Maastricht University

Patient-oriented outcome assessment in rheumatic diseases

Citation for published version (APA):

Bakker, C. H. (1995). *Patient-oriented outcome assessment in rheumatic diseases*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. https://doi.org/10.26481/dis.19950616cb

Document status and date:
Published: 01/01/1995

DOI:
10.26481/dis.19950616cb

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 26 Apr. 2024

SUMMARY

This thesis concerns patient-oriented outcome assessment in rheumatic diseases.
Chapter 1 provides an introduction and guide to this thesis. Measuring outcome or health status is important for assessing the impact of chronic diseases on the patient. Comprehensive outcome or endpoint measuring should include all those components of health status that are important to the patient and physician and that are relevant to the intervention assessed.
Health status can be described in 5 dimensions: Death, Disability, Discomfort, Drug (or therapeutic toxicity), Dollar costs. The scope of health status can be either specifically oriented, focussing on only one dimension, or broadly oriented, focussing on several dimensions. Further, different diseases may deduct different aspects of health status, whereas individual patients may have different priorities regarding their health status. Physician-assessed and patient-reported measures can be classified as either specific or generic in each of 3 areas of focus: health status, disease and patient (See cover). A measure is specific in the health status area when it measures only one dimension, and generic when it measures several dimensions. A measure is specific in the disease area when it is applicable to only 1 disease (e.g., ankylosing spondylitis). It becomes less focussed when it is applied in a group of diseases (e.g., all arthritides, all cancers) and completely generic when applicable to all diseases. Similarly, measures that are specific in the patient area refer to single patients. Less specifically focussed measures refer to subgroups (e.g., the elderly) and generic measures refer to all possible patients. Obviously, outcome measures are multidimensional and can be generic in one area and focussed in another.

Chapter 2 reviews outcome measures used today in ankylosing spondylitis. Of the 43 studies reviewed, 79% contained physician assessed measures. Patient reported measures were mentioned in 67%. Most physician assessed measures (67%) focussed on disability (spinal mobility), most patient reported measures (65%) focussed on discomfort (pain and stiffness). Single item global assessment by physician or patient, the most generic measure, was reported in 16% and in 40% of the studies respectively. Other general measures of health status were only occasionally used. In particular, self-assessed measurement of functional ability was frequently lacking. Only 1 study reported a measure which specifically addressed the patient's priorities regarding treatment risks, and only 3 studies included an economic analysis. Drug toxicity reports missed informative detail. This review clearly showed that some areas of the total spectrum of health status were not well covered, whereas other areas were overrepresented.

The chapters 3 to 8 deal with the measurement of patient priorities.
Chapter 3 gives an introduction to utility assessment in rheumatology. The unique feature of utility measurement is that it is generic in the health and disease area of the cube figure (See cover). In the patient area it is focussed when measuring utilities of individual patients and generic when eliciting utilities from the general population. The utility allows for comparison across different diseases and interventions. This chapter describes 2 approaches to utility measurement. The first approach is to classify patients into categories based on their responses to questions about their functional status (for example, the Quality of Well being questionnaire and the Sickness Impact Profile). The second approach to utility measurement is to ask the patients to directly assign one value to their overall health. The 4 methods used most frequently (rating scale, standard gamble, time trade-off and willingness to pay) were described.

Chapter 4 describes the feasibility of utility assessment by rating scale and standard gamble method in patients with ankylosing spondylitis (AS) or fibromyalgia (FMS). In the context of 2 randomized controlled trials the Maastricht Utility Measurement Questionnaire (MUMQ) was applied by trained interviewers to 59 patients with AS at the 9 months' followup and 86 patients with FMS at baseline. The MUMQ is a Dutch translation and adaptation of the McMaster Utility Measurement Questionnaire developed by Bennett and Torrance and consists of a rating scale and standard gamble method. The feasibility of these methods in both patients with AS and FMS was generally satisfactory. All patients completed the interview. Four (2.8%) of all 143 patients gave inconsistent answers: one on the rating scale and 3 on the standard gamble. The duration of the baseline interview was about 9-12 minutes (SD 3.6-5.4) for the rating scale and about 12-14 minutes (SD 2.9-3.8) on the standard gamble. Four weeks test-retest reliability for the patient's own health state measured by intraclass correlation coefficients was 0.56 for the rating scale and 0.66 for the standard gamble technique. Patients with AS valued their personal health state on the rating scale (0-100) considerably higher than patients with FMS (AS:69 and FMS:54). Standard gamble utility values (0-1), however, were about the same at a higher level (AS:0.86 and FMS:0.83). These data supported the view that utility measurement is sensitive to the method chosen to elicit patient well being.

Chapter 5 and 6 compare utilities derived by rating scale and standard gamble method in patients with AS and FMS, relate these values to other outcome measures, and describe the sensitivity to change of utilities relative to changes in other outcomes.
In both patient groups the MUMQ was applied twice: first at baseline and again after 6 (FMS) or 9 (AS) months' followup. In AS the analysis was restricted to the 59 patients who were seen by the same interviewer at both assessments. In FMS a total of 73 patients was assessed by the same interviewer at both assessments.
Construct validity seemed to be higher for the rating scale than for the standard gamble method in both patients with AS or FMS. For the rating scale Spearman correlation coefficients ranged from 0.25 to 0.68 in AS and from 0.23 to 0.62 in FMS; for the standard gamble it ranged from 0.23 to 0.36 in both diseases. Regression analyses showed that patient's global assessment explained 41-59% of rating scale utilities and 10-11% of standard gamble utilities in patients with AS or FMS. These findings support the view that rating scale utilities more closely resemble global assessment. The methodological advantage of standardized rating scale utility measurement over non-standardized global assessment is that utilities provide numerical values which allows patient outcomes of different diseases or resulting from various health care interventions to be compared across patients and diseases.

In both diseases, changes in rating scale utilities correlated to a higher degree with changes in other outcomes than changes in standard gamble utilities.

Chapter 7 provides a detailed description of methodological issues of patient utility measurement in 2 randomized controlled trials involving 144 patients with AS and 85 patients with FMS. In both trials the MUMQ was applied at baseline and at 2 followup assessments. Patients were asked to value their own health state and a light, moderate and severe marker or reference state by means of the rating scale and standard gamble method.
It was confirmed that standard gamble scores are consistently higher than rating scale scores for both the patient's own health state and the marker states. The 3 months' test-retest reliability for the reference states measured by intraclass correlation coefficients ranged from 0.24 to 0.33 using the rating scale and 0.43 to 0.70 on the standard gamble. Although the reproducibility is not high, the mean scores are fairly stable over time. Mean standard gamble scores tend to differ depending on how the measurements are taken. Utilities elicited with 'chained gambles' were significantly higher than utilities elicited with 'basic reference gambles'. On the individual level some inconsistent responses occurred. However, more than 70% of them fell within the bounds of the measurement error which ranged from 0.11 to 0.13 on the standard gamble (0-1 scale) and from 8 to 10 on the rating scale (0-100 scale). The large number of negative utilities for the severe marker state, which was used as an anchor point in the chained gambles, and the magnitude of these negative utilities (down to -19) lead us to favor using death as the anchor point in the standard gamble.

Chapter 8 describes another way to assess patient priorities by means of the Problem Elicitation Technique (PET) questionnaire. This questionnaire deals only with activities that are directly limited by the disease and judged important by the patient. The PET questionnaire was applied by trained interviewers to 134 patients with AS and 73 patients with FMS at baseline and at 9 and 6 months' followup respectively. A PET score was calculated at each assessment.
Patients with FMS identified more problems (mean 6.8) than patients with AS (mean 4.4). Moreover, patients with AS could more often not identify any problem at baseline than patients with FMS (10% compared to 1%). The PET score improved from 14.9 to 11.3 ($p=0.0001$) in AS patients but did not change from 21.8 to 21.1 ($p=0.24$) in FMS patients. Construct validity testing of the PET score showed statistically significant ($p<0.05$) correlations with Arthritis Impact Measurement Scale (AIMS), utilities, Sickness Impact Profile (SIP), Health Assessment Questionnaire (HAQ), pain, stiffness, and patient's global health in both AS and FMS patients (r's varying from 0.22 to 0.66). By multiple regression pain explained 29% of the variance in PET scores among patients with AS. In FMS, patient's global assessment accounted for 39% of total variance of PET scores, whereas pain explained another 15%. Changes in PET scores correlated significantly ($p<0.05$) with changes in AIMS, utilities, pain, stiffness, and patient's global health in both patients with AS and FMS (r's varying from 0.22 to 0.51). Some 6% of the variance in changes in PET scores was explained by changes in pain in patients with AS and for 35% by changes in pain and subjective health in patients with FMS. Assessment of sensitivity to change revealed that the efficiency of the PET score was 0.6 in patients with AS and 0.09 in patients with FMS. Compared to other outcomes this was reasonable in patients with AS but low in patients with FMS.
It was generally feasible to obtain patient priorities. In both patients with AS or FMS construct validity of the PET questionnaire was satisfactory. The PET was much more sensitive to change in patients with AS than in patients with FMS.

Chapter 9 provides a comparison of the cost-effectiveness of supervised group therapy with that of exercises at home in 144 patients with AS. Compared to exercises at home, group therapy produces an extra increase of 7% in spinal mobility and of 5% in fitness, and an extra improvement of 28% in global health at a cost of $531 per year, while reducing direct medical costs by $122 per year. Hence, the beneficial effects of group therapy cost $409 per patient with AS on a yearly base or about $10 per weekly therapy session (40 sessions per year).

Chapter 10 provides a general discussion of the study described in this thesis.

SAMENVATTING

Dit proefschrift behandelt patiënt-georiënteerde uitkomstmetingen bij reumatische aandoeningen. **Hoofdstuk 1** geeft een inleiding tot dit proefschrift. Het is van belang de gezondheidstoestand of de 'uitkomst' (outcome) te meten, om zo de impact van chronische aandoeningen op de patiënt vast te stellen. Om de uitkomst, of het eindpunt goed te kunnen bepalen moet het die aspecten van de gezondheidstoestand omvatten die belangrijk zijn voor de patiënt en de arts en tevens moeten ze relevant zijn voor de meting. Uitkomstmetingen kunnen worden ingedeeld als door de arts gemeten, door de patiënt gerapporteerd of 'overig' (bijvoorbeeld laboratorium waarden en röntgen foto's).
De gezondheidstoestand kan beschreven worden in vijf dimensies: dood, verminderd functioneren, klachten of symptomen zoals pijn en stijfheid, bijwerkingen van medicijnen of van de behandeling, en kosten van allerlei aard. De gezondheidstoestand kan specifiek worden benaderd door te focussen op één dimensie, of breed worden benaderd door verschillende dimensies tegelijk te onderzoeken. Bovendien kunnen verschillende ziektes verschillende aspecten van de gezondheidstoestand aantasten en individuele patiënten kunnen verschillende prioriteiten hebben ten aanzien van hun gezondheidstoestand. Daarom kunnen uitkomstmetingen geclassificeerd worden als specifiek of algemeen in elk van drie gebieden: gezondheidstoestanden, ziekten of aandoeningen, en patiënten (zie omslag van dit proefschrift). Een maat is specifiek in het gezondheidstoestand-gebied als het slechts één dimensie meet, en algemeen als het verschillende dimensies meet. Een maat is specifiek in het ziekte- of aandoening-gebied als het toepasbaar is op maar één aandoening (bijvoorbeeld spondylitis ankylopoetica). De maat wordt minder specifiek als het toepasbaar is op een groep van aandoeningen (bijvoorbeeld alle vormen van artritis, of alle vormen van kanker) en volledig algemeen als het toepasbaar is op alle mogelijke aandoeningen. Zo ook refereren maten die specifiek zijn in het patiënt-gebied naar individuele patiënten. Minder specifiek gerichte maten refereren naar subgroepen (bijvoorbeeld de ouderen) en algemene maten refereren naar alle mogelijk patiënten. Uitkomstmetingen kunnen algemeen zijn in één gebied en specifiek in een ander gebied.

Hoofdstuk 2 geeft een overzicht van uitkomstmetingen met betrekking tot de gezondheidstoestand die vandaag de dag worden gebruikt in spondylitis ankylopoetica. Van de 43 besproken studies werd in 79% gebruik gemaakt van maten die door de arts werden gemeten. In 67% werd gebruik gemaakt van maten die door de patiënt werden gerapporteerd. De meeste van de door de arts gemeten maten (67%) richtten zich op fysieke beperkingen (beweeglijkheid van de wervelkolom). De meeste van de door de patiënt gerapporteerde maten (65%) richtten zich op symptomen zoals pijn en stijfheid. De meest algemene maat, 'global

assessment' door arts of patiënt uitgedrukt in één item, werd in respectievelijk 16% en in 40% van de studies gerapporteerd. Andere algemene maten die de gezondheidstoestand meten werden maar af en toe gebruikt. Met name ontbrak het vaak aan de maten die door de patiënt worden gerapporteerd en die het fysieke functioneren van de patiënt meten. Er was slechts één studie die de prioriteiten van de patiënt ten aanzien van behandelingsrisico's rapporteerde en slechts drie studies omvatten een economische analyse. De rapportage van bijwerkingen van medicijnen of van de behandeling liet te wensen over. Dit literatuuroverzicht laat duidelijk zien dat sommige gebieden van de gezondheidstoestand niet worden gedekt met de gebruikte uitkomstmetingen en dat andere gebieden relatief veel aandacht krijgen.

De hoofdstukken 3 tot en met 8 behandelen prioriteiten van patiënten.
Hoofdstuk 3 geeft een inleiding tot utiliteitsmeting in de reumatologie. Het unieke van utiliteitsmaten is dat deze algemeen zijn in zowel het gezondheidstoestand-gebied als in het ziekte- of aandoening-gebied. In het patiënt-gebied echter is het specifiek wanneer utiliteiten van individuele patiënten worden gemeten, doch algemeen wanneer utiliteiten van de algemene bevolking worden verkregen. Met de utiliteit kunnen vergelijkingen worden gemaakt tussen verschillende aandoeningen en interventies. Dit hoofdstuk beschrijft twee methoden van utiliteitsmeting. Bij de eerste methode worden patiënten geclassificeerd in categorieën die gebaseerd zijn op antwoorden op vragen over het functioneren van patiënten (bijvoorbeeld, de 'Quality of Well-Being questionnaire' en de 'Sickness Impact Profile'). Bij de tweede methode van utiliteitsmeting wordt de patiënt zelf direct gevraagd een waardering te geven over zijn of haar algehele gezondheid. De vier meest gebruikte methoden (rating scale, standard gamble, time trade-off en willingness to pay) worden beschreven.

Hoofdstuk 4 beschrijft de bruikbaarheid van utiliteitsmeting met de rating scale en de standard gamble methode bij patiënten met spondylitis ankylopoetica (SA) of fibromyalgie (FMS). In het kader van twee gerandomiseerde gecontroleerde studies (trials) werd door getrainde interviewers de 'Maastricht Utility Measurement Questionnaire' (MUMQ) afgenomen bij 59 patiënten met SA ten tijde van de negen maanden followup en bij 86 patiënten met FMS op de baseline meting. De MUMQ is een Nederlandse aangepaste vertaling van de 'McMaster Utility Measurement Questionnaire' die door Bennett en Torrance werd ontwikkeld. De MUMQ omvat zowel de rating scale als de standard gamble methode. De bruikbaarheid van beide methodes bij zowel patiënten met SA als patiënten met FMS was over het algemeen bevredigend. Alle patiënten voltooiden het interview. Vier (2.8%) van alle 143 patiënten gaven inconsistente antwoorden: één op de rating scale en drie op de standard gamble. De duur van het interview was ongeveer 9-12 minuten (SD 3.6-5.4) met de rating scale en ongeveer 12-14 minuten (SD 2.9-3.8) met de standard gamble methode. De vier-week test-hertest betrouwbaarheid voor de gezondheidstoestand van de patiënt, uitgedrukt als intraclass correlatie coëfficiënt, was 0.56 voor de rating scale en 0.66 voor de standard gamble. Op de rating scale (0-100) waardeerden patiënten met SA hun eigen gezondheidstoestand aanzienlijk hoger dan patiënten met FMS (SA:69 en FMS:54). Echter, de standard gamble utiliteitswaarden (0-1) waren voor beide aandoeningen ongeveer gelijk doch op een hoger niveau (SA: 0.86 en FMS:0.83). Deze data ondersteunen de opvatting dat het resultaat van de utiliteitsmeting afhankelijk is van de methode die gekozen wordt om het welbevinden van de patiënt te meten.

In **hoofdstuk 5 en 6** worden utiliteiten van patiënten met SA of FMS die verkregen zijn met de rating scale methode, vergeleken met utiliteiten van deze patiënten die verkregen zijn met de

standard gamble methode. Vervolgens worden deze utiliteiten gerelateerd aan andere uitkomstmaten, en wordt de gevoeligheid voor verandering van utiliteiten vergeleken met de gevoeligheid voor verandering van andere uitkomstmaten. De MUMQ werd twee keer afgenomen bij beide patiëntengroepen: ten eerste op baseline en ten tweede na zes (FMS) of negen (SA) maanden followup. Bij SA werd de analyse beperkt tot slechts 59 van de 135 patiënten die door dezelfde interviewer bij beide metingen werden gezien. Bij FMS werden alle 73 patiënten door dezelfde interviewer gemeten op beide meetmomenten. De construct validiteit leek bij zowel patiënten met SA als patiënten met FMS hoger voor de rating scale dan voor de standard gamble methode. Voor de rating scale methode varieerden de Spearman correlatie coëfficiënten van 0.25 tot 0.68 bij patiënten met SA en van 0.23 tot 0.62 bij patiënten met FMS; voor de standard gamble methode varieerden deze van 0.23 tot 0.36 bij beide aandoeningen. Regressie analyses lieten zien dat de 'global assessment' van de patiënt 41 - 59% van de rating scale utiliteiten verklaarden en 10 - 11% van de standard gamble utiliteiten bij patiënten met SA of FMS. Deze bevindingen ondersteunen de gedachte dat utiliteiten verkregen via de rating scale methode de 'global assessment' dichter benaderen. Het methodologische voordeel van gestandaardiseerde rating scale utiliteitsmeting boven niet-gestandaardiseerde 'global assessment' is dat utiliteiten numerieke waarderingen opleveren die het mogelijk maakt patiënt-uitkomsten van verschillende aandoeningen of patiënt-uitkomsten die resulteren van verschillende interventies te vergelijken. Bij beide aandoeningen correleerden veranderingen in rating scale utiliteiten beter met veranderingen in andere uitkomstmaten dan veranderingen in standard gamble utiliteiten.

Hoofdstuk 7 behandelt diverse methodologische aspecten van utiliteitsmetingen met de rating scale en de standard gamble methode. Deze methodologische aspecten werden gedestilleerd uit twee gerandomiseerde klinische trials waaraan respectievelijk 85 FMS patiënten en 144 SA patiënten deelnamen. In beide trials vond één baseline meting en twee followup metingen plaats waarin patiënten gevraagd werden hun waarderingen uit te spreken over hun eigen gezondheidstoestand en enkele hypothetische gezondheidstoestanden (licht, matig, en ernstig ziek).
De standard gamble methode bleek consistent tot hogere scores te leiden dan de rating scale, zowel voor de eigen gezondheidstoestand als voor de hypothetische gezondheidstoestanden. De drie-maand test-hertest betrouwbaarheid voor de hypothetische toestanden uitgedrukt als intraclass correlatie coëfficiënt varieerde van 0.24 tot 0.33 voor de rating scale en van 0.43 tot 0.70 voor de standard gamble. Hoewel de reproduceerbaarheid niet hoog was, bleven de groepsgemiddelden in de loop van de tijd relatief constant. De gemiddelde standard gamble score lijkt afhankelijk te zijn van de wijze waarop de metingen worden uitgevoerd. Utiliteiten gemeten met de zogenaamde 'chained gambles' waren significant hoger dan utiliteiten gemeten met de 'basic reference gamble'. Op het individuele niveau werden inconsistente antwoorden gevonden. Meer dan 70% van deze inconsistente antwoorden viel echter binnen de grenzen van de meetfout, die varieerde van 0.11 tot 0.13 op de standard gamble (0-1 schaal) en van 8 tot 10 op de rating scale (0-100 schaal). Het grote aantal negatieve utiliteiten voor de hypothetische toestand van een ernstig zieke - de toestand die als ankerpunt in de 'chained gambles' was gebruikt - en de waarde van de negatieve utiliteiten (tot -19) onderstrepen de noodzaak betere methoden te ontwikkelen voor het meten van negatieve utiliteiten en het hanteren van deze utiliteiten in 'Quality Adjusted Life Years' (QALY) berekeningen.

Hoofdstuk 8 beschrijft een andere manier om prioriteiten van patiënten vast te stellen en wel met behulp van de 'Problem Elicitation Technique (PET) questionnaire'. Deze vragenlijst behandelt alleen die activiteiten die direct door de aandoening worden beperkt en als belangrijk worden ervaren door de patiënt. De PET vragenlijst werd door getrainde interviewers afgenomen bij 134 patiënten met SA en 73 patiënten met FMS op de baseline meting en na respectievelijk negen en zes maanden followup. Op iedere meting werd een PET score berekend.
Patiënten met FMS gaven meer problemen aan (gemiddeld 6.8) dan patiënten met SA (gemiddeld 4.4). Bovendien konden patiënten met SA vaker geen probleem noemen op de baseline meting dan patiënten met FMS (10% in vergelijking tot 1%). De PET score verbeterde van 14.9 naar 11.3 ($p=0.0001$) bij patiënten met SA, maar veranderde niet (van 21.8 naar 21.1) ($p=0.24$) bij patiënten met FMS.
De construct validiteit van de PET score werd getoetst middels correlaties tussen de PET score en andere effectmaten. De PET score correleerde statistisch significant met de 'Arthritis Impact Measurement Scale' (AIMS), utiliteiten, 'Sickness Impact Profile' (SIP), 'Health Assessment Questionnaire' (HAQ), pijn, stijfheid, en de algehele gezondheidstoestand van de patiënt bij zowel patiënten met SA als FMS (r's varieerden van 0.22 tot 0.66). In een multiple regressie analyse verklaarde pijn 29% van de variantie in de PET scores bij patiënten met SA. Bij FMS verklaarde de 'global assessment' van de patiënt 39% van de totale variantie in PET scores, terwijl pijn nog eens 15% verklaarde. Veranderingen in PET scores correleerden significant ($p<0.05$) met veranderingen in AIMS, utiliteiten, pijn, stijfheid en de algehele gezondheidstoestand van de patiënt. Dit betrof zowel patiënten met SA als patiënten met FMS (r's varieerden van 0.22 tot 0.51). Veranderingen in PET scores werden voor 6% verklaard door veranderingen in pijn bij patiënten met SA en voor 35% door veranderingen in pijn en subjectieve gezondheid bij patiënten met FMS. Het meten van de gevoeligheid voor verandering toonde aan dat de efficiëntie van de PET score 0.6 was bij patiënten met SA en 0.09 bij patiënten met FMS. Vergeleken met andere uitkomstmaten is dit redelijk voor patiënten met SA, maar laag voor patiënten met FMS.
Over het algemeen was het goed mogelijk om prioriteiten van de patiënt in beeld te krijgen. Bij zowel patiënten met SA als FMS was de construct validiteit van de PET vragenlijst bevredigend. De PET was veel gevoeliger voor verandering bij patiënten met SA dan bij patiënten met FMS.

Hoofdstuk 9 bespreekt de kosten-effectiviteit van groepsoefentherapie bij patiënten met SA. Vergeleken met thuis oefenen levert groepsoefentherapie na negen maanden een extra verbetering op van 7% in beweeglijkheid van de wervelkolom, 5% extra verbetering in fitheid en 28% extra verbetering in globale gezondheid voor een bedrag van 1001 gulden per jaar. De directe medische kosten (bezoeken aan de huisarts, specialist en paramedicus, het gebruik van medicijnen en het aantal dagen opgenomen in het ziekenhuis) worden met 230 gulden per jaar gereduceerd. De effecten van groepsoefentherapie kosten dus 771 gulden per jaar, oftewel bijna 20 gulden per week.

Hoofdstuk 10 plaatst enkele kanttekeningen bij het onderzoek beschreven in dit proefschrift.