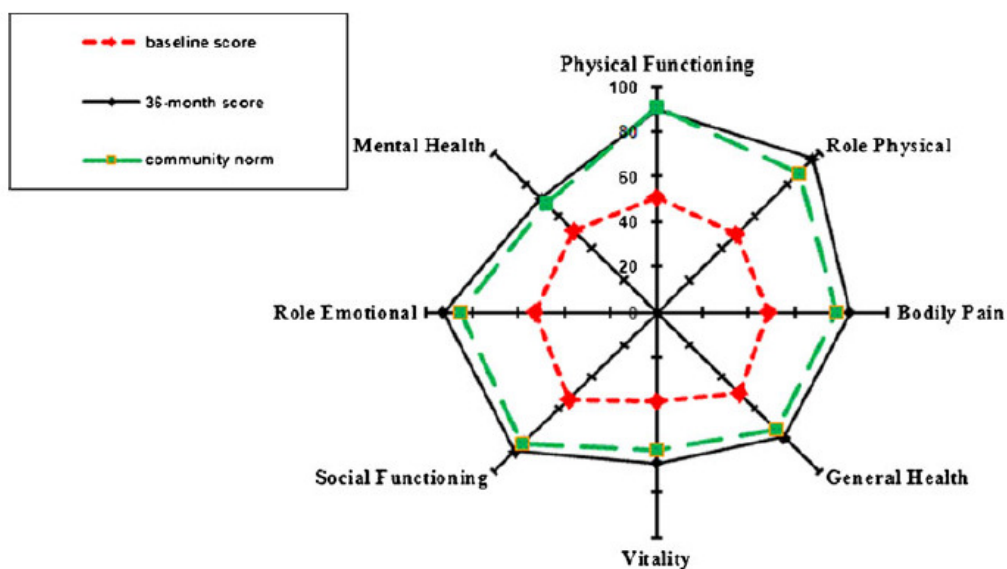
Figur 5 er et eksempel på, hvordan middelværdier for SF36-skalaerne kan afbilledes for flere grupper. Her er afbilledet værdier for forskellige patientgrupper, men det kunne lige så vel være for forskellige regioner eller afdelinger, som er den relevante enhed for kvalitetsovervågning. Figuren illustrer kun én måling svarende til baseline, og der er ikke taget hensyn til f.eks. alder eller øvrige sygdomme.

Figur 5. Middelverdier for SF36-subskaler målt ved baseline blandt deltagere i et kohortestudie af genetisk rådgivning for arvelig bryst- og ovariecancer. Ellen M. Mikkelsen, Upubliceret data



I figur 6 illustreres livskvalitetsdata for patienter, der har fået foretaget en gastric banding operation jf. tabel 10. Figuren giver et samlet billede af alle subskalaer og illustrerer ud over resultaterne fra baseline og 36 måneders opfølgning populationsbaserede normdata. Herved kan ændringer fra baseline til opfølgningstidspunktet evalueres, samtidig med at resultaterne kan sammenlignes med tilsvarende resultater for en normpopulation, der ikke har fået foretaget den pågældende operation.

Figur 6. Forandringer i helbredsrelateret livskvalitet efter fedmeoperation målt prospektivt med SF36 Robert M et al. *Obes Surgery*, 2013 (68)



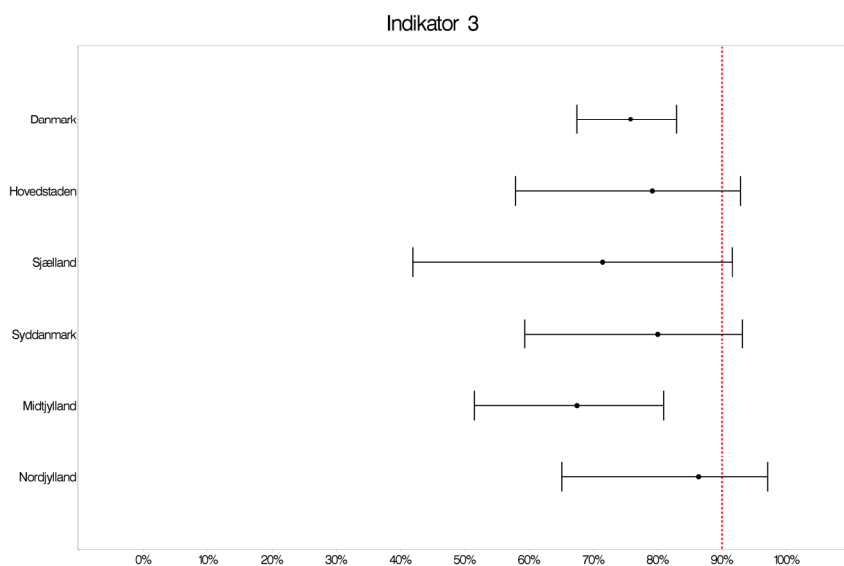
Tabel 11 er et eksempel fra Årsrapport 2014 fra Dansk Skulder Register, som har en indikator, der evaluerer skulderfunktion efter indoperation af skulderprotese. Patienterne besvarer postoperativt et PRO-instrument (WOOS), hvor en score > 40 point for 90 % af patienterne indikerer god kvalitet. Tabellen viser, at 128 patienter på landsplan har besvaret skemaet, og 66 patienter har ikke besvaret. På landsplan har 76 % scoret > 40 point, og det resultatet varierer mellem 67 % og 86 % for de 5 inkluderede regioner. Tabellen giver således et billede af patienternes protesefunktion efter operation i de fem regioner. Men kun ca. 2/3 af de opererede har besvaret skemaet, og det er ukendt om de er repræsentative for hele populationen, hvilket må tages i betragtning ved fortolkning af data. Derudover lægger tabellen op til en sammenligning mellem regionerne, men det vides ikke, om de inkluderede patienter er sammenlignelige i forhold til f.eks. alder eller andre faktorer, der kan have indflydelse på protesens funktion. Samtidig er der kun foretaget en måling 1 år efter operation, hvorfor det ikke er muligt at vurdere ændringer over tid eller tage højde for præoperative skulderfunktion.

Tabel 11. WOOS score >40 hos patienter med skulderalloplastik indsat under indikationen frisk fraktur (< 14 dage). Dansk Skulder Register, Årsrapport 2014 (s. 14)(69).

	Std. 90% opfyldt	Tæller/nævner	Aktuelle 2012			Tidligere år	
			Uoplyst Antal (%)	%	95% CI	2011 % (95% CI)	2010 % (95% CI)
Danmark	nej	97 / 128	66(34)	76	(67 - 83)	79 (72-85)	75 (69-81)
Hovedstaden	ja*	19 / 24	11(31)	79	(58 - 93)	87 (74-95)	77 (61-88)
Sjælland	ja*	10 / 14	7(33)	71	(42 - 92)	83 (59-96)	75 (58-88)
Syddanmark	ja*	20 / 25	8(24)	80	(59 - 93)	62 (44-78)	76 (59-89)
Midtjylland	nej	29 / 43	26(38)	67	(51 - 81)	77 (60-90)	72 (59-83)
Nordjylland	ja*	19 / 22	14(39)	86	(65 - 97)	83 (67-94)	78 (61-90)

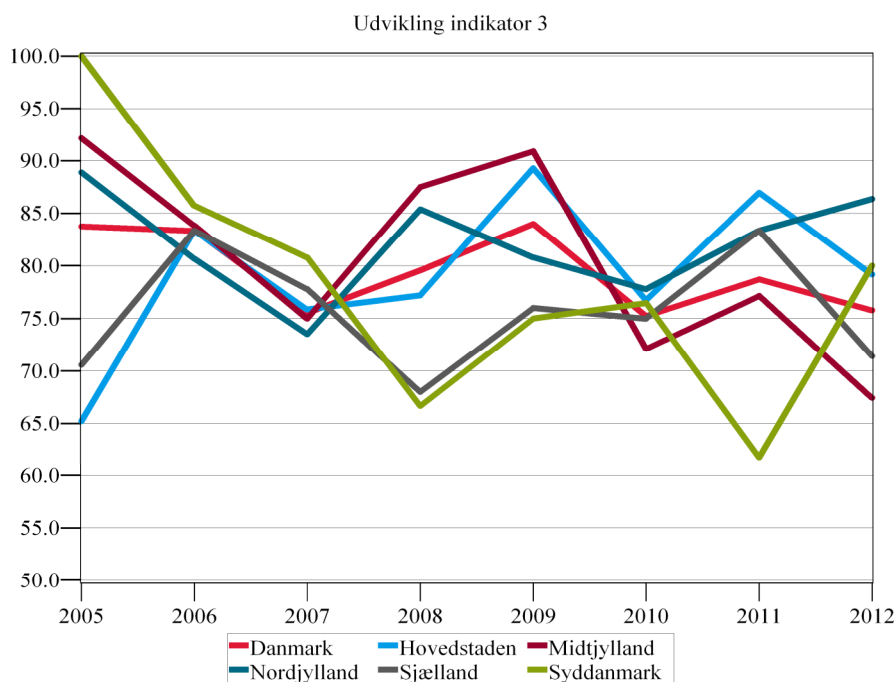
Figur 7 illustrerer de samme data som tabel 11. Figuren giver et tydeligt billede af den manglende standardopfyldelse på landsplan og i de 5 regioner, men figuren rapporterer færre detaljer end tabellen. F.eks. giver figuren ingen oplysninger om antal ubesvarede skemaer.

Figur 7. Andel af patienter med WOOS score >40 hos patienter med skulderalloplastik indsat under indikationen frisk fraktur (< 14 dage). Dansk Skulder Register, Årsrapport 2014 (s. 14)(69).



Figur 8 illustrerer ligeledes data svarende til tabel 11 og figur 7, men er udvidet med data for en længere årrække (2005-2012). Dette giver mulighed for at vurdere udvikling over tid, men det må tages i betragtning om antallet af uoplyste og patientpopulationen har ændret sig over tid f. eks. er blevet ældre på operationstidspunktet. Ellers kan man fortolke data således, at kvaliteten har været faldende i nogle regioner, uden at det nødvendigvis er korrekt.

Figur 8. Andel af patienter med WOOS-score >40 for patienter med skulderalloplastik indsat under indikationen frisk fraktur (< 14 dage), 2005-2012. Dansk Skulder Register, Årsrapport 2014 (s. 15)(69).



Tabel 12 er et eksempel på en tabel, der viser PRO-data fra det svenske hofteregisters årsrapport 2012. Tabellen viser gennemsnitsværdier for smerte (Pain), livskvalitet (EQ VAS og EQ-5D) samt tilfredshed (Satisf) målt både præ- og postoperativt med undtagelse af tilfredshed, der kun er målt postoperativt. Tabellen indeholder desuden data (Charnley Class C) vedrørende andelen af patienter, der har multiple ledproblemer eller komorbiditet. I rapporten afreporteres for en lang række hospitaler, men her er kun medtaget nogle få hospitaler. På trods af at tabellen inkluderer data for gentagne målinger, er det svært at sammenligne data fra de forskellige hospitaler, da patientgrundlaget kan variere betydeligt, hvilket Charnley Class C værdierne også indikerer. Hvis en afdeling hovedsageligt behandler patienter, der generelt er sunde, vil de have bedre livskvalitet og færre smerte både før og efter operationen, og den gennemsnitlige forbedring fra præ- til postoperativt bliver lav sammenholdt med landsgennemsnittet. Dette kan fejlagtigt tolkes som dårlig kvalitet. Modsat kan den afdeling, der behandler patienter, der som udgangspunkt har et dårligt helbred med flere komplicerende sygdomme – fremstå med bedre kvalitet, da disse patienter har et stort potentiale for forbedring, som gennemsnitligt kan være større end landsgennemsnittet.

Tabel 12. Patient rapporterede oplysninger pr. hospital. Svensk Hofte Alloplastik Register. Årsrapport 2012, s. 87(36)

Patient-reported outcome per hospital 2010–2011

	Preop, 2010–2011					One-year postop, 2011–2012					Six-year postop, 2011–2012				
	Number	C-cat ¹⁾	Pain	EQ VAS	EQ-5D	Number	Satisf. ²⁾	Pain	EQ VAS	EQ-5D	Number	Satisf. ²⁾	Pain	EQ VAS	EQ-5D
University or regional hospitals															
Karolinska/Huddinge	402	57%	79	59	0.42	391	19	17	73	0.73					
Karolinska/Solna	275	51%	64	48	0.33	299	19	16	72	0.73					
Linköping	46	35%	69	46	0.33	71	16	12	74	0.75					
SU/Mölndal	463	49%	65	56	0.34	545	22	18	71	0.71	66	21	20	66	0.65

1) Proportion Charnley class C.

2) Satisfaction (VAS, 0 = Completely satisfied, 100 = Dissatisfied).

The table presents result in the form of number of patients, mean values of pain VAS and EQ-5D index pre-operatively, together with the proportion of Charnley class C patients (i.e. patients with multiple joint disease and/or co-morbidity). Departments with a high proportion of C patients most frequently show lower average values for all parameters both pre-operatively and after one year. However, the prospectively gained values are most often not equal affected by C affiliation. Results are presented for units with 40 or more registrations.

Tabel 13 er baseret på data fra det svenske hofteregisters årsrapport 2012 tilsvarende tabel 12. I denne tabel afrapporteres der tilfredshed og livskvalitet målt 1 år efter operation. For at imødekomme problemet med vanskeligt sammenlignelige patientpopulationer (case mix), er der for hver afdeling, på baggrund af en regressionsmodel, estimeret den forventede værdi for tilfredshed og livskvalitet ved 1 år justeret for alder, køn, komorbiditet (Charnley C), diagnose samt baselineværdien for henholdsvis livskvalitet og tilfredshed. Den forventede værdi og den aktuelle værdi er sammenholdt, og afvigelsen fra det forventede er beregnet. For tilfredshed indikerer en negativ værdi, at resultatet er bedre end forventet, og for livskvalitet indikerer en værdi større end 0, at resultatet er bedre end forventet. Denne afrapportering giver således mulighed for en mere reel sammenligning mellem afdelingerne, da der er justeret for en række vigtige faktorer. Men der er naturligvis også andre faktorer, som kunne have betydning, og som der ikke er taget højde for ved justeringen f.eks. uddannelsesniveau, kulturelle faktorer samt komorbiditet, der ikke er indeholdt i Charnley C-variablen. Derudover må man være opmærksom på, om der er regionale forskelle i besvarelsen af de respektive PRO-instrumenter.

Tabel 13. Patient rapporterede oplysninger pr. hospital. Svensk Hofte Alloplastik Register. Årsrapport 2012, s. 87(36)

Expected outcome and improvement index per hospital 2010–2011

Hospital	Satisfaction VAS 1 year			Number (EQ-5D, EQ VAS, Pain)	EQ-5D index 1 year		
	Number (Tilff)	Actual	Deviation from expected		Actual	Deviation from expected	Improvement index
University or regional hospitals							
Karolinska/Huddinge	391	19.2	3.9	356	0.733	-0.055	54.60
Karolinska/Solna	299	19.4	3.0	249	0.732	-0.031	59.65
Linköping	71	15.7	2.2	35	0.750	-0.052	56.41
SU/Mölndal	545	22.3	5.8	432	0.711	-0.064	58.39
SUS/Lund	75	15.9	-1.1	55	0.673	-0.061	58.45
SUS/Malmö	86	13.3	-4.3	79	0.780	0.053	71.82

Number (Satisf.) = Number of registrations per hospital with satisfaction VAS one year postoperatively.

Actual = Mean outcome one year postoperatively

Deviation from expected = Difference between the actual mean value and the expected value calculated using regression coefficients in a modell that includes age, gender, Charlley class, and preoperative level of the respective PROM variable.

For EQ-5D index and EQ VAS values above zero indicate better outcome than expected and for satisfaction and pain negative values indicate better outcomes than expected.

Number (EQ-5D index, EQ VAS, Pain) = Number of registrations per hospital with EQ-5D index, EQ VAS and Pain-VAS both pre- and one year postoperative.

Improvement index = The ratio between the mean improvement divided by the maximum possible improvement (see text for details).

Konklusion:

- PRO-data skal præsenteres både i tabelform og grafisk for at fremme fortolkning og forståelse af data.
- Både tabeller og grafer skal være selvforklarende.

18. Dokumentation

Ved udvælgelse og fastlæggelse af traditionelle kvalitetsindikatorer er det et krav, at indikatorerne måler de vigtigste og alvorligste aspekter af behandlingskvaliteten, og at indikatorerne er evidensbaserede (1). Det er ligeledes en forudsætning, at standarden eller det ønskelige niveau af kvalitet fastlægges på baggrund af videnskabelige studier. Når det ikke er muligt, fordi der ikke findes relevante studier, må indikatorer og standarder fastlægges på baggrund af konsensus og bedst mulig praksis. Principielt gælder dette uanset om indikatorer og standarder anvender selvrapporterede data eller andre data. Men for selvrapporterede data må der også foreligge evidens for, at PRO-instrumentet er validt i den givne patientpopulation, og at det valgte PRO-mål giver mening i den aktuelle sammenhæng. Kravene til dokumentation af 1) patientpopulationen (in- og eksklusion af patienter), 2) dataindsamlingsprocedurer, 3) missing values, 3) variabelbeskrivelser samt 4) indikatorspecifikationer og beregningsregler er ligeledes enslydende for PRO-baserede indikatorer og traditionelle indikatorer.

19. Konklusioner og anbefalinger

- Definition af PRO-data:
PRO-data kommer direkte fra patienten uden fortolkning fra f.eks. sundhedsfaglige personale. Det er væsentligt at skelne mellem PRO-data, -instrumenter og – kvalitetsindikatorer, selvom der er en tæt sammenhæng mellem begreberne. PRO-kvalitetsindikatorer bygger på PRO-data, som indsamles gennem PRO-instrumenter.
- Formål med PRO-data:
PRO-data kan anvendes på flere niveauer og i forskellige sammenhænge. I de kliniske kvalitetsdatabaser anvendes data på gruppeniveau, men det kan overvejes om data også kan anvendes på individniveau.
- Der er svag evidens for at anvendelsen af PRO-data på individniveau i klinisk praksis kan medføre et bedre patientforløb inkl. sygdomsudfald. Der er endnu ingen evidens for at brug af PRO-data på gruppeniveau, herunder brug af PRO-baserede kvalitetsindikatorer, medfører en bedre behandlingskvalitet og bedre sygdomsudfald. Der er dog tale om et umodent forskningsfelt med ganske få studier, som er behæftet med væsentlige metodologiske begrænsninger.
- Datatilladelser og datasikkerhed:
Indsamling og anvendelse af PRO-data til de godkendte kliniske kvalitetsdatabaser skal foregå efter de samme principper, som gælder for alle andre data i de kliniske databaser, if. de to bekendtgørelser (Bekendtgørelse nr. 459 af 16/05/2006 og Bekendtgørelse nr. 1725 af 21.12.2006) for kliniske kvalitetsdatabaser samt persondataloven.
- Erfaringer fra Danmark:
PRO-data indsamles og anvendes kun i relativt få nationale kliniske databaser i Danmark. Databaserne har typisk anvendt allerede eksisterende PRO-instrumenter, som regel efter en officiel validering og oversættelse til dansk. De fleste databaser forsøger at indsamle PRO-data både før og efter operation/ behandling med henblik på at kunne måle ændringer efter intervention og tage højde for forskelle ved baseline. Data anvendes primært som resultatindikatorer. Lave besvarelsesprocenter har i nogle databaser givet udfordringer i forhold til fortolkning af resultater og sammenligninger på tværs af afdelinger. Ingen af databaserne har dog i analysefasen taget stilling til håndtering af lav besvarelsesprocent, kun en måling af PRO-data, eller anvendt PRO-data til justering for case-mix. Elektronisk indsamling af PRO-data er mere undtagelsen end reglen, og de indsamlede PRO-data er derfor ikke en del af den elektroniske patientjournal.
- Erfaringer fra udlandet:
Nogle lande har adskillige års erfaringer med anvendelse af PRO-data i kvalitetssammenhæng. Generelt er det først inden for de seneste 10-15 år, at udviklingen har taget fart. De beskrevne eksempler på udenlandske databaser anvender oftest validerede instrumenter, typisk en kombination

af et generisk og et sygdomsspecifikt. I nogle tilfælde har man tilføjet ad hoc-spørgsmål. Databaserne anvender primært PRO som resultatindikator. Der må forventes en stigning i publikationer på området de kommende år, i takt med at flere databaser implementerer PRO.

- Hvis man overvejer at indsamle PRO-data i en klinisk kvalitetsdatabase, er det relevant først at tage stilling til følgende spørgsmål:

- Findes der aspekter (domæner) som ikke kan måles med eksisterende outcomes/indikatorer?
- Er PRO-data egnede til at måle dette aspekt (domæne)?
- Hvilke PRO-data er nødvendige for at måle dette?
- Skal PRO-data også anvendes på individniveau?

- Valg af type af PRO-instrument:

Hvis man skal vælge et PRO-instrument, til brug i en klinisk kvalitetsdatabase, vil det være relevant at medtænke følgende:

- Generiske instrumenter giver mulighed for sammenligning på tværs af patientgrupper.
- Sygdomsspecifikke instrumenter vil ofte opleves meget relevante af patienterne.
- Det anbefales at bruge en grundstamme af generiske og/eller domænespecifikke instrumenter, som kan kombineres med et sygdomsspecifikt instrument eller ad hoc-spørgsmål.
- Enkelstående supplerende ad hoc-spørgsmål kan anvendes, mens ad hoc-instrumenter med sum-scoringsalgoritmer kun bør anvendes efter en detaljeret og specifik valideringsproces.

- Eksisterende eller nye PRO-instrumenter:

Undersøg muligheden for at anvende et eksisterende instrument, da der altid vil være fordele ved at anvende et instrument, der er afprøvet, og hvor der evt. findes referencedata.

- PRO-instrumentets egenskaber:

Indhent dokumentation for validitet og reliabilitet og medtænk i hvilken kontekst instrumentet er testet. Er disse forhold sammenlignelige med de omstændigheder instrumentet skal anvendes under i forbindelse med kvalitetsovervågning? Overvej, om de instrumenter der vælges, kan identificere ændringer i patienternes tilstand og dermed dokumentere, hvorvidt der sker en forbedring/forværring i patientens tilstand over tid.

- Inklusionskriterier:

Der skal tages stilling til, om alle patienter i den pågældende diagnosegruppe skal inkluderes eller der skal udvælges en stikprøve.

- Pilotfasen:

En pilotfase er tilrådelig.

- Ved indsamling af PRO-data skal der tages stilling til:

- Hvor skal PRO-data indsamles - på hospitalet eller hjemme hos patienten?
- Hvornår skal PRO måles?

- Hvordan skal PRO-data indsamles - papirbaseret, internetbaseret eller ved mixed-mode?
- Patientkontakt:
Det skal afgøres, hvordan patienten kontaktes: personligt, via e-mail, digital post, brev eller sms.
- Statistiske analyser:
 - Databasen skal på forhånd overveje hvilke statistiske metoder der skal anvendes, afhængig om det er 1, 2 eller flere PRO-målinger tilgængelige.
 - Der skal tages stilling til valg af referencegruppe før dataindsamlingen påbegyndes.
 - Missing PRO-data er vanskelige at undgå. Procedurer for at begrænse missing data samt procedurer for håndtering af missing data i analysefasen skal besluttes før indsamlingsfasen begynder.
 - Det skal besluttes, hvordan der tages højde for forskelle i patientsammensætning fra afdeling til afdeling.
- Præsentation af PRO-data:
 - PRO-data skal præsenteres både i tabelform og grafisk for at fremme fortolkning og forståelse af data.
 - Både tabeller og grafer skal være selvforklarende.

20. Referencer

Reference List

- (1) Mainz J, Krog BR, Bjornshave B, Bartels P. Nationwide continuous quality improvement using clinical indicators: the Danish National Indicator Project. *Int J Qual Health Care* 2004 Apr;16 Suppl 1:i45-i50.
- (2) Slevin ML, Stubbs L, Plant HJ, Wilson P, Gregory WM, Armes PJ, et al. Attitudes to chemotherapy: comparing views of patients with cancer with those of doctors, nurses, and general public. *BMJ* 1990 Jun 2;300(6737):1458-60.
- (3) Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014 Jun;23(6):508-18.
- (4) U.S. Department of Health and Human Services Food and Drug Administration. Guidance for industry. patient-reported outcome measures: Use in medical product development to support labeling claims. 2009.
- (5) Fayers PM, Machin D. *Quality of life: The assessment, analysis and interpretation of patient-reported outcomes*. Chichester: John Wiley & Sons Ltd; 2007; 2007.
- (6) National Quality Forum. Patient reported outcomes (PROs) in performance measurement. 2013.
- (7) Cella D, Hanh E, Jensen S, Butt Z, Nowinski C, Rothrock N. Methodological issues in the selection, administration and use of patient-reported outcomes in performance measurement in health care settings. 2012.
Ref Type: Online Source
- (8) Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA* 1995 Jan 4;273(1):59-65.
- (9) International classification of functioning, disability and health (ICF) . 2014.
Ref Type: Online Source
- (10) Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. *Qual Life Res* 2008 Nov;17(9):1125-35.
- (11) Digitalisering med effekt: national strategi for digitalisering af sundhedsvæsenet 2012-2017. Regeringen, KL, Danske Regioner; 2013 Jan 1.
- (12) Databasernes fællessekretariat. PRO-målinger og RKKP. Oversigt over projekter og aktiviteter. Anbefalinger af fremtidig strategi. 2014.
- (13) Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Serv Res* 2013;13:211.
- (14) Boyce MB, Browne JP. Does providing feedback on patient-reported outcomes to healthcare professionals result in better outcomes for patients? A systematic review. *Qual Life Res* 2013 Nov;22(9):2265-78.
- (15) Coates A, Porzsolt F, Osoba D. Quality of life in oncology practice: prognostic value of EORTC QLQ-C30 scores in patients with advanced malignancy. *Eur J Cancer* 1997 Jun;33(7):1025-30.

- (16) Grosse FK, van der Meulen J, Black N. Relationship between patients' reports of complications and symptoms, disability and quality of life after surgery. *Br J Surg* 2012 Aug;99(8):1156-63.
- (17) Rothwell AG, Hooper GJ, Hobbs A, Frampton CM. An analysis of the Oxford hip and knee scores and their relationship to early joint revision in the New Zealand Joint Registry. *J Bone Joint Surg Br* 2010 Mar;92(3):413-8.
- (18) Varaganam M, Hutchings A, Neuburger J, Black N. Impact on hospital performance of introducing routine patient reported outcome measures in surgery. *J Health Serv Res Policy* 2014 Apr;19(2):77-84.
- (19) Varaganam M, Hutchings A, Black N. Do patient-reported outcomes offer a more sensitive method for comparing the outcomes of consultants than mortality? A multilevel analysis of routine data. *BMJ Qual Saf* 2015 Mar;24(3):195-202.
- (20) Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346:f167.
- (21) Weingarten SR, Kim CS, Stone EG, Kristopaitis RJ, Pelter M, Sandhu M. Can peer-comparison feedback improve patient functional status? *Am J Manag Care* 2000 Jan;6(1):35-9.
- (22) Manual for udvikling af sygdomsspecifikke indikatorer, standarder og prognostiske faktorer i Det Nationale Indikatorprojekt. http://www.kcks-vest.dk/siteassets/manualer/20090609_indikatormanual_v3.pdf 2015
- (23) Lo IK, Griffin S, Kirkley A. The development of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder: The Western Ontario Osteoarthritis of the Shoulder (WOOS) index. *Osteoarthritis Cartilage* 2001 Nov;9(8):771-8.
- (24) Moorehead MK, Ardelt-Gattinger E, Lechner H, Oria HE. The validation of the Moorehead-Ardelt Quality of Life Questionnaire II. *Obes Surg* 2003 Oct;13(5):684-92.
- (25) Digitaliseringsstyrelsen. Vejledning i informations-sikkerhedsstyring (ISMS). 2015.
- (26) Rasmussen JV, Jakobsen J, Olsen BS, Brorson S. Translation and validation of the Western Ontario Osteoarthritis of the Shoulder (WOOS) index - the Danish version. *Patient Relat Outcome Meas* 2013;4:49-54.
- (27) Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes* 2003;1:64.
- (28) Groenvold M, Petersen MA, Aaronson NK, Arraras JI, Blazeby JM, Bottomley A, et al. The development of the EORTC QLQ-C15-PAL: a shortened questionnaire for cancer patients in palliative care. *Eur J Cancer* 2006 Jan;42(1):55-64.
- (29) Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. *Am J Obstet Gynecol* 2003 Jul;189(1):98-101.
- (30) Hetland ML. DANBIO--powerful research database and electronic patient record. *Rheumatology (Oxford)* 2011 Jan;50(1):69-77.
- (31) Hutchings A, Neuburger J, Grosse FK, Black N, van der Meulen J. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. *Health Qual Life Outcomes* 2012;10:34.
- (32) Health and Social Care Information Centre. Provisional monthly patient reported outcome measures (PROMs) in England. A guide to PROMs methodology. 2013.
- (33) Devlin N, Appleby J. Getting the most out of PROMS. Putting health outcomes at the heart of NHS decision-making. The Kings Fund 2010; 2010.

- (34) Emilsson L, Lindahl B, Koster M, Lambe M, Ludvigsson JF. Review of 103 Swedish healthcare quality registries. *J Intern Med* 2015 Jan;277(1):94-136.
- (35) Nilsson E, Lindblom H. Oversikt av PROM och PREM inom de nationella kvalitetsregistren 2015. 2015.
- (36) Annual report 2012. Swedish Hip Arthroplasty Register. 2013.
- (37) Rolfson O, Karrholm J, Dahlberg LE, Garellick G. Patient-reported outcomes in the Swedish Hip Arthroplasty Register: results of a nationwide prospective observational study. *J Bone Joint Surg Br* 2011 Jul;93(7):867-75.
- (38) Svenske gastro- og oesophagal cancer register 2014. 2015.
- (39) New Zealand Joint Registry. New Zealand Orthopaedic Association. Fifteen year report. 2014.
- (40) Basch E, Snyder C, McNiff K, Brown R, Maddux S, Smith ML, et al. Patient-reported outcome performance measures in oncology. *J Oncol Pract* 2014 May;10(3):209-11.
- (41) National Quality Forum. National Quality Forum. Patient Reported Outcomes (PROs) Workshop 1, July 30.31, 2012. Workshop summary. 2015.
- (42) Advances in the use of patient reported outcome measures in electronic health records. Including case studies. November 7, 2013. 2015.
- (43) Bowling A. Measuring disease: A review of disease specific quality of life measurement scales. 2 ed. 2001.
- (44) Hjollund NH, Andersen JH, Bech P. Assessment of fatigue in chronic disease: a bibliographic study of fatigue measurement scales. *Health Qual Life Outcomes* 2007;5:12.
- (45) Cappelleri JC, Zou KH, Bushmakina A, Alvir MJM, Symonds T. Patient-reported outcomes: Measurement, implementation and interpretation. CRC Press 2013; 2013.
- (46) Cook KF, O'Malley KJ, Roddey TS. Dynamic assessment of health outcomes: time to let the CAT out of the bag? *Health Serv Res* 2005 Oct;40(5 Pt 2):1694-711.
- (47) Process of translation and adaptation of instruments. Available from. http://www.who.int/substance_abuse/research_tools/translation/en 2014
- (48) Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2(14):i-74.
- (49) Aadahl M, Lund H. Grundliggende principper for valg og anvendelse af test og målemetoder i fysioterapi. *Forskning i Fysioterapi* 2003;20(October 1):1-9.
- (50) Rolstad S, Adler J, Ryden A. Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value Health* 2011 Dec;14(8):1101-8.
- (51) Nilsson E, Wenemark M, Bendtsen P, Kristenson M. Respondent satisfaction regarding SF-36 and EQ-5D, and patients' perspectives concerning health outcome assessment within routine health care. *Qual Life Res* 2007 Dec;16(10):1647-54.
- (52) Hjollund NH, Larsen LP, Biering K, Johnsen SP, Riiskjaer E, Schougaard LM. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. *Interact J Med Res* 2014;3(1):e5.
- (53) Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome

- (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009 Jun;12(4):419-29.
- (54) User's guide to implementing patient-reported outcomes assessment in clinical practice. Available from. 2011.
Ref Type: Online Source
- (55) Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a randomized comparison (2). *J Med Internet Res* 2004 Sep 24;6(3):e30.
- (56) Balter KA, Balter O, Fondell E, Lagerros YT. Web-based and mailed questionnaires: a comparison of response rates and compliance. *Epidemiology* 2005 Jul;16(4):577-9.
- (57) Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. *J Med Internet Res* 2007;9(3):e25.
- (58) Zuidgeest M, Hendriks M, Koopman L, Spreeuwenberg P, Rademakers J. A comparison of a postal survey and mixed-mode survey using a questionnaire on patients' experiences with breast care. *J Med Internet Res* 2011;13(3):e68.
- (59) van den Berg MH, Overbeek A, van der Pal HJ, Versluys AB, Bresters D, van Leeuwen FE, et al. Using web-based and paper-based questionnaires for collecting data on fertility issues among female childhood cancer survivors: differences in response characteristics. *J Med Internet Res* 2011;13(3):e76.
- (60) Hohwu L, Lyshol H, Gissler M, Jonsson SH, Petzold M, Obel C. Web-based versus traditional paper questionnaires: a mixed-mode survey with a Nordic perspective. *J Med Internet Res* 2013;15(8):e173.
- (61) Kirkwood BR, Sterne AC. *Essential Medical Statistics*. second ed. 2003.
- (62) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992 Jun;30(6):473-83.
- (63) Carpenter J, Kenward M. *Multiple imputation and its application*. 2013.
- (64) King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res* 2011 Apr;11(2):171-84.
- (65) Tubach F, Giraudeau B, Ravaud P. The variability in minimal clinically important difference and patient acceptable symptomatic state values did not have an impact on treatment effect estimates. *J Clin Epidemiol* 2009 Jul;62(7):725-8.
- (66) Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Ann Rheum Dis* 2005 Jan;64(1):29-33.
- (67) Paulsen A, Roos EM, Pedersen AB, Overgaard S. Minimal clinically important improvement (MCII) and patient-acceptable symptom state (PASS) in total hip arthroplasty (THA) patients 1 year postoperatively. *Acta Orthop* 2013 Nov 29.
- (68) Robert M, Denis A, Badol-Van SP, Jaisson-Hot I, Guillaat C. Prospective longitudinal assessment of change in health-related quality of life after adjustable gastric banding. *Obes Surg* 2013 Oct;23(10):1564-70.
- (69) Dansk Skulderalloplastik Register. Årsrapport 2014. 2015.