

Laatupainotetut elinvuodet terveydenhuollon vaikuttavuuden arvioinnissa:

järjestelmällinen kirjallisuuskatsaus

Use of Quality-Adjusted Life Years for the Estimation of Effectiveness of Health Care:

A Systematic Literature Review

Laatupainotettuja elinvuosia pidetään yhtenä parhaimmista menetelmistä arvioida hoidon vaikuttavuutta. Tästä huolimatta tutkimuksia, joissa hoidon lopputulos ilmaistaan laatupainotettuina elinvuosina siten, että vaikuttavuuden arvioijina toimivat potilaat, löytyy vielä vähän.

Tämän raportin tarkoituksena on kuvata tutkimukset, jotka ovat käyttäneet potilaan itsensä arvioimaa elämänlaatua QALY- laskelmien perustana. Järjestelmälliseen kirjallisuuskatsaukseen on kerätty tieto tutkimusten lääketieteen erikoisalasta, tutkitusta toimenpiteestä, tutkimuksen tuloksista, käytössä olleesta elämänlaatumittarista, tutkimuksen laadusta ja alkuperämaasta, todetuista laatupainotetuista lisäelinvuosista ja tulosten tulkinnasta kustannusvaikuttavuuden suhteen.

Raporttia voivat hyödyntää sekä terveydenhuollon ammattilaiset että päättäjät arvioidessaan eri hoitomuotojen vaikuttavuutta ja päättäessään terveydenhuollon voimavarojen kohdentamisesta.

The quality-adjusted life year is considered one of the most important indicators of effectiveness of health care.

Even so, the number of studies in which treatments outcomes are based on actual measurements of patients' health-related quality of life, is still fairly limited.

The objective of this systematic literature review was to identify published studies having used quality-adjusted life years based on actual measurements of patients' health-related quality of life, and to characterize the studies regarding the health-related quality of life instrument used, medical specialty, intervention studied, results obtained, quality, country of origin, quality-adjusted life year gain observed, and interpretation of results regarding cost-effectiveness.

We hope that health care professionals as well as decision makers both can benefit from this report when evaluating various treatments and deciding on allocation of resources.

Raporttia voi tilata maksutta Finohtasta./ The report can be ordered free of charge from Finohta.

STAKES / FinOHTA • P. O. Box 220. FI-00531 Helsinki
Tel. (09) 3967 2396, fax: (09) 39672278. E-mail: finohta@stakes.fi

Raportti on saatavilla Finohtan sivuilla/The report is available for downloading from the Finohta website: <http://www.stakes.fi/finohta>



FinOHTA
Terveydenhuollon menetelmien
arviointiyksikkö
Finnish Office for Health
Technology Assessment

ISBN 951-33-1850-8
ISSN 1239-6273



Sosiaali- ja terveysalan
tutkimus- ja kehittämiskeskus
National Research and Development
Centre for Welfare and Health

postimyynti/
sales Stakes / Asiakaspalvelut
PL 220, 00531 Helsinki
+358 (9) 3967 2190,
puhelin/tel.: +358 (9) 3967 2308
(automaatti/automate)
+358 (9) 3967 2450

faksi/fax:
internet: www.stakes.fi



Pirjo Räsänen, Eija Roine, Harri Sintonen, Virpi Semberg-Konttinen,
Olli-Pekka Ryyänen, Risto P. Roine

Laatupainotetut elinvuodet terveydenhuollon
vaikuttavuuden arvioinnissa:
järjestelmällinen kirjallisuuskatsaus

Use of Quality-Adjusted Life Years for the
Estimation of Effectiveness of Health Care:
A Systematic Literature Review

Finohtan raportti/Finohta's report

29/2006



TYÖRYHMÄ

Pirjo Räsänen	tutkija, Helsingin ja Uudenmaan sairaanhoitopiiri ja Terveystieteiden tutkimuskeskitys, Helsingin ja Uudenmaan sairaanhoitopiiri ja Terveystieteiden tutkimuskeskitys
Eija Roine	lääketieteen opiskelija
Harri Sintonen	professori, Terveystieteiden tutkimuskeskitys, Helsingin ja Uudenmaan sairaanhoitopiiri ja Terveystieteiden tutkimuskeskitys
Virpi Semberg-Konttinen	kuntayhtymätarkastaja, Helsingin ja Uudenmaan sairaanhoitopiiri
Olli-Pekka Ryyänen	professori, Kuopion yliopisto, Kansanterveystieteen ja yleislääketieteen laitos
Risto P. Roine	arviointiylilääkäri, Helsingin ja Uudenmaan sairaanhoitopiiri

Käyttäessäsi tätä raporttia lähteenä, merkitse viitteeksi: Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryyänen O-P, Roine RP. Laatu- ja vaikuttavuusarvioinnissa: järjestelmällinen kirjallisuuskatsaus. *Finohtan raportti 2006; 29.*

Tutkimus on julkaistu alkuperäisartikkelina (suppeampi versio ilman liitteitä) *Int J Technol Assess Health Care* -lehdessä 22:2(2006).

WORKING GROUP

Pirjo Räsänen	Researcher, Helsinki and Uusimaa Hospital Group and Finnish Office for Health Technology Assessment
Eija Roine	Medical Student
Harri Sintonen	Professor of Health Economics, Finnish Office for Health Technology Assessment and Helsinki University, Department of Public Health
Virpi Semberg-Konttinen	Officer for Audit Board, Helsinki and Uusimaa Hospital Group
Olli-Pekka Ryyänen	Professor, University of Kuopio, Department of Public Health and General Practice
Risto P. Roine	Chief Physician, Helsinki and Uusimaa Hospital Group

This report should be referenced as follows: Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryyänen O-P, Roine RP. Use of quality-adjusted life years for the estimation of effectiveness of health care: a systematic literature review. *Finohta's report 2006; 29.*

This research has also been published as an original article (limited version without appendices) in the *Int J Technol Assess Health Care* 22:2(2006).

© Kirjoittajat ja Stakes

Authors and National Research and Development Centre for Welfare and Health Stakes

Taitto/Layout: Terhi Ilonen

Kannen suunnittelu/Cover design: Harri Heikkilä

Kannen kuva/Photo: Ingram Publishing

ISBN 951-33-1850-8

ISSN 1239-6273

Gummerus Kirjapaino Oy
Vaajakoski 2006, Finland

SISÄLTÖ/CONTENTS

Esipuhe	4
Kiitokset	5
Tiivistelmä	6
Sammandrag	7
Taulukko- ja liiteluettelo	8
1 Johdanto	9
2 Menetelmät	11
2.1 Kirjallisuushaku	11
2.2 Julkaisujen valinta	11
2.3 Katsaukseen kelpuutettujen tutkimusten laatu	11
3 Tulokset	13
3.1 Löydetyt artikkelit	13
3.2 Tutkimusten luokittelu	14
3.3 Tutkimusten asetelma ja laatu	14
3.4 Raportoidut QALYt ja niiden kustannukset	14
3.5 Tutkimusten johtopäätökset	15
4 Pohdinta	17
5 Johtopäätökset	19
Foreword	22
Acknowledgements	23
Tables and appendices	24
Abstract	25
1 Introduction	26
2 Methods	28
2.1 Literature search	28
2.2 Selection of publications	28
2.3 Quality of included studies	28
3 Results	30
3.1 Retrieved articles	30
3.2 Study classification	31
3.3 Study design and quality	31
3.4 Reported number of QALYs and costs per QALY	31
3.5 Study conclusions	32
4 Discussion	34
5 Conclusions	36
Lähteet/References	37
Liite/Appendix 1. Literature search summary	38
Liite/Appendix 2. Details of selected studies	41
Artikkelit, joita siteerattu liitteessä 2/Articles cited in Appendix 2	80
Sidonnaisuudet/Conflict of interests	86

ESIPUHE

Teknologinen kehitys ja väestön ikääntyminen pakottavat terveydenhuollon palvelujärjestelmät nopeaan muutokseen. Tutkimusrintamalla panostus hoidon tuloksellisuuden arviointiin (evaluation) alkaa saada näkyvyyttä uusien tutkimus- ja hoitomenetelmien keksimisen (discovery) kustannuksella. Vastaavasti palvelutuotannon mittaamisessa aletaan lyhytjänteisen kaupallisen hyödyn tavoittelun ja markkinaohjautuvuuden asemesta työltä yhä useammin edellyttää nimenomaan hyvän elämän tuottamista. Koska edes vauraimmat maat eivät ole kyenneet lisäämään väestönsä elinajan odotetta kovin paljon terveydenhuoltoon panostamalla, elämän laadullinen parantaminen saa yhä suuremman sijan elämän pituuden tavoittelemisen rinnalla; Maailman Terveysjärjestön hieman kauhtuneen iskulauseen tavoin palvelujärjestelmien odotetaan siis tuovan myös elämää vuosin eikä vain vuosia elämään. – Näin potilaan elämän laadun painoarvo kasvaa siis siitä riippumatta, tarkastellaanko tilannetta tieteellisen tutkimustyön, palvelumarkkinoiden tai terveystalouden muutoksen näkökulmasta.

Tämän perusteella tuntuu oudolta, että potilaan subjektiivisen, holistisen elämänlaatukokemuksen mittaaminen poikki erilaisten hoitointerventioiden on vielä aika uusi ja ohutkin ilmiö. Tämän raportin laatijat ovat ottaneet tehtäväkseen kerätä raporttiin kaikki tieteellisen tutkimusasetelman vaatimukset täyttävät julkaisut, jotka koskevat hoidon vertailukelpoisin yksiköin mitattua vaikutusta potilaan itsensä kokemaan elämän laatuun. Sovellettu hakustrategia tuotti saaliiksi muutamia kymmeniä alkuperäisjulkaisuja, valtaosa Isosta-Britanniasta, loputkin käytännössä englantilaiselta kielialueelta.

Tutkimus antaa hyvän kuvan nykytilanteesta: Vain pieni osa nykyaikaisen terveydenhuollon hoitokäytännöistä on arvioitu koetun elämän laadun mittapuulla, suurin osa arvioinnin ”kartasta” on valkeata aluetta, vailla ensimmäistäkään piirtoa.

Kaikki palvelujärjestelmät joutuvat tuottamaan joitakin hoitoja toisista hoidoista tinkien. Terveydenhuollon laajalla kentällä ensisijaistaminen, priorisointi, voidaan tehdä muutamien perusturvapaalutusten jälkeen hoidon tuottaman ja sen kustannuksiin suhteutetun utiliteetin perusteella. Myös tämä pakottaa palvelujärjestelmän selvittämään nykyistä kattavammin, mistä hoidosta on hyötyä, minkälaista, kenelle, miten paljon, miten varmasti ja minkälaisin kustannuksin. Edelleen on tiedossa, että hoitoaiheiden vähittäinen muuttuminen voi ajan mittaan rapauttaa tehokkaaksikin osoittautuneen hoitomenetelmän suhteellista vaikuttavuutta.

Kumpikin näistä tekijöistä, priorisointitarve ja hoitoindeksaatioiden liukuminen, pakottavat ajan oloon rakentamaan kattavan tietokannan, jossa ennen-jälkeen-asetelmassa selvitetään eri hoitovaihtoehtojen tuottama yksilöllinen hyvä eli elämän kokemuksellisen laadun muutos. Aika ajoin tätä karttuvaa rutiinitietokantaa joudutaan kalibroimaan satunnaistetuin potilassarjain, aivan kuten tähän raporttiin kerätyissä alkuperäistutkimuksissa on tehty.

Järkevämpää tapaa tulevaisuuden kohtaamiseen on vaikea kuvitella.

Martti Kekomäki
Professori

KIITOKSET

Esitämme lämpimät kiitokset informaatikko Merja Jauhiaiselle Työterveyslaitokselta ja erikoislääkäri, tutkija Iris Pasternackille Stakes/Finohtasta.

Tutkimus sai rahallista tukea Stakesin Terveystieteiden tutkimuskeskuksen menetelmien arviointiyksiköltä, Finohtalta.

TIIVISTELMÄ

Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryyänen O-P, Roine RP.

Laatupainotetut elinvuodet terveydenhuollon vaikuttavuuden arvioinnissa: järjestelmällinen kirjallisuuskatsaus.

Finohtan raportti 29/2006. Terveydenhuollon menetelmien arviointiyksikkö (Finohta)/Stakes. Helsinki 2006. ISBN 951-33-1850-8. ISSN 1239-6273.

Tavoite. Tavoitteena oli laatia järjestelmällinen kirjallisuuskatsaus julkaistuista tutkimuksista, joissa on käytetty laatupainotettuja elinvuosia (QALYs, Quality Adjusted Life Years) siten, että ne perustuvat potilaan terveyteen liittyvän elämänlaadun (HRQoL, Health-Related Quality of Life) todelliseen mittaamiseen sekä määrittellä, mitä HRQoL-mittareita laatupainotettujen elinvuosien laskennassa on käytetty. Lisäksi luonnehditaan tutkimukset kuvaamalla niiden lääketieteen erikoisala, tutkittu toimenpide, tutkimuksen tulokset, laatu ja alkuperämaa, havaitut laatupainotetut lisäelinvuodet ja tulosten tulkinta kustannusvaikuttavuuden suhteen.

Menetelmät. Järjestelmällinen kirjallisuushaku toteutettiin seuraavista sähköisistä tietokannoista: Medline, Embase, CINAHL, SCI ja Cochrane-kirjasto. Aluksi kaksi raportin kirjoittajista seuroi artikkelit itsenäisesti niiden tiivistelmien perusteella. Tämän jälkeen kaksi kirjoittajaa arvioi valitut artikkelit kokonaisuudessaan ja päätti niiden mukaan ottamisesta.

Tulokset. Haun tuloksena löydettiin 3882 artikkelia, joista 624 kelpuutettiin lähempään tarkasteluun. Näistä 70 tutkimusta käytti laatupainotettujen elinvuosien määrittelyssä todellista ennen-jälkeen-asetelmaa ja hyödynsi validia HRQoL-mittaria. Yleisimmin käytetty oli EQ-5D (47,5 %). Muita mittareita olivat HUI (8,8 %), Rosser-Kind Index (6,3 %), QWB (6,3 %), SF-6D (5,0 %) ja 15D (2,5 %). Loput (23,8 %) käyttivät suoraa arviointimenetelmää: Time-Trade-Off (10,0 %), Standard Gamble (5,0 %), Visual analogue scale (5,0 %) tai Rating scale (3,8 %). Useimmin tutkitut lääketieteen erikoisalajat olivat ortopedia (15,5 %), keuhkosairaudet (12,7 %) ja kardiologia (9,9 %). Yhdeksänkymmentä prosenttia tutkimuksista tuli neljästä maasta: Isosta-Britanniasta, Yhdysvalloista, Kanadasta ja Alankomaista. Noin puolet oli menetelmällisesti korkealaatuisia satunnaistettuja kokeita. Alkuperäistutkimuksista 49 prosentissa tutkittu interventio katsottiin tutkimuksen tekijöiden arvioimana kustannusvaikuttavaksi; vain 13 prosentissa toimenpide määriteltiin ei-kustannusvaikuttavaksi.

Pohdinta. Laatupainotetut elinvuodet ovat tärkeä terveydenhuollon vaikuttavuuden mittari, mutta tutkimukset, joissa QALYt perustuvat potilaan terveyteen liittyvän elämänlaadun todelliseen mittaamiseen ennen-jälkeen-asetelmassa ovat vielä melko harvinaisia.

SAMMANDRAG

Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryyänänen O-P, Roine RP.

Kvalitetsjusterade levnadsår i utvärderingen av effektivitet i hälsovården: en systematisk litteraturoversikt.

Finohtas rapport 29/2006. Finska enheten för utvärdering av medicinsk metodik (Finohta)/Stakes. Helsingfors 2006. ISBN 951-33-1850-8. ISSN 1239-6273.

Syfte. Syftet var att göra upp en systematisk litteraturoversikt över publicerade undersökningar där kvalitetsjusterade levnadsår (QALY, Quality Adjusted Life Years) använts på ett sådant sätt att de baserar sig på verkliga mätningar av patienters hälsorelaterade livskvalitet (HRQoL, Health-Related Quality of Life), och definiera vilka HRQoL-instrument som använts vid beräkandet av kvalitetsjusterade levnadsår. Dessutom ges en karakteristik av undersökningarna genom att man beskriver den medicinska specialiteten, den åtgärd som undersökts, undersökningens resultat, kvalitet och ursprungsland, de observerade kvalitetsjusterade levnadsåren och tolkningen av resultaten avseende kostnads-effektiviteten.

Metoder. Den systematiska litteratursökningen genomfördes i följande elektroniska databaser: Medline, Embase, CINAHL, SCI och Cochrane-biblioteket. Först sällade två av rapportförfattarna artiklarna självständigt på basis av sammandragen. Därefter bedömde två författare de valda artiklarna i sin helhet och fattade beslut om vilka som skulle ingå.

Resultat. Resultatet av sökningen var 3 882 artiklar av vilka 624 godkändes för närmare granskning. Av dessa använde 70 undersökningar ett verkligt före-efter-upplägg vid definitionen av kvalitetsjusterade levnadsår och utnyttjade valida HRQoL-instrument. Det vanligaste var EQ-5D (47,5 %). Andra instrument var HUI (8,8 %), Rosser-Kind Index (6,3 %), QWB (6,3 %), SF-6D (5,0 %) och 15D (2,5 %). Övriga (23,8 %) använde en direkt utvärderingsmetod: Time-Trade-Off (10,0 %), Standard Gamble (5,0 %), Visual Analogue Scale (5,0 %) eller Rating Scale (3,8 %). De medicinska specialiteter som undersöktes var ortopedi (15,5 %), lungsjukdomar (12,7 %) och kardiologi (9,9 %). Nittio procent av undersökningarna kom från fyra länder: Storbritannien, USA, Kanada och Nederländerna. Ungefär hälften var metodiskt högklassiga randomiserade studier. I 49 procent av originalstudierna bedömdes den undersökta interventionen vara kostnadseffektiv, i endast 13 procent konstaterades att åtgärden inte hade kostnadseffekter.

Diskussion. Trots att de kvalitetsjusterade levnadsåren är en viktig mätare av hälsovårdens effekt är undersökningar där QALY baserar sig på verkliga mätningar av patienters livskvalitet med anknytning till hälsa i ett före-efter-upplägg fortfarande rätt sällsynta.

TAULUKKO- JA LIITELUETTELO

Taulukot

1.	Kriteerit, joiden perusteella artikkelit jaettiin viiteen eri ryhmään	13
2.	Tutkimuksen toteutuksen luokittelu (soveltaen lähdettä: Hailey ym. 2004)	15
3.	Tutkimuksessa käytetty HRQoL-mittari (joissakin tutkimuksissa useita)	16
4.	Raporttiin sisällytettyjen tutkimusten lääketieteelliset erikoisalajat	16
5.	Tutkimusten alkuperämaat	16

Liitteet

1.	Kirjallisuushaun yhteenveto	38
2.	Valittujen tutkimusten analyysi	41

JOHDANTO

Terveysthuoltoon panostetaan perinteisesti voimavaroja ilman yksityiskohdasta näyttöä sen tuottamista terveyshyödyistä. Voimavarojen ollessa rajalliset ne tulisi jakaa mahdollisimman tehokkaasti, mutta ilman vertailevaa tietoa vaikuttavuudesta päätöksenteko on usein epävakaa pohjalla. Erityisesti puutetta on tiedosta, joka mahdollistaa eri toimenpiteiden vaikuttavuuden vertailun yli lääketieteen eri erikoisalojen. Tämä johtuu siitä, että useimmat vertailevat tutkimukset ovat käyttäneet vaikuttavuuden arvioinnissa vain sairausspesifisiä mittareita.

Jos voimavarojen kohdentamisesta päätetään ainoastaan kliinisten tulosten perusteella, saattavat ne jakautua yhteiskunnallisen hyvinvoinnin kannalta epä-tarkoituksenmukaisesti. Eri vaihtoehtoja punnittaessa päätöksenteossa tulisi mahdollisten saavutettujen hyötyjen lisäksi ottaa huomioon myös menetetyt mahdollisuudet. Voimavarojen kohdentamisen pitäisi ihanteellisimmillaan tuottaa yhteiskunnalle mahdollisimman paljon hyötyä, mutta erityisesti terveydenhuollossa voimavarojen kohdentamista koskevaan päätöksentekoon liittyy usein hyvinkin suuri epävarmuus. Kun tavoitteena on laadukas ja kustannustietoinen terveydenhuolto, tulisi erilaisten interventoiden kustannusvaikuttavuuden luotettavan mittaamisen olla avainasemassa.

Viime vuosina on aikaisempaa enemmän painotettu, että elämän pituuden lisäksi myös sen laadulla on merkitystä. Tästä syystä on kehitetty geneerisiä mittaamenetelmiä, joilla hoidon tulosten arvioinnissa voidaan ottaa huomioon myös potilaan omat arvostukset. Mittausten vertailukelpoisuuden takaamiseksi, terveys-talouden tutkijat ovat kehittäneet terveyteen liittyvän elämänlaadun (HRQoL) käsitteen. HRQoL on yksilöllisen hyvinvoinnin indikaattori; mittari, jolla voidaan arvioida eri hoitomuotojen tuottamia terveyshyötyjä. HRQoL kuvaa jonkin sairauden vaikutusta potilaan elämänlaatuun sekä kliinisen toimenpiteen vaikuttavuutta terveyteen ja yleiseen hyvinvointiin (1). Sairauden ja sen hoidon lisäksi terveyteen liittyvään elämänlaatuun vaikuttavat myös kyseisen yksilön yleiset olosuhteet, muut mahdolliset terveysongelmat, sairauden kokeminen, hänen elämänvaiheensa, sekä tehtävänsä ja tavoitteensa.

Käytössä on kahdenlaisia elämänlaatumittareita: geneerisiä ja sairausspesifisiä mittareita. Sairauspesifejä mittareita käytetään, kun halutaan tietoa tietyn sairauden merkittävimmistä vaikutuksista. Siten ne eivät sovellu eri sairauksien hoitotulosten väliseen vertailuun. Niiden pääasiallinen tarkoitus onkin tukea päätöksentekoa potilashoidossa; ne ovat yleensä tiettyjen rajattujen sairauksien herkkiä mittareita. Hyvinä esimerkkeinä sairausspesifeistä mittareista voidaan mainita mm. Knee Society Score (KSS), jolla arvioidaan polvi-ongelman potilaan kipua ja liikkuvuutta sekä Harris Hip Score (HHS), joka on suunniteltu lonkkaoireiden arviointiin (2–3).

Geneerisiä mittareita voidaan käyttää hyvin erilaisille potilasryhmille. Menetelmällisesti ne luokitellaan profiili- ja/tai yhden indeksiluvun tuottaviksi mittareiksi. Profiilimittarit, esimerkiksi yleisesti käytössä oleva SF-36 -instrumentti, kuvaavat

terveydentilaa erilaisten fyysisten ja henkisten muuttujien kautta, joita voivat olla elinvoima, tunnetila, ruumiillinen kipu, yleinen terveydentila, sosiaalinen toiminta jne. Yhden indeksiluvun tuottavat mittarit käyttävät 0–1-asteikkoa (jotkut mittarit voivat tuottaa myös negatiivisia indeksilukuja). Tämä on edellytys sille, että niitä voidaan käyttää laatupainotettujen elinvuosien (QALY) laskemisessa ja terveydenhuollon menetelmien kustannusvaikuttavuuden vertailussa. HRQoL-instrumentin valinnassa on kiinnitettävä erityistä huomiota sen empiirisiin, teoreettisiin ja teknisiin ominaisuuksiin, joita ovat esimerkiksi mittarin validiteetti, luotettavuus, herkkyys, käyttökelpoisuus ja tulkittavuus (1,4,5). Geneerisistä, yhden indeksiluvun tuottavista mittareista voidaan mainita EQ-5D (EuroQol), SF-6D (kehitetty RAND-36/SF-36 -mittarin pohjalta), HUI (Health Utilities Index Mark II/Mark III), AQoL (Assessment of Quality of Life) ja 15D (1,6).

QALYn avulla terveydenhoidon vaikuttavuus voidaan ilmaista mittarilla, joka ottaa huomioon sekä elämän pituuden että laadun muutoksen. Viime vuosina QALY on saavuttanut yleisesti tunnustetun aseman terveydenhuollon vaikuttavuuden mittarina. Tämä näkyy esimerkiksi siten, että Britanniassa kansallisia hoitosuosituksia antava National Institute of Health and Clinical Excellence (NICE) käyttää QALYja terveysvaikutusten tärkeimpinä mittareina (7). Toinen osoitus laatupainotettujen lisäelinvuosien enenevästä käytöstä vaikuttavuuden mittarina on Medline-tietokannasta hakusanalla "QALY" löytyvien kirjallisuusviitteiden määrä; tällä vuosikymmenellä niiden määrä on lisääntynyt vuosittain keskimäärin kymmenellä prosentilla.

QALYja vaikuttavuuden arvioinnissa käyttävät tutkimukset perustuvat usein taloudelliseen mallinnukseen, jossa elämänlaatumittarit on koottu monista eri lähteistä tai ne perustuvat terveydenhuollon ammattilaisten arvioihin tiettyihin sairauksiin liittyvästä elämänlaadusta. Tällaiset arviot voivat kuitenkin olla harhaanjohtavia, koska ne edustavat palvelun tuottajan, eivät potilaan omia näkemyksiä. Siksi on tärkeää, että QALY-laskelmat perustuvat potilaan elämänlaadun todelliseen mittaamiseen. Tähän tarkoitukseen voidaan käyttää esimerkiksi jotakin valmista monimuuttujaista geneeristä elämänlaatumittaria ja/tai suoraa menetelmää, jossa potilas itse arvioi ja arvottaa omaa terveydentilaansa.

Tämän järjestelmällisen kirjallisuuskatsauksen tavoitteena oli löytää tutkimukset, jotka ovat käyttäneet potilaan itsensä arvioimaa elämänlaatua QALY-laskelmien perustana sekä luonnehtia tutkimuksia kuvaamalla niiden lääketieteen erikoisala, tutkittu toimenpide, tutkimuksen tulokset, käytössä ollut elämänlaatuinstrumentti, tutkimuksen laatu ja alkuperämaa, tutkimuksessa todetut laatupainotetut lisäelinvuodet ja tulosten tulkinta kustannusvaikuttavuuden kannalta.

2.1 Kirjallisuushaku

Kirjallisuushaku toteutettiin ilman kielirajoitusta seuraavista sähköisistä tietokannoista: Medline (1966–kesäkuu 2004), Embase (1966–kesäkuu 2004), CINAHL (1982–kesäkuu 2004) ja Science Citation Index (1982–kesäkuu 2004) sekä Cochrane-kirjastosta (numero 2, 2004) käyttämällä liitteessä A kuvattuja hakukriteerejä. Lisäksi joitakin artikkeleja löydettiin käymällä läpi katsauksen valittujen tutkimusten lähdeluettelot, etsimällä Medlinesta käyttämällä hakusanoina mukaan hyväksytyjen tutkimusten pääkirjoittajien nimiä sekä konsultoimalla taloudellisen arvioinnin asiantuntijoita. Lopuksi vertasimme haun tulosta Harvardin yliopiston ylläpitämään rekisteriin kustannusvaikuttavuuden tunnusluvuista (Cost Effectiveness Analysis Registry, <http://www.tufts-nemc.org/cearegistry/index.html>).

2.2 Julkaisujen valinta

Ensimmäinen seulonta tehtiin tutkimusten tiivistelmien perusteella. Vähintään kaksi katsauksen kirjoittajista luki tiivistelmät toisistaan riippumatta. Jatkotarkasteluun artikkelit valittiin tiivistelmien pohjalta katsauksen kirjoittajien yhteisen neuvottelun tuloksena. Jos tiivistelmä ei antanut tietoa tutkimuksesta tai tieto ei ollut riittävän yksityiskohtaista, koko artikkeli otettiin tarkempaan tarkasteluun.

Kaksi katsauksen kirjoittajista (PR, ER tai RR) luki valitut täyspitkät artikkelit itsenäisesti ja jaotteli ne ennalta määriteltujen kriteerien avulla viiteen luokkaan (taulukko 1). Jos kirjoittajat olivat eri mieltä siitä, mihin luokkaan jokin tutkimus tulisi sijoittaa, se annettiin kolmannen lukijan arvioitavaksi. Tämän jälkeen kaikki kolme kirjoittajaa neuvottelivat yhdessä kriteerit täyttävän ratkaisun löytämiseksi.

Katsaukseen kelpuutetun artikkelin tuli tieteellisesti pätevällä tavalla verrata terveyteen liittyvää elämänlaatua ennen–jälkeen-asetelmassa potilaiden oman arvioinnin pohjalta käyttämällä geneeristä, yhden indeksiluvun tuottavaa elämänlaatumittaria (15D, EQ-5D, SF-6D, HUI, AQoL, QWB, Rosser-Kind) tai suoraa arviointimenetelmää (TTO, SG, VAS tai RS) laatupainotettujen elinvuosien arvioimiseksi.

2.3 Katsaukseen kelpuutettujen tutkimusten laatu

Valittujen tutkimusten näytön vahvuutta arvioitiin aiemmin kuvattujen kriteerien mukaan (8–9) ottamalla huomioon sekä tutkimuksen asetelma että sen toteutus. Tutkimusasetelman luokittelussa käytettiin asteikkoa viidestä nollaan. Laajat satunnaistetut kokeet (vähintään 50 tutkittavaa kussakin ryhmässä) saivat viisi pistettä, pienet satunnaistetut kokeet kolme, etenevät ei-satunnaistetut kokeet kaksi, historialliset vertailututkimukset yhden ja ei-kontrolloidut tutkimukset nolla pistettä.

Tutkimusten toteutusta arvioitiin sillä, miten hyvin niissä oli otettu huomioon taulukossa 2 kuvatut viisi kohdealuetta. Jokaiselta alueelta oli mahdollista saada nollassa kahteen pistettä. Nolla pistettä merkitsi sitä, että olennainen tieto puuttui tai se ei ollut riittävän yksityiskohtaista, yksi piste kertoi, että tieto oli melko yksityiskohtaista, mutta siinä oli merkittäviä rajoituksia ja kaksi pistettä annettiin, jos tutkimuksen tuottama tieto oli tyydyttävää ilman merkittäviä rajoituksia. Siten tutkimuksen toteutuksesta oli mahdollista saada maksimissaan 10 pistettä.

3.1 Löydetyt artikkelit

Kirjallisuushaku tuotti tulokseksi 4878 julkaisua. Näistä 996 oli joko katsauksia, kirjeitä tai pääkirjoituksia ja suljettiin pois lähemmästä tarkastelusta, koska mukaan kelpuutettiin vain alkuperäistutkimukset. Tämän lisäksi mainitut 996 artikkelia sisälsivät myös julkaisuja, jotka käsittelivät ennaltaehkäisyä tai seulon-
taa, aihealueita, jotka jo etukäteen oli rajattu pois katsauksesta. Tarkasteltavaksi jäi siten kaikkiaan 3882 artikkelia, jotka mahdollisesti käyttivät QALYja tulosmittarina. Tiivistelmien arvioinnin perusteella lähempään tarkasteluun valittiin 624 artikkelia. Näistä 72 (joissa 70 erillistä tutkimusta) täytti valintakriteerit ja kelpuutettiin katsaukseen (liite B). 624 artikkelista 80:n kohdalla (13 %) kahden toisistaan riippumattoman arvioitsijan mielipiteet erosivat koskien sitä, perustuiko artikkeli selvästi identifioitavissa olevalla, validilla mittarilla (ryhmät A tai B taulukossa 1) saatuun elämänlaatumietoon. Näissä tapauksissa kyseisen artikkelin arvioi lisäksi kolmas kirjoittaja ja lopullinen päätös tehtiin yhteisessä kokouksessa. Näistä 80 artikkelista katsaukseen valittiin lopulta 18.

Kun vertasimme kirjallisuushaun tuloksia Cost Effectiveness Analysis Registry -tietokantaan, löysimme vielä 59 mahdollista lisäartikkelia. Niiden tarkempi tutkiminen osoitti kuitenkin, ettei yksikään täyttänyt katsauksemme sisäänottokriteereitä.

Taulukko 1. Kriteerit, joiden perusteella tarkastellut artikkelit jaettiin viiteen eri ryhmään.

-
- A** HRQoL mitattu geneerisellä, QALYn laskemisen mahdollistavalla mittarilla (15D, EQ-5D, SF-6D, HUI, AQoL, GWB, tai Rosser-Kind), tai suoralla määrittymenetelmällä (TTO, SG, VAS, tai RS) joko itse raportoitavassa tutkimuksessa tai selvästi identifioitavissa olevassa muussa tutkimuksessa, ja HRQoL määritetty sekä ennen interventiota että sen jälkeen.
 - B** HRQoL mitattu geneerisellä, QALYn laskemisen mahdollistavalla mittarilla (15D, EQ-5D, SF-6D, HUI, AQoL, GWB, tai Rosser-Kind), tai suoralla määrittymenetelmällä (TTO, SG, VAS, tai RS), mutta HRQoL-mittaus tehty vain joko ennen interventiota tai sen jälkeen.
 - C** HRQoL tieto huonosti määritellystä lähteestä tai määritetty mittarilla, joka ei mahdollista QALYn laskemista, vaikka artikkelissa raportoitaisiinkin erilaisten transformaatiomenetelmien avulla laskettuja QALYja.
 - D** HRQoL arvioitu (pääosin tai kokonaan) asiantuntijapaneelien tai vapaaehtoisten toimesta, tai kirjallisuuteen perustuen
 - E** Katsausartikkeli, joka ei raportoi alkuperäisiä HRQoL määrittymisen tuloksia tai riittämätön tai huonosti määritelty tutkimusasetelma.
-

Taulukossa käytetyt lyhenteet: 15D = 15D HRQoL mittari; AQoL = Assessment of Quality of Life; EQ-5D = EuroQol HRQoL mittari; HRQoL = Terveysteen liittyvä elämänlaatu; HUI = Health Utilities Index Mark II/ Mark III; QWB = Quality of Well-Being Scale; Rosser-Kind = Rosser-Kind HRQoL mittari; RS = Rating Scale; SF-6D = SF-6D HRQoL mittari; SG = Standard Gamble; TTO = Time Trade-off; VAS = Visual analogue Scale

3.2 Tutkimusten luokittelu

Mukaan valitut 70 tutkimusta ryhmiteltiin käytetyn elämänlaatumittarin (taulukko 3), 18 lääketieteen erikoisan (taulukko 4) ja julkaisun alkuperämaan mukaan (taulukko 5). Valituista artikkeleista 71 prosenttia oli julkaistu oman erikoisan tieteellisissä sarjoissa, 20 prosenttia oli julkaistu yleisissä lääketieteellisissä aikakauslehdissä ja kahdeksan prosenttia pääasiassa terveystaloustieteeseen, terveydenhuollon menetelmien arviointiin tai terveydenhuollon hallintoon keskittyvissä aikakauslehdissä. Yksi tutkimus oli julkaistu väitöskirjana. Artikkeleista 67 oli englanninkielisiä, yksi norjan-, yksi hollannin- ja yksi espanjankielinen.

Tutkimuksista 31 prosenttia käsitteli pääasiassa lääkehoitoa ja 26 prosenttia kirurgisia interventioita. Loput tutkimuksista käsitelivät erilaisia konservatiivisia hoitoja, kuntoutusta, lääketieteellistä kuvantamista ja sekundaarista ehkäisyä. Tutkimusinterventiot olivat moninaiset: elinsiirtokirurgiasta kylpyläterapiaan. Yleisimmin tutkitut interventiot olivat sepelvaltimotaudin hoito, lonkan tekonielleikkaus ja sisäkorvaistute, joista kustakin oli neljä tutkimusta.

Tutkimuksista 86 prosenttia sisälsi taloudellisen osion, yhdeksän tutkimusta raportoi vain elämänlaatu- ja QALY-tulokset.

Katsaukseen sisällytetyt tutkimukset on esitetty yksityiskohtaisesti englanninkielellä liitteessä 2, josta käy ilmi niiden kliininen erikoisala, interventio, tavoite, tutkimuspopulaatio, käytetty arviointimenetelmä, taloudellisen arvioinnin näkökulma, tutkimuksessa käytetyt kustannustiedot, elämänlaadun mittaustulokset, toimenpiteen tuottamien laatuainotettujen lisäelinvuosien määrä ja kustannukset, tutkimuksen laatu sekä menetelmälliset tai muut rajoitukset. Kuten aiemmin todettiin, kahta tutkimusta käsiteltiin useissa artikkeleissa; näissä tapauksissa niiden tulokset on yhdistetty.

3.3 Tutkimusten asetelma ja laatu

Tutkimusten laadun arviointi osoitti, että noin puolet artikkeleista perustui satunaistettuihin kontrolloituihin kokeisiin. Suurin osa muistakin tutkimuksista oli asetelmaltaan vertailevia. Kun tutkimusten toteuttamista arvioitiin 0–10 pisteen asteikoilla (taulukko 2), hyvän arvosanan (8–10 pistettä) sai 49 prosenttia, tyydyttävän (6–7 pistettä) 29 prosenttia ja välttävän (4–5 pistettä) 22 prosenttia tutkimuksista. Yhtään tutkimusta ei arvioitu laadultaan huonoksi (alle 4 pistettä). Neljä tutkimusta käytti taloudellista mallinnusta, kolme Markovin mallia ja yksi perustui päätöspuuteoriaan.

3.4 Raportoidut QALYt ja niiden kustannukset

Laatupainotettujen lisäelinvuosien määrä vaihteli tutkimuksissa suuresti: negatiivisesta kahdeksaan riippuen tutkitusta interventiosta sekä osittain siitä, kuinka pitkälle aikavälille saavutetut QALYt oli mallinnettu. Myös kustannukset QALY

kohti vaihtelivat erittäin paljon: alle tuhannesta eurosta aina yli miljoonaan euroon asti.

3.5 Tutkimusten johtopäätökset

QALY-tulosten raportoinnin lisäksi useimmat tutkimukset käsittelevät pohdintaosiossa QALYjen kustannuksia yhteiskunnallisen hyväksyttävyyden kannalta. Alkuperäistutkimuksista 49 prosentissa tutkitun intervention katsottiin olevan kustannusvaikuttava; vain 13 prosentissa interventio määriteltiin ei-kustannusvaikuttavaksi.

Taulukko 2: Tutkimuksen toteutuksen luokittelu (soveltaen lähdettä: Hailey ym. 2004).

	Tarkastelun kohde	Huomioitavat seikat
1	Potilasvalinta	Satunnaistamisessa/valinnassa käytetyt menetelmät; interventio- ja vertailuryhmän samankaltaisuus; ennen tutkimuksen alkamista pois jääneiden lukumäärä
2	Intervention kuvaus/määrittely	Interventio ja vertailuryhmän riittävä kuvaus
3	Tutkimuksen määrittely ja analyysissä käytetyt menetelmät	Otoskoko; käytetyt tilastolliset menetelmät; lopputulosmuuttujien selvä määrittely
4	Käytettävissä oleva potilasmäärä	Seurannan pituus, poisjääneiden määrä, tutkimusprotokollan noudattaminen
5	Lopputulosmuuttujien raportointi	Raportoinnin riittävyys ja selkeys; puuttuvat tulokset; tilastollinen yhteenveto; vastaavatko johtopäätökset tuloksia

Jokainen tarkastelun kohde pisteytetään seuraavien havaintojen perusteella asteikolla 0–2 (maksimipistemäärä yhteensä 10):

0 = Artikkelista puuttuu merkityksellistä informaatiota tai sitä on vain minimaalisesti

1 = Yksityiskohdat on kuvattu kohtuullisen tarkasti, mutta kuvauksessa on joitakin tärkeitä puutteita

2 = Informaatiota on riittävästi, ei merkittäviä rajoituksia

Taulukko 3. Tutkimuksessa käytetty HRQoL-mittari (joissakin tutkimuksessa useita).

Instrumentti	N	% kaikista	%
HRQoL-mittarit			76
15D	2	2,5	
EQ-5D (EuroQol)	37	46,8	
HUI (Health Utilities Index Mark II/Mark III)	7	8,9	
QWB (Quality-of-Well-Being Scale)	5	6,3	
Rosser-Kind	5	6,3	
SF-6D	4	5,1	
Suora arviointi			24
SG (Standard Gamble)	4	5,1	
TTO (Time-trade-off)	7	10,9	
Rating Scale	2	3,8	
VAS	4	5,1	

Taulukko 4: Raporttiin sisällytettyjen tutkimusten lääketieteelliset erikoisalalat.

Erikoisala	N (%)
Ortopedia	11 (15,7)
Keuhkosairaudet	9 (12,9)
Kardiologia	7 (10,0)
Neurologia	6 (8,6)
Reumatologia	6 (8,6)
Korva-, nenä- ja kurkkutaudit	5 (7,1)
Elinsiirtokirurgia	5 (7,1)
Psykiatria	4 (5,7)
Syöpätaudit	4 (5,7)
Gynekologia	2 (2,9)
Tehohoito	2 (2,9)
Urologia	2 (2,9)
Endokrinologia	2 (2,9)
Infektiotaudit	1 (1,4)
Munuaistaudit	1 (1,4)
Hammaskirurgia	1 (1,4)
Gastroenterologia	1 (1,4)
Yleiskirurgia	1 (1,4)

Taulukko 5: Tutkimusten alkuperämaat.

Maa	N (%)
Iso-Britannia	23 (32,9)
Alankomaat	18 (25,7)
Kanada	11 (15,7)
Yhdysvallat	11 (15,7)
Espanja	2 (2,9)
Saksa	1 (1,4)
Hongkong	1 (1,4)
Norja	1 (1,4)
Ruotsi	1 (1,4)
Monikansallinen	1 (1,4)

Vaikka laatupainotettuja elinvuosia pidetään tärkeänä vaikuttavuuden mittarina terveydenhuollossa, vain pieni osa tutkimuksista perustuu potilaan terveyteen liittyvän elämänlaadun todelliseen mittaamiseen. Useissa tätä katsausta varten analysoiduissa tutkimuksissa elämänlaatudiedot oli saatu epämääräisesti määritellyistä lähteistä tai ne perustuivat terveydenhuollon ammattilaisten omiin arvioihin. Vaikka he varmasti ovatkin tietoisia sairauden kliinisestä luonteesta ja siitä taakasta, jonka se potilaalle aiheuttaa, on hyvin epätodennäköistä että he, ilman omakohtaista sairauden kokemista, pystyisivät tarkasti arvioimaan potilaan terveyteen liittyvää elämänlaatua. Tähän viittaa mm. hiljattain julkaistu tutkimus, jossa todettiin eturauhassyöpöpotilaiden ja kliinikoiden käsitysten hoidon hyödyistä korreloivan huonosti ja ilman tilastollista merkitsevyyttä (10). Tätä havaintoa tukevat myös jotkut muut aikaisemmat tutkimukset (11–13). Onkin oletettavaa, että tutkimukset, jotka perustuvat todelliseen, potilailta kysyttyyn tietoon ovat resurssien jakoa pohtivalle päättäjälle paljon hyödyllisempiä.

Toisaalta löysimme myös useita tutkimuksia, joissa elämänlaatu oli arvioitu kellollisessa ennen–jälkeen-asetelmassa, mutta QALY-termiä ei raporteista löytynyt lainkaan. Suurimmassa osassa näistä tutkimuksista QALYjen laskeminen olisi ollut mahdollista, mutta koska näin ei ollut tehty, emme voineet sisällyttää näitä tutkimuksia katsauksemme. Hakustrategiamme mukaisesti etsimme nimenomaan tutkimuksia, jotka raportoivat lopputuloksenaan laatupainotettuja elinvuosia. Niinpä hakumme ei välttämättä olisi ollut kattava tällaisten muiden artikkelien suhteen. On mahdollista, että tutkimukset, joissa interventiolla ei ole lainkaan tai vain vähäisiä vaikutuksia terveyteen liittyvään elämänlaatuun, saattavat jättää QALYt raportoimatta. Jos tulokset ovat positiivisia, QALYt raportoidaan ehkä herkemmin. Siten katsauksemme sisällytetyt tutkimukset saattavat antaa jonkin verran harhaisen ja ylioptimistisen kuvan lääketieteellisillä interventioilla saavutettavista laatupainotetuista lisäelinvuosista.

Lähestymistapamme eroaa aiemmin tehdyistä kirjallisuuskatsauksista siinä, että etsimme ainoastaan tutkimuksia, joissa raportoidut QALYt on laskettu validilla elämänlaatumittarilla ennen–jälkeen-asetelmassa. Tästä syystä vertailu aiempiin, muita hyväksymiskriteereitä käyttäneisiin katsauksiin ei ole välttämättä kovin hedelmällistä. On totta, ettei aina ole mahdollista saada asianmukaista elämänlaatudietoa suoraan potilaalta (esim. lapsilta, dementoituneilta jne.). Näissä tapauksissa on käytettävä ulkopuolista arvioijaa tai muita menetelmiä. Tällaisten tutkimusten pois sulkeminen katsauksestamme ei merkitse sitä, etteivät niiden tulokset olisi arvokkaita. Se ei myöskään merkitse sitä, että toisentyyppisten tutkimustulosten järjestelmälliset kuvaukset, kuten kattava Cost Effectiveness Analysis Registry (<http://www.tufts-nemc.org/cearegistry/index.html>), olisivat vähemmän tärkeitä kuin meidän lähestymistapamme. Usein potilaat ovat kuitenkin kykeneviä itse arvioimaan omaan terveyteensä liittyvää elämänlaatua. Mielestämme näissä tapauksissa ulkopuoliseen arviointiin perustuva tieto saattaa johtaa harhaan. Siksi halusimme kartoittaa ja kuvata olemassa olevan tieteellisen

kirjallisuuden, jossa raportoidut QALYt perustuvat todelliseen potilailta saatuun tietoon.

Tutkimusten laatu vaihteli pääosin kohtalaisesta hyvään ja noin puolet niistä perustui satunnaistettuihin kontrolloituihin kokeisiin. Siten useimpien tutkimusten tuloksia voidaan pitää luotettavina ja niitä voidaan suoraan hyödyntää päätöksenteossa. Aiemmissä raporteissa on todettu kustannus-utiliteettianalyysien laadun edelleen vaihtelevan (14–15). Vaikka katsauksemme pääpaino ei ollutkaan tutkimusten laadussa eikä lähestymistapamme laadunarviointiin ollut yhtä tarkkaa kuin Gerardilla (14–15), tuloksemme saattavat viitata siihen, että vuosien saatossa kustannus-utiliteettianalyysit ja niiden raportointi ovat hiljalleen muuttuneet laadukkaimmiksi. Tämä johtopäätös on yhtäpitävä Neumannin ym. äskettäisen tutkimuksen tulosten kanssa, jossa vertailtiin vuosina 1998–2001 julkaistuja artikkeleita 1976–1997 julkaistuihin ja todettiin, että niiden laatu oli parantunut melkein kaikissa arvioituissa kategorioissa (16).

Katsaukseen sisällytetyissä tutkimuksissa käytetyin elämänlaatumittari oli selvästi EQ-5D. Se on helppokäyttöinen, mutta sillä on omat rajoituksensa: se määrittelee esimerkiksi vain 243 terveydentilaa. EQ-5D on kehitetty useiden maiden terveystaloustieteilijöiden yhteistyönä, jossa on ollut vahva edustus erityisesti Ison-Britanniasta ja Alankomaista, joka on todennäköisesti edistänyt EQ-5D -mittarin käyttöä näissä maissa. Tämän katsauksen tutkimuksista suurin osa tuli juuri näistä maista.

Useampaa kuin yhtä geneeristä elämänlaatumittaria käyttäneissä tutkimuksissa tulokseksi saatujen QALYjen määrä vaihteli melko paljon riippuen käytössä olleesta mittarista. Jos tutkimuksissa on käytetty eri mittareita, rajoittaa se eri interventioiden vaikuttavuuden vertailua QALYjen perusteella. Toinen tekijä, joka häiritsee tulosten vertailua, on vaihtelu arvioitavien ajanjaksojen pituudessa. Joissakin tutkimuksissa QALYjen laskennassa käytetty aikaväli oli vuotta lyhyempi, toisissa taas useita vuosia tai se ulottui kuolemaan asti. Samanaikaisesti tulosten vertailua hankaloittaa se, että kustannuksia ja hyötyjä on arvioitu eri lähtökohdista (esimerkiksi yhteiskunnallisesta tai palvelun tarjoajan näkökulmasta). Jos halutaan tehdä kustannus-utiliteettianalyysien tuloksista käyttökelpoisempia, tarvitaan yhteistä ja yleisesti hyväksyttyä metodologiaa.

Melkein puolta tutkituista toimenpiteistä voitiin pitää yhteiskunnallisesti hyväksyttävänä kustannukset/QALY -arviolla. Yhteiskunnallinen QALYja koskeva maksuhalukkuuskynnys kuitenkin vaihteli viitaten siihen, ettei tällä hetkellä ole yleisesti hyväksyttyä tasoa, jota voitaisiin käyttää voimavarojen kohdentamista koskevassa päätöksenteossa. Tässä suhteessa tuloksemme ovat yhteneviä äskettäin julkaistun tutkimustuloksen kanssa, jossa on käsitelty yhteiskunnan maksuhalukkuutta laatupainotetusta elinvuodesta. (17)

5

JOHTOPÄÄTÖKSET

Laatupainotettuja lisäelinvuosia pidetään tärkeänä terveydenhuollon vaikuttavuuden mittarina, mutta tutkimukset, joissa QALYjen raportointi perustuu potilaan terveyteen liittyvän elämänlaadun todelliseen mittaamiseen ovat vielä melko harvinaisia. Tällaisia tutkimuksia kuitenkin tarvitaan varmistamaan, että terveydenhuollon voimavarojen kohdentaminen perustuu tieteelliseen näyttöön eri interventioiden hyödyistä ja niiden mahdollisuudesta tuottaa hyvinvointia.

Pirjo Räsänen, Eija Roine, Harri Sintonen, Virpi Semberg-Konttinen,
Olli-Pekka Ryyänen, Risto P. Roine

Use of Quality-Adjusted Life Years for the Estimation of Effectiveness of Health Care: A Systematic Literature Review

Finohta's report

29/2006



FOREWORD

The advance of medical technologies and ageing demographics prompt rapid changes in post-industrialised countries. In research, evaluation gains importance even at the cost of new discoveries. Correspondingly, the health services market seeks measurable health benefits, is ready to pay for effective care instead of mere service production. At the level of health politics, yesterday's slogan "not just years to life but also life to years" seems to grow in weight. Thus, whether we view the change from the scope of research, service market or national health policies, the importance of the change in the experienced quality of life across the intervention is increasing.

Against this background it is somewhat surprising that the tradition to measure the subjective, holistic, quality-of-life -related patient experience across different care interventions is a rather new, even fragile phenomenon. The authors of this report have collected all those original publications, which deploy global measures of health-related quality of life and fulfil the strictest criteria of randomised, prospective controlled trials. This search strategy yielded less than one hundred publications, most of them coming from the U.K. and even the rest usually from the English-speaking world.

The present analysis provides a good picture of the present situation: Only a small proportion of the modern care entities have been scrutinised using perceived quality of life as a measure. Most part of the map is just white, lacking any markings thus far.

All service systems are pressed to set care priorities, in one way or another. After some basic issues relating to human rights have been covered, these priorities will base on the utility gains provided by care, as related to the costs incurred. This fact also makes it mandatory for the health care system to monitor far more widely than is currently done, which care benefits whom, how much, how probably, and what costs are tied up by such a care. Furthermore, a growing body of evidence seems to point out that medical indications seem to creep gradually over time, which may seriously erode the measured utility-cost ratio of almost any treatment modality.

Both of these factors, the need to set priorities and the creeping of clinical indications, should lead to the construction of massive clinical databases, which, using a simple before-after -setting quantify the experienced quality-of-life change across the medical intervention. From time to time, this routine database is then calibrated using randomised, prospective patient series, as has been in the original publications collected into this document.

I cannot envision a more rational way to meet the future.

Martti Kekomäki
Professor

ACKNOWLEDGEMENTS

The authors wish to thank Merja Jauhiainen, Information Officer at the Institute of Occupational Health Institute and Iris Pasternack, MD, Occupational health physician, researcher at the National Research and Development Centre for Welfare and Health (STAKES) and the Finnish Office for Health Care Technology Assessment (FinOHTA).

The study was financially supported by the Finnish Office for Health Technology Assessment.

TABLES AND APPENDICES

Tables

1.	Criteria for classification of reviewed full-text articles into one of five categories	30
2.	Classification of study performance according to Hailey et al. 2004 (b)	32
3.	HRQoL instrument used in the study (in some studies several instruments were used)	33
4.	Clinical specialities of the included studies	33
5.	Country of origin of the studies	33

Appendices

1.	Literature search summary	38
2.	Details of selected studies	41

ABSTRACT

Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryyänen O-P, Roine RP.

Use of quality-adjusted life years for the estimation of effectiveness of health care: a systematic literature review.

Finohta's report 29/2006. Finnish Office for Health Technology Assessment (Finohta)/National Research and Development Centre for Welfare and Health STAKES. Helsinki, Finland 2006. ISBN 951-33-1850-8. ISSN 1239-6273.

Objective. To identify, in a systematic literature review, published studies having used quality-adjusted life years (QALYs) based on actual measurements of patients' health-related quality of life (HRQoL), and to determine which HRQoL instruments have been used to calculate QALYs. Furthermore, to characterise studies with regard to medical speciality, intervention studied, results obtained, quality, country of origin, QALY gain observed, and interpretation of results regarding cost-effectiveness.

Methods. Systematic search of the literature using the Medline, Embase, CINAHL, SCI and Cochrane Library electronic databases. Initial screening of identified articles based on abstracts read independently by two of the authors, full-text articles again evaluated by two authors, who made the final decision on which articles should be included.

Results. The search identified 3882 articles, 624 were obtained for closer review. Of the reviewed full-text articles, 70 reported QALYs based on actual before-after measurements using a valid HRQoL instrument. The most frequently used instrument was EQ-5D (47.5 %). Other instruments used were HUI (8.8 %), Rosser-Kind Index (6.3 %), QWB (6.3 %), SF-6D (5.0 %), and 15D (2.5 %). The rest (23.8 %) employed a direct valuation method: Time-Trade-Off (10.0 %), Standard Gamble (5.0 %), Visual analogue scale (5.0 %), or Rating scale (3.8 %). The most frequently studied medical specialities were orthopaedics (15.5 %), pulmonary diseases (12.7 %), and cardiology (9.9 %). 90 per cent of the studies came from four countries: United Kingdom, USA, Canada, the Netherlands. Approximately half of the papers were methodologically high quality randomised trials. Forty-nine per cent of the studied interventions were viewed by the authors of the original studies as being cost-effective; only 13 per cent of interventions were deemed not to be cost-effective.

Conclusions. Although QALYs are considered an important measure of effectiveness of health care, the number of studies in which QALYs are based on actual measurements of patients' HRQoL is still fairly limited.

Keywords

Cost-effectiveness, health-related quality of life, quality-adjusted life years

INTRODUCTION

Investments into health care have traditionally been made without detailed information on the health gains produced. As resources are scarce, they should be allocated in the most cost-effective way, but without comparative effectiveness data, decision making is often on a shaky ground. Especially data allowing the comparison of the effectiveness of various interventions across different medical specialities have been scarce due to the fact that most comparative studies have used disease-specific outcome measures.

Allocation decisions based on clinical results only may lead to inappropriate distribution of resources regarding societal welfare. Thus, when considering various alternatives, one should, in addition to the expected gains, also take into account the lost opportunities that inevitably follow an investment decision. Resource allocation should, under optimal conditions, generate maximal benefits for the society, but especially in health care allocation decisions are often combined with significant uncertainty. To be able to reliably measure the cost-effectiveness of various interventions is thus one of the key targets in the pursuit for good quality, cost-conscious health care.

During recent years it has been acknowledged that in addition to the length of life also its quality is of importance. This has resulted in attempts to develop new, generic methods for the estimation of treatment results that also take into account patient preferences. To solve the problem of comparability of measurements, health economists have introduced the concept of health-related quality of life (HRQoL) as an indicator of individual well-being and as a potential yardstick for the estimation of health gains produced by treatments. HRQoL can be used to describe the effects of an illness on the quality of life and the effect of clinical interventions on health and general well-being (1). In addition to the disease and its treatment, HRQoL is affected by the general condition of the individual in question, other health problems and sickness experiences he may have, his phase of life as well as the tasks and goals he has.

Two kinds of HRQoL instruments exist, generic and disease-specific ones. Disease-specific instruments are used for studying the most important effects of a given disease. They are thus not suited for comparison of treatment results across various diseases. Their main purpose is to assist clinical decision making and they are usually sensitive in measuring results of specific treatments. Good examples of disease-specific instruments are for instance the Knee Society Score (KSS) evaluating pain and mobility in patients with knee problems, and the Harris Hip Score (HHS) designed for the assessment of symptoms of hip disorders (2–3).

The generic instruments can be used for diverse patient groups independent of the underlying disease or disability. Generic instruments can be methodologically classified into profile and single index score measures. The former describe the health state from the stand point of various physical and emotional dimensions such as vitality, role emotional, bodily pain, general health, social function, etc. as in the widely used SF-36 instrument. The latter produce a single index score on a 0-1 scale (although some instruments produce also negative scores), which is a necessary requirement for the calculation of quality-adjusted life-years

(QALYs) used for commensurate appraisal of the cost-effectiveness of various health care interventions. When choosing a HRQoL instrument special attention needs to be paid to its empirical, theoretical, and technical characteristics such as validity, reliability, sensitivity, usability, and interpretability (1,4,5). Generic, single index score instruments include for instance the EQ-5D (EuroQol), the SF-6D (derived from RAND-36/SF-36), the HUI 3 (Health Utilities Index Mark II/Mark III), the AQoL (Assessment of Quality of Life) and the 15D (1, 6).

The QALY makes it possible to express the effectiveness of health care as a combination of a change both in the length and quality of life. During recent years the QALY has been recognised as the currently most important indicator of effectiveness of health care interventions. This is reflected for instance in the standpoint of the National Institute of Health and Clinical Excellence (NICE), providing national guidance on treatments and care for those using the NHS in England and Wales, that it uses the QALY as its principal measure of health outcome (7). The increasing utilisation of QALYs gained as a measure of effectiveness is also evidenced by the fact that the number of references found with the search term QALY in the Medline database has during this decade increased by approximately 10 per cent every year.

Many of the articles reporting QALYs as end-points, however, are based on economic modelling using HRQoL data obtained from many different sources or derived from health care professionals' estimates of the HRQoL associated with certain disease states. Such estimates, however, are likely to be biased as they represent the care providers' views, not those of patients. Consequently it is of importance that QALY calculations are based on actual measurements of patients' HRQoL by either multi-attribute (like the available generic HRQoL instruments) and/or direct (the patients' assessment and valuation of their own health status) measures. The aim of this systematic literature review was to identify articles having used patient-derived HRQoL as the basis for the QALY calculations and to characterise the studies with regard to medical speciality, intervention studied, results obtained, HRQoL instrument employed, quality, country of origin, QALY gain observed, and interpretation of results regarding cost-effectiveness.

2.1 Literature search

Computerized literature searches were performed, without any language restrictions, using the Medline (1966–June 2004), Embase (1966–June 2004), CINAHL (1982–June 2004), and Science Citation Index (1982–June 2004) databases and the Cochrane library (Issue 2, 2004) and the search strategies described in Appendix A. In addition some articles were identified by scanning reference lists of included articles, running a Medline search using the name of the principal author of each included article as the search term, and consulting experts in the field of economic evaluation. Finally, we also compared the results of our search with the listing of cost-effectiveness ratios published in the Cost Effectiveness Analysis Registry (<http://www.tufts-nemc.org/cearegistry/index.html>).

2.2 Selection of publications

Initial screening of the identified articles was based on their abstracts. All abstracts were read independently by at least two of the authors. Selection of relevant articles was based on the information obtained from the abstracts and was agreed upon in discussion between the authors. When an abstract did not give sufficiently precise information about the study or such information was not available at all, the full-text article was obtained for further review.

Full-text articles obtained for closer inspection were independently read by at least two of the authors (PR, ER, or RR) and placed in one of five categories according to predefined criteria (Table 1). If the two readers disagreed about the category the article belonged to, the article was read by a third person, and all three evaluators then discussed the article together to reach consensus using the criteria discussed in Table 1 and below.

Included were articles that in a scientifically valid manner compared HRQoL of patients in a before-after setting and in which HRQoL had been assessed by a generic HRQoL instrument recognized to produce a valid single index score for the calculation of QALYs (15D, EQ-5D, SF-6D, HUI, AQoL, QWB, Rosser-Kind), or in which HRQoL had been assessed by a direct valuation method (TTO, SG, VAS, or RS).

2.3 Quality of included studies

The strength of evidence given in selected papers was considered with regard to the study design used and study performance as described earlier (8–9). For study design, scores were assigned to five classifications. Large randomized controlled trials (RCTs), defined as those with at least 50 subjects in each arm, were given a score of 5. Small RCTs had a score of 3, prospective non-randomized studies 2, retrospective comparative studies 1, and non-controlled series 0.

For study performance, five areas of interest were considered, as shown in Table 2. When reviewing a study, each of these five areas was given a score of 0, 1 or 2. A score of 0 applied when relevant information was missing or given in only minimal detail; 1 indicated that reasonable detail was provided but there were some important limitations; and a score of 2 was allocated when information provided was satisfactory, with no significant limitations. Each study therefore had a possible maximum score of 10 for performance.

3.1 Retrieved articles

The literature search identified 4878 publications. However, 996 were either reviews, letters, or editorials and, as we were looking for original studies, not included for further review. Furthermore, the 996 excluded articles also included publications dealing with prevention or screening, topics which had been decided to be excluded from the review. Thus we were left with altogether 3882 articles potentially reporting QALYs as outcome measures. After screening of abstracts, 624 full-text articles were selected for closer inspection. Of them 72 (representing 70 separate studies) were deemed to fulfil the selection criteria and were included in the review (Appendix B). In 80 cases (13 % of the 624 full-text articles) the initial evaluation of the two independent reviewers differed regarding whether the article was based on clearly identifiable HRQoL data obtained with a valid instrument (groups A or B) or not. In those cases the article was also evaluated by a third person and the final decision was made in a consensus meeting. Of those 80 articles, 18 were finally deemed to merit inclusion in the review.

Comparison of our search result with the Cost Effectiveness Analysis Registry database identified 59 additional candidate articles, none of which however, when studied in more detail, turned out to fill the inclusion criteria of the review.

Table 1. Criteria for classification of reviewed full-text articles into one of five categories.

A	HRQoL measured with a generic instrument allowing calculation of QALYs (15D, EQ-5D, SF-6D, HUI, AQoL, GWB, or Rosser-Kind), or assessed by a direct valuation method (TTO, SG, VAS, or RS) either in the study reported, or in a clearly identifiable other study, and HRQoL was assessed both before and after the intervention.
B	HRQoL measured with a generic instrument allowing calculation of QALYs (15D, EQ-5D, SF-6D, HUI, AQoL, GWB, or Rosser-Kind), or assessed by a direct valuation method (TTO, SG, VAS, or RS), but HRQoL was assessed only before or after the intervention.
C	HRQoL data obtained from poorly defined sources or determined with an instrument not suitable for calculation of QALYs although the article reports QALYs using diverse transformation processes.
D	HRQoL estimated (mainly or entirely) by expert panels or based on literature.
E	Review article not reporting original assessment of HRQoL or inadequate or poorly defined research setting.

Abbreviations used in the table: 15D = 15D HRQoL instrument; AQoL = Assessment of Quality of Life; EQ-5D = EuroQol HRQoL instrument; HRQoL = Health-related Quality of life; HUI = Health Utilities Index (Mark II/Mark III); QWB = Quality of Well-Being Scale; Rosser-Kind = Rosser-Kind HRQoL instrument; RS = Rating Scale; SF-6D = SF-6D HRQoL instrument; SG = Standard Gamble; TTO = Time Trade-off; VAS = Visual analogue Scale

3.2 Study classification

The 70 selected publications were grouped by the HRQoL instrument employed in the study (Table 3), by the 18 medical specialities they represented (Table 4), and by the country of origin (Table 5). Of the included articles 71 per cent had been published in speciality journals, 20 per cent in general medical journals, and 8 per cent in journals mainly devoted to health economics, assessment of health care technologies, or health care administration. One included study had been published as a dissertation. Sixty-seven of the articles were in English, one in Norwegian, one in Dutch, and one in Spanish.

A total of 31 per cent of the studies were mainly concerned with pharmacological therapy and 26 per cent with surgical interventions. The rest covered various types of conservative treatment, rehabilitation, diagnostic imaging or preventive services. The interventions studied covered a wide range from transplantation surgery to spa-exercise therapy. The most commonly studied interventions were treatment of coronary heart disease, total hip arthroplasty and cochlear implant with four studies concerned with each of them. An economic analysis was present in 86 per cent of the studies, nine studies reported only HRQoL and QALY results.

Further details of each study are given in Appendix B, in which studies are considered in terms of clinical speciality, intervention, aim, data used, method used, perspective of economic analysis, cost data used, results concerning HRQoL assessment, number of and cost per QALYs gained by intervention, quality of study and any methodological or other limitations. Two of the studies were the subject of more than one paper, in those cases the results of separate articles were combined in the table.

3.3 Study design and quality

Qualitative analysis showed that approximately half of the articles were based on randomised controlled trials. Also most of the remaining studies were comparative. Study performance on the scale of 0–10 (Table 2) was considered good (8–10 points on the scale) in 49 per cent of the studies, fair (6–7 points) in 29 per cent of the studies, and fair-to-poor (4–5 points) in 22 per cent of the studies. None of the studies was deemed to be of poor quality (< 4 points on the scale). Four studies used economic modelling, three the Markov model, and one a decision analytic model.

3.4 Reported number of QALYs and costs per QALY

The reported number of QALYs gained by various treatments varied widely from negative to eight depending on the intervention studied and, partly, on over how many years the QALY gain was extrapolated. Also the cost per QALY showed great variation from less than €1000/QALY to over €1 000 000/QALY.

3.5 Study conclusions

Apart from reporting QALY results, most of the studies also discussed the cost per QALY in terms of acceptability for the society. Forty-nine per cent of the studied interventions were viewed by the authors of the original studies as being cost-effective; only 13 per cent of the interventions were deemed not to be cost-effective.

Table 2. Classification of study performance according to Hailey et al. 2004 (b).

Areas of interest	Points considered
1 Patient selection	Methods of randomization/ selection. Equivalence of intervention and control groups. Drop-outs prior to commencement of intervention
2 Description/ specification of the interventions	Adequate description for both intervention and control groups
3 Specification and analysis of study	Sample size; statistical methods used; clear specification of outcome measures
4 Patient disposal	Length of follow up; drop-outs; compliance failures
5 Outcomes reported	Fullness and clarity of reporting. Missing results; statistical summary. Whether conclusions were consistent with data.

Each of these five areas was given a score from 0 to 2 (maximum score 10) based on the following observations:

0 = Relevant information was missing or given in only minimal detail

1 = Reasonable detail was provided but there were some important limitations

2 = Information was satisfactory, there were no significant limitations

Table 3. HRQoL instrument used in the study (in some studies several instruments were used).

Instrument	N	% of all	%
HRQoL instruments			76
15D	2	2.5	
EQ-5D (EuroQol)	37	46.8	
HUI (Health Utilities Index - Mark II/III)	7	8.9	
QWB (Quality-of-Well-Being Scale)	5	6.3	
Rosser-Kind	5	6.3	
SF-6D	4	5.1	
Direct valuation			24
SG (Standard Gamble)	4	5.1	
TTO (Time-trade-off)	7	10.1	
Rating Scale	2	3.8	
VAS	4	5.1	

Table 4. Clinical specialities of the included studies.

Speciality	N (%)
Orthopaedics	11 (15,7)
Pulmonary disease	9 (12,9)
Cardiology	7 (10)
Neurology	6 (8,6)
Rheumatology	6 (8,6)
Otorhinolaryngology	5 (7,1)
Transplantation surgery	5 (7,1)
Psychiatry	4 (5,7)
Oncology	4 (5,7)
Gynaecology	2 (2,9)
Intensive care	2 (2,9)
Urology	2 (2,9)
Endocrinology	2 (2,9)
Infectious disease	1 (1,4)
Nephrology	1 (1,4)
Dental surgery	1 (1,4)
Gastroenterology	1 (1,4)
General surgery	1 (1,4)

Table 5. Country of origin of the studies.

Country	N (%)
United Kingdom	23 (32,9)
The Netherlands	18 (25,7)
Canada	11 (15,7)
USA	11 (15,7)
Spain	2 (2,9)
Germany	1 (1,4)
Hong Kong	1 (1,4)
Norway	1 (1,4)
Sweden	1 (1,4)
Multinational	1 (1,4)

DISCUSSION

Although QALY is considered an important measure of effectiveness of health care, only a fairly limited number of studies are really based on actual measurements of patients' HRQoL. In many studies identified during the review process, HRQoL data were obtained from vaguely defined sources or estimated by health care professionals. Even though health care professionals certainly are aware of the clinical nature of a disease and the burden it can cause for their patients, it is unlikely that they – having never experienced the disease themselves – would really be able to judge patients' HRQoL properly. This is evidenced by a recent study in which correlations between prostate cancer patients' and clinicians' utilities were very low and not statistically significant (10), and also in line with some earlier findings (11–13). Therefore, studies based on real patient data probably are of much more value for the decision maker pondering over allocation of resources.

On the other hand we identified several studies in which HRQoL had been studied in a proper before-after setting, but which did not include the term QALY in their reports. Most of those studies would have provided a possibility to calculate also QALYs, but as they did not, they were not included, as we could not be sure whether we would have covered them extensively with the search strategy specifically designed to identify studies reporting QALYs. It is possible that studies in which the effect of an intervention on HRQoL is absent or minimal are more likely not to report QALYs as those with a more positive result and that, therefore, the studies included in our review may give a somewhat biased and over-optimistic impression about the QALY gains that can be achieved by medical interventions.

Our approach differs from earlier reviews as we were explicitly looking only for studies in which the reported QALYs were based on a before – after assessment of HRQoL with a valid instrument. Consequently, it is not possible to compare – at least in a very productive manner – our results to those of earlier reviews having employed different inclusion criteria. We acknowledge that it is not always possible to obtain valid HRQoL data directly from the patients (for instance in children, in demented patients etc.) and in such instances HRQoL must be estimated using proxies or other methods. Although such studies were excluded from the present analysis, it does not mean that their results would not be valuable. Neither does it imply that earlier attempts to systematically review such studies and to describe their results, like for instance the comprehensive Cost Effectiveness Analysis Registry (<http://www.tufts-nemc.org/cearegistry/index.html>) would be of lesser importance than our approach. However, in most cases patients can certainly independently rate their HRQoL themselves. We feel that in such cases the use of estimated data may be misleading and, therefore, wanted to scrutinize the existing scientific literature reporting QALYs to be able to list and describe the studies that really are based on patient-derived data.

Study quality was in the majority of cases fair to good and approximately half of the studies were based on randomised controlled trials. The results of the studies can thus in most cases be considered reliable and may have a direct influence on decision making. Previous studies have indicated continuing variation in

the quality of cost-utility studies (14–15). Although the main emphasis of our review was not on the quality of the included studies and our approach to quality assessment was therefore not as stringent as that of Gerard (14–15), our results may indicate that there has been slow improvement in the quality and reporting of cost-utility analyses over the years. This is in agreement with the recent results of Neumann et al. who, comparing articles published in 1998 to 2001 with those published in 1976 to 1997, reported an improvement in almost all quality categories assessed (16).

Of the HRQoL instruments employed in the reviewed studies, the EQ-5D was by far the most popular one. Although the instrument has limitations, as for instance the fact that it can only define 243 health states, it is simple to use. Furthermore, it has been developed in co-operation of several health economist groups in different countries with a strong representation from the UK and the Netherlands, which probably has furthered its adoption into research practice in those countries, from which most of the reviewed studies indeed come from.

In studies where more than one generic HRQoL instrument had been used concurrently, the number of QALYs gained differed fairly much depending on the instrument used. This is certainly a limitation for the use of QALY data based on different instruments for comparison of the effectiveness of interventions. Another limitation preventing meaningful comparisons regarding the utility of various interventions is the fact that there is great variation in the way the QALY results are expressed especially regarding the extrapolation of the results over time. In some articles the time horizon on which the QALY calculation was based on was less than a year, in some others several years or until death. At the same time, the comparison of cost-utility results is hampered by variation in the costing methods and the many different perspectives (e.g. societal or health-care provider perspective) from which cost-utility is considered. Consequently, to make the results of cost-utility analyses more valuable, there clearly is a need for common, widely agreed methodology.

Almost half of the studied interventions were considered acceptable for the society in terms of cost per QALY. The threshold used for the society's willingness to pay per QALY, however, varied indicating that at the moment, there is no universally accepted standard for an appropriate threshold for resource-allocation decisions. In this respect, our results are in line with a recent report about the prevailing judgments about society's willingness to pay for a QALY (17).

5

CONCLUSIONS

Although QALYs gained are considered an important measure of effectiveness of health care, the number of studies reporting QALYs based on actual measurement of patients' HRQoL is still fairly limited. Such studies, however, are urgently needed to ensure that allocation of health care resources is based on scientific evidence on the value of various interventions regarding their ability to produce societal welfare.

LÄHTEET/REFERENCES

1. Drummond MF, O'Brien BJ, Stoddard GL, Torrance GW. 1997. *Methods for the Economic evaluation of Health Care Programmes*, Oxford University Press.
2. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res* 1989;248:13-4.
3. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: Treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am* 1969;51:737-755.
4. Gold MR, Russell LB, Siegel JE, Weinstein MC. 1996. *Cost-Effectiveness in Health and Medicine*. Oxford University Press. Oxford.
5. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2:i-iv: 1-74.
6. Kopec JA, Willison KD. A comparative review of four preference-weighted measures of health-related quality of life. *J Clin Epidemiol* 2003;56:317-325.
7. Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgements. *BMJ* 2004;329:224-227.
8. Hailey D, Ohinmaa A, Roine R (a). Evidence for the benefits of telecardiology applications: A systematic review. Alberta Heritage Foundation for Medical Research, 2004.
9. Hailey D, Ohinmaa A, Roine R (b). Study quality and evidence of benefit in recent assessments of telemedicine. *J Telemed Telecare* 2004;10:318-324.
10. Elstein AS, Chapman GB, Chmiel JS et al. Agreement between prostate cancer patients and their clinicians about utilities and attribute importance. *Health Expect* 2004;7:115-25.
11. Epstein AM, Hall JA, Tognetti J, Son LH, Conant L. Using proxies to evaluate quality of life. *Med care* 1989;27(Suppl):91-98.
12. Sintonen H. The 15D-measure of health-related quality of life. I. Reliability, validity and sensitivity of its health state descriptive system. National Centre for Health Program Evaluation, Working Paper 41, Melbourne 1994 (<http://www.buseco.monash.edu.au/centres/che/publications.php#rp>).
13. Slevin ML, Plant H, Lynch D, Drinkwater J, Gregory WM. Who should measure quality of life, the doctor or the patient? *Br J Cancer* 1989;57:109-112.
14. Gerard K. Cost-utility in practice: a policy maker's guide to the state of the art. *Health Policy*. 1992;21:249-79.
15. Gerard K, Smoker I, Seymour J. Raising the quality of cost-utility analyses: lessons learnt and still to learn. *Health Policy*. 1999;46:217-38.
16. Neumann PJ, Greenberg D, Olchanski NV, Stone PW, Rosen AB. Growth and quality of the cost-utility literature, 1976-2001. *Value Health* 2005;8:3-9.
17. Greenberg D, Winkelmayr WC, Neumann PJ. Prevailing judgments about society's willingness to pay for QALY or life-year gained. *Ital J Public Health* 2005;2(Supplement 1):301.

LIITE/APPENDIX 1. LITERATURE SEARCH SUMMARY

	Database	Platform	Search Terms
1	The Cochrane Database of Systematic Reviews The Cochrane Central Register of Controlled Trials The Cochrane Methodology Register The NHS Economic Evaluation Database Health Technology Assessment Database The Database of Abstracts of Reviews of Effects	Update-Software: Cochrane Library, Issue 2, 2004 Wiley InterScience - 2004 Licensed Resource	#1 quality adjusted life years (MeSH) #2 "quality adjusted life year" OR "quality adjusted life years" #3 qaly OR qalys #4 #1 OR #2 OR #3 #5 eq-5d OR eq5d OR 15d OR sf6d OR sf-36 OR sf36 OR euroqol OR aqol OR "health utilities index" OR "standard gamble" OR tto OR "time trade off" OR "time tradeoff" OR "Rosser Kind" (T1 OR AB) #6 qualit* OR qaly OR qalys #7 #5 AND #6 #8 EXP economics (MeSH) #9 cost OR costs OR econom* #10 #8 OR #9 #11 #7 AND #10 #12 #4 OR #11
2	PubMed, National Library of Medicine: 1966–June 2004	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi	#1 quality adjusted life years (MeSH) #2 "quality adjusted life year" OR " quality adjusted life years" (TW) #3 qaly OR qalys #4 quality of life (MeSH, Major Topics) Limits: 1966-1995 #5 #1 OR #2 OR #3 OR #4 #6 eq-5d OR eq5d OR 15d OR sf6d OR sf-36 OR sf36 OR euroqol OR aqol OR "health utilities index" OR "standard gamble" OR tto OR "time trade off" OR "time tradeoff" OR "Rosser Kind" (TW) #7 qualit* OR qaly OR qalys #8 #6 AND #7 #9 EXP economics (MeSH) #10 cost OR costs OR econom* (TW) #11 outcome assessment (health care) (MeSH) #12 quality assurance, health care (MeSH) #13 patient acceptance of health care (MeSH) #14 #9 OR #10 OR #11 OR #12 OR #13 #15 #5 OR #8 #16 #14 AND #15
3	Science Citation Index, Institute of Scientific Institute 1986–June 2004	ISI Web of Knowledge Licensed Resource	#1 "quality adjusted life year" OR " quality adjusted life years" #2 qaly OR qalys #3 #1 OR #2 #4 eq-5d OR eq5d OR 15d OR sf6d OR sf-36 OR sf36 OR euroqol OR aqol OR "health utilities index" OR "standard gamble" OR tto OR "time trade off" OR "time tradeoff" OR "Rosser Kind" (TS) #5 qualit* OR qaly OR qalys #6 #4 AND #5 #7 econom* #8 cost OR costs #9 #7 OR #8 #10 question* #11 #3 AND (#9 OR #10) #12 #6 AND #10 #13 #11 OR #12

	Database	Platform	Search Terms
4	EMBASE, Elsevier: Excerpta Medica 1966–June 2004	http://www.embase.com Licensed Resource	#1 quality adjusted life year (KW) #2 "quality adjusted life year" OR "quality adjusted life years" (TW) #3 qaly OR qalys (TW) #4 #1 OR #2 OR #3 #5 eq-5d OR eq5d OR 15d OR sf6d OR sf-36 OR sf36 OR euroqol OR aqol OR "health utilities index" OR "standard gamble" OR tto OR "time trade off" OR "time tradeoff" OR "Rosser Kind" (TI, AB, KW) #6 quality OR qaly OR qalys OR "quality adjusted life years" (TI, AB, KW) #7 #5 AND #6 #8 EXP cost #9 EXP cost benefit analysis #10 EXP cost effectiveness analysis #11 EXP economic evaluation #12 EXP health economics #13 EXP health status #14 #8 OR #9 OR #10 OR #11 OR #12 #15 questionnaire (KW) OR question* #16 #4 AND (#14 OR #15) #17 #7 AND #14 #18 #16 OR #17
5	CINAHL, Cumulative Index to Nursing & Allied Health Literature Cinahl Information System 1982–Jun 2004	OVID Licensed Resource	#1 quality adjusted life year.mp. #2 quality adjusted life years.mp. #3 qaly OR qalys.mp. #4 EXP quality of life (KW) OR quality of life.mp. #5 #1 OR #2 OR #3 OR #4 #6 eq-5d OR eq5d OR 15d OR sf6d OR sf-36 OR sf36 OR euroqol OR aqol OR "health utilities index" OR "standard gamble" OR tto OR "time trade off" OR "time tradeoff" OR "Rosser Kind" OR short form-36 health survey.mp. #7 qaly OR qalys OR quality #8 #6 AND #7 #9 EXP questionnaire #10 EXP health care costs #11 EXP costs and cost analysis #12 EXP quality assurance #13 "outcome assessment health care".mp. #14 "patient acceptance".mp. #15 #9 OR #10 OR #11 OR #12 OR #13 OR #14 #16 #5 AND #15 #17 #8 AND #15 #18 #16 OR #17

Notes:

Truncation: The * symbol is a truncation character that retrieves possible suffix variations of the root word e.g. surg* retrieves surgery, surgical, surgeon, etc. In databases accessed via the OVID platform the truncation character is \$;

LIITE/APPENDIX 2. DETAILS OF SELECTED STUDIES

Lukiessasi raporttia tietokoneen näytöltä käytä Adobe Readerin aukeamatoimintoa nähdäksesi taulukon molemmat sivut samanaikaisesti.

When reading this report on screen please use the Adobe Reader's double-page spread function to be able to see both sides of the table simultaneously.

Abbreviations:

15D: 15D HRQoL -instrument
AQoL: Assessment of Quality of Life
CE: Cost-effectiveness
CI: Confidence interval
CIS: Checklist Individual Strength
CUA: Cost-utility analysis
DM: Deutsche Mark
€: Euro
EQ-5D: EuroQol HRQoL -instrument
ESP: Spanish peseta
Hfl: Dutch florin = Netherlands guilder
HRQoL: Health-Related Quality of Life
HUI: Health Utilities Index HRQoL -instrument (same as HUI3, HUI-Mark III)
ICER: Incremental cost-effectiveness ratio
NHS: National Health Service
NLG: Netherlands guilder
NOK: Norwegian krone
N.S.: Statistically not significant
QADS: Quality-adjusted days
QoL: Quality of Life
QALE: Quality-adjusted life expectancy
QALM: Quality-adjusted life months
QALY: Quality-adjusted life years
QWB: Quality of Well-being Scale
Rosser-Kind: Rosser-Kind HRQoL -instrument (same as The Charing Cross health indicator and Rosser index)
RS: Rating scale
SF-6D: SF-6D HRQoL -instrument
SG: Standard Gamble
SKr: Swedish krone
TTO: Time Trade-off
VAS: Visual analogue Scale
Vs: Versus
\$: US dollar
Can\$: Canadian dollar
£: Great Britain pound

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
CARDIOLOGY					
Campbell et al 2001, United Kingdom	Transmyocardial laser revascularization (TMLR) + medical management	To compare the cost-utility of laser revascularization + medical management to that of continued medical therapy alone	188 patients with refractory angina and not suitable for conventional forms of revascularization randomized to receive TMLR (n=94) or medical management alone (n=94)	HRQoL assessment was completed at baseline, and at 3, 6, and 12 month	NHS perspective. Costs to NHS of TMLR, and all secondary care contacts and cardiac related medication in the 12 month following randomization
Eefting et al 2003, The Netherlands	Coronary stenting or off-pump bypass surgery	To compare QoL and CE of stenting to that of off-pump by-pass surgery	280 patients referred for angioplasty randomly assigned to stenting (n=138) or off-pump bypass surgery (n=142) (131 and 136 patients, respectively, received the assigned treatment)	HRQoL assessed at baseline and at 1, 6, and 12 month after treatment	Societal. Both direct (during the initial hospitalization and follow-up until 1 year) and indirect (calculated by the "friction cost" method) costs assessed
Mitton et al 1999, Canada	Permanent pacemaker therapy	To determine the cost-utility of pacemaker therapy compared with conventional treatment	38 patients with frequent neurally mediated syncope (NMS) split to 2 groups according to the frequency of syncopal spells, a more severe (19) and a less severe group (19)	Non-randomized, before - after design. Pre- and post-pacemaker follow-up interval 12 month. HRQoL measured before and between 3-24 (average 12) month after intervention	Payer's perspective. Only direct health care costs related with syncope included
Nathoe et al 2003, The Netherlands	Off-pump coronary bypass surgery	To compare cardiac outcome and CE of on-pump and off-pump surgery one year after operation	139 low-risk patients assigned to on-pump surgery and 142 to off-pump surgery	Multi-centre, randomized trial. HRQoL assessed at baseline and 1, 3, 6 and 12 month after surgery	Direct medical costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	Over the year from hospitalization TMLR was £8901 more expensive than medical therapy	The mean QALY gain in TMLR 0.486 (95% CI 0.429–0.528) and in medical management 0.447 (0.393–0.492). The mean QALY difference, in favour of TMLR was 0.039 (95% CI -0.033 to 0.113, N.S.)	Analysis only considers costs and QALYs to 1 year	The incremental cost per QALY was over £228 000. For the studied patients TMLR was considered inefficient use of resources
EQ-5D	Overall costs for stenting \$8276, for off-pump surgery \$11 209	At 1 month HRQoL significantly higher after stenting than after off-pump surgery but comparable at 1 year. Number of QALYs 0.82 after stenting, 0.79 after off-pump surgery (p=0.09)	The open design of the study may have affected the assessment of outcome	Stenting produced cost savings of \$97 767 for QALY gained compared to off-pump surgery. Stenting can be recommended as a first choice revascularisation strategy
EQ-5D	Mean cost per patient during pre-pacemaker year \$2032, during 1st post-pacemaker year (including pacemaker insertion) \$10 905	The mean incremental EQ-5D VAS - score 13.5 (an improvement from 62.7 to 76.2). Incremental EQ-5D index score 0.089 (an improvement from 0.748 to 0.837). Extrapolated over 10 years, 1.02 QALYs gained per patient	Non-controlled, small sample size, short follow-up	Mean cost per QALY in the whole group Can\$13 159 considered acceptable
EQ-5D	On-pump surgery associated with \$1839 in additional direct costs per patient	EQ-5D-score increased in both groups from baseline 0.65 to 0.84 3 month after surgery. Subsequent scores remained within normal limits. QALYs increased by 0.83 after on-pump, and by 0.82 after off-pump surgery	Relatively low-risk patient material, results cannot be extrapolated to patients with more advanced coronary artery disease. Difference in HRQoL at one year follow-up between the treatment groups N.S.	ICER for on-pump surgery \$183 000/QALY, clearly above the general societal willingness-to-pay threshold of \$20 000. Off-pump surgery is more cost-effective with 95% certainty

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Rawles et al 1993, United Kingdom	Thrombolytic treatment either at home or in hospital	To measure HRQoL and the loss of QADS after suspected myocardial infarction (AMI) in patients who received thrombolytic treatment	311 patients with suspected AMI and no contraindications to thrombolytic treatment seen at home within 4 hours of the onset of symptoms randomized to receive thrombolysis at home (n=163) or at hospital (n=148)	Randomized double-blind parallel group trial. Preadmission HRQoL assessed in an interview before discharge from hospital. Follow-up interview conducted 3 month after suspected AMI	
Schulman et al 1996, USA	Epoprostenol	To assess costs and outcomes of epoprostenol therapy	471 patients with severe congestive heart failure, randomized to receive epoprostenol (n= 239) or best usual care (n=232)	Prospective economic evaluation that served as a secondary end-point for the FIRST study, a randomized international multicentre trial. HRQoL assessed at time of randomization, and 2 week, 1 and 3 month and every 3 month thereafter	Health service perspective, direct medical costs excluding costs of medication

DENTAL SURGERY

Cunningham et al 2003, United Kingdom	Orthognathic treatment	To determine the cost-utility of orthognathic treatment	21 patients, 15 having bimaxillary surgery and 6 mandibular surgery alone	Patients interviewed at start and end of treatment and 3 times during treatment	Direct health service costs of the intervention discounted at 6% a year
---------------------------------------	------------------------	---	---	---	---

GASTROENTEROLOGY

Kennedy et al 2003, United Kingdom	Evidence-based self-help guidebook and patient-centred consultations. Patients prepared a written self-management plan and self-referred to services based on a self-evaluation of their need for advice	To determine if a whole systems approach to self-management improves clinical outcomes and leads to cost-effective use of services	700 ulcerative colitis or Crohn's disease patients (297 at intervention sites and 403 at controls sites). For economic analysis 651 patients having returned either or both questionnaires were included	19 hospitals randomized to 10 control and 9 intervention sites. HRQoL assessed at baseline and at 1 year	NHS perspective
------------------------------------	--	--	--	--	-----------------

QALY Instrument	Costs	Results	Limitations	Cost/QALY
Rosser-Kind	No cost data reported	Mean premorbid HRQoL-scores similar in home and hospital groups (0.91 vs 0.89). Difference in proportion of survivors who had not achieved their pre-morbid HRQoL -score by 100 days N.S. Difference in median loss of QADS N.S.	Short follow up time	Not reported
EQ-5D (apparently VAS only)	Costs of usual care \$11 797 (\$22 476 when extrapolated to 12 month), costs of usual care + epoprostenol \$16 819 (\$27 747 when extrapolated to 12 month)	Mean EQ-5D score at baseline 46.0 in the usual care group and 50.0 in the epoprostenol group. Usual care produced 2.3 QALMs, epoprostenol 2.2 QALMs (extrapolated to 12 month: 4.35 and 3.7 QALMs)	Early conclusion of the trial, follow-up time ranged from few week to 18 month (mean less than 5 month). Some data quality assurance steps were not undertaken for data collected late in the trial. HRQoL assessment based on VAS data only	Not reported
TTO	Discounted total cost was £3182/ patient	Mean QALYs gained 5.67	Small sample size, lack of a formal control group	Incremental cost for each additional QALY was £561. High probability of treatment being cost-effective
EQ-5D	Total costs over the 12 month period £922 in the intervention group and £1070 in the control group	Mean EQ-5D scores were reduced in both groups at follow-up compared to baseline. Mean QALY gain in the intervention group -0.01892, in the control group -0.018780, i.e. a reduction of QALYs in both groups	There may have been some contamination regarding selftreatment in the control group	Usual care results in QALY gain of 0.00022 at increased cost of £148 per patient. Incremental cost per QALY of usual care thus £676 417. Self-management model more likely to be cost-effective than existing practice

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
GENERAL SURGERY					
van Gemert et al 1999, The Netherlands	Vertical banded gastroplasty (VBG)	To determine CE of VBG	21 morbidly obese patients	2 alternatives compared: VBG and no treatment. In no treatment group costs of treatment of morbid obesity-related comorbidities determined by cost-of-illness analysis using costs and QoL values obtained from preoperative patients. Analysis of the VBG group incorporated total (direct and indirect) costs, QoL outcome and life-years gained after surgical treatment. HRQoL questionnaires administered by mail 1 month before, and 1 and 2 years after surgery	Societal. In the CUA total incremental costs determined as the costs attributable to morbid obesity (cost-of-illness) multiplied by the life expectancy of a morbidly obese population minus the direct and indirect costs of surgical treatment and the costs of failure of operation
GYNAECOLOGY					
Garry et al 2004 and Sculpher et al 2004 (two reports) United Kingdom	Hysterectomy	To compare both standard abdominal (AH) and vaginal hysterectomy (VH) with laparoscopic hysterectomy (LH)	Women requiring a hysterectomy for reasons other than malignancy. 292 patients had AH, 168 patients VH, and 920 patients LH	Women were randomized between the selected conventional procedure and laparoscopic procedure in 2 parallel trials. HRQoL assessed at baseline and at 6 week, 4 month and 1 year after hospital discharge	NHS perspective. Resource use data included theatre resources, hospital stay and costs incurred during the postoperative period

QALY Instrument	Costs	Results	Limitations	Cost/QALY
VAS	Total direct costs of VBG \$5865. According to the cost-of-illness analysis, lifelong costs of illness of morbidly obese persons ranged from \$8304 to \$9367	1 year after VBG significant improvements on the NHP-1, NHP-2 and VAS outcomes. 2 years after VBG the QoL outcomes on NHP-1, NHP-2, and VAS were still significantly improved, but slightly less than at 1 year. 3.6 life years were gained. The mean improvement on the VAS scale was 0.25. VBG gained 12 QALYs when a lifelong scenario was considered	In the cost-of-illness analysis 2 different prevalences for morbid obesity used since real prevalence is not known. Also prevalences for morbid obesity-related comorbidities estimated. The 2-year results are extrapolated to a lifelong scenario so the results have to be interpreted with care. Only 21 patients. HRQoL assessment based on VAS data only	CE of VBG established by a decrease in costs and gain in QALYs
EQ-5D	LH was £401 more expensive than VH and £186 more expensive than AH	Compared with VH, LH produced 0.0015 more QALYs, compared with AH it produced 0.007 more QALYs		Compared to VH, the incremental cost per QALY gained by LH was £267 333. The probability that LH is cost-effective <50% for a large range of willingness to pay values. Compared to AH, the incremental cost per QALY gained by LH was £26 571. If the NHS willing to pay £30 000 for additional QALYs, probability that LH is cost-effective compared to AH is 56%

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Manca et al, 2003, United Kingdom	Tension-free vaginal tape	To assess CE of tension-free vaginal tape compared with open colposuspension	Base-case analysis conducted on 214 patients (117 with tension-free vaginal tape and 97 with colposuspension). (Only observations with complete data (62%) included in the base-case analysis)	CUA alongside a multicentre, randomized, comparative trial. HRQoL assessed at baseline, and at 6 week and 6 month after hospital discharge	Data on resource utilization prospectively recorded from hospitalization to 6 month from discharge. Direct, undiscounted medical costs only
ENDOCRINOLOGY					
Kaplan et al 1988 USA	Behavioural interventions: diet+exercise versus education-control	To estimate the cost-utility of behavioural intervention in the management of non-insulin dependent diabetes mellitus (NIDDM)	76 subjects with diagnosed NIDDM recruited through public radio announcements, newspaper notices, and physician referrals HRQoL measured at initial interview, and at 3, 6, 12, and 18 month	Patients were randomly assigned to one of 4 experimental conditions: diet, exercise, diet+exercise or education control. Here, only the differences between the best treatment (diet+exercise) and the education-control are considered	Costs of diet+exercise program estimated using 1986 clinical charges, only direct costs considered. No discounting as costs and benefits were evaluated in the short term
Kristiansen et al 1997 Norway	Alendronate	To establish CE of 5 years intervention with alendronate	31 patients (11 with femoral fracture, 10 with forearm fracture, and 10 with vertebral fracture)	Simulation model for women aged 65 years with a bone mineral density (BMD) of the femoral neck 2.5 SD below peak bone mass. HRQoL measured at baseline and through a simulation model after treatment	Societal. Direct medical costs, discounted by 7%
INFECTIOUS DISEASES					
O'Brien et al 2003 Canada	Oseltamivir	To establish CE of oseltamivir in treatment of influenza when it is on formulary compared to off-formulary	A cohort of otherwise healthy adults aged 16 to 64 years	Decision analytic model. Utility data pooled from 2 published and 2 unpublished trials	Direct medical costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	Total cost £1058 in the tension-free vaginal tape group and £1301 in the colposuspension group, respectively	HRQoL increased from baseline 0.778 to 0.806 6 month after operation in the tension-free vaginal tape group, and from 0.785 to 0.794 in the colpo-suspension group. 0.397 QALYs gained in the tension-free vaginal tape and 0.387 QALYs in the colposuspension group, respectively	Significant number of dropouts before intervention. The 6-month follow-up is rather short	Probability of tension-free vaginal tape being more cost-effective than colposuspension 94.6% if the decision maker is willing to pay £30 000 per additional QALY
QWB	The total costs of the program were approximately \$1000. Costs for the education-control program are not given	At 18 month, the diet+exercise group had improved by 0.06 on the QWS compared with -0.04 for the controls, yielding a difference of 0.092 between the groups. The well-year gains across treatment periods yields a utility of 0.092 well-years	Only short-term benefits are considered. The observations were based on a single experimental study with self-selected patients. The number of participants was small and their representativeness in relation to all diabetic patients is questionable. The follow-up period was limited to 18 month	Using the average cost of \$1000, the cost-utility was \$10 870 per well-year. Cost-utility considered to compare favourably with many widely accepted medical interventions
15D	Cost of treatment for 5 years was NOK19 958/patient. Avoided future costs estimated at NOK5630 resulting in a net cost of treatment of NOK14 329	Baseline 15D: femoral fracture 0.75, forearm fracture 0.78, vertebral fracture 0.85. After treatment values using simulation model 0.80, 0.92 and 0.99, respectively. Discounted number of QALYs gained 0.049	HRQoL data derived from a small sample of interviewed patients (and adjusted for analysis), no before-after data.	Discounted cost per QALY NOK528 000, NOK291 000 and NOK147 000 when BMD was 1.5, 2.5 and 3.5 SD below peak bone mass, respectively
VAS	If oseltamivir is available for prescription the additional cost per patients is CAN\$20.50	Treatment yields 0.01173 QALYs for oseltamivir not on formulary and 0.01208 QALYs for oseltamivir on formulary	Model based on data from several trials and certain assumptions. VAS values normalized into 0 to 1 interval to allow calculation of QALYs. HRQoL assessment based on VAS data only	Incremental CE of oseltamivir Can\$57 863 per QALY. Oseltamivir considered potentially cost-effective

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
INTENSIVE CARE					
Hamel et al 2000 USA	Mechanical ventilation and intensive care	To estimate CE of providing mechanical ventilation and intensive care	1005 patients with acute respiratory failure requiring ventilator support (963 patients received ventilator support, in 42 cases it was withheld)	Trained interviewers gathered information 2 week before and 6 month after study entry. Unavailable patient's reports were substituted by a surrogate's report. Markov model was used to estimate expected lifetime costs and quality-adjusted life expectancy	Direct medical costs. Discount rate 3% per year. Health care providers perspective
Ridley et al 1994 United Kingdom	Intensive therapy unit (ITU)	To measure HRQoL in survivors from ITU, relate changes in HRQoL to individual treatment costs and estimate the average costs/QALY gained compared with other health care programmes	56 survivors of critical illness 1 year after their admission to ITU	HRQoL assessed 1 year after admission. Baseline HRQoL apparently assessed retrospectively	
NEPHROLOGY					
Kroeker et al 2003 Canada	Home hemodialysis	To compare the economics of short daily, long nocturnal and conventional thrice weekly hemodialysis (HD)	N=44, retrospective analysis of patients' conventional HD costs for 12 month before study were compared to the costs when on quotidian HD. 10 daily, 12 nocturnal and 22 conventional HD patients	HRQoL assessed before and immediately after training, 1 month after the study began, and then at 3-month intervals	Costs borne by the public health care system (an operating cost comparison, patient training and home installation costs etc. excluded)
NEUROLOGY					
Gudex et al 1997 United Kingdom	Botulinum toxin (BT)	To assess the cost-utility of BT therapy for patients with dystonias	130 patients with current diagnosis of dystonia	Patients completed at least 5 consecutive HRQoL questionnaires over 6 month	Direct costs to the health service and patients

QALY Instrument	Costs	Results	Limitations	Cost/QALY
TTO	Mean health care costs from entry through 1 year varied from \$77 000 to \$95 000 depending on the risk group	Mean HRQoL at 6 month (n=225) 0.88±0.22	Some of the health care costs and life expectancy based on estimations. HRQoL determined for only a part of the patients. It was assumed that QoL values for patients not interviewed were similar to those of the interviewed at 6 month	For patients with > 70% probability of surviving 2 month, the incremental cost of QALY was \$29 000, for those with 51% to 70% probability of surviving 2 month \$44 000, and for those with < 50% probability of surviving 2 month, \$110 000. Treatment of the first 2 groups considered economically worthwhile
Rosser-Kind	Direct hospital costs. Discounting 5% for 1 year. Mean cost per day £550 (mean duration of admission 4.4 days)	For the majority of patients HRQoL did not alter after critical illness. HRQoL decreased in patients admitted with respiratory problems and in those who previously enjoyed a good QoL	Possible "recall bias" i.e. the patients' memories may be coloured by present perceptions of their health. Small number of patients and wide range of diagnoses	Total hospital costs per QALY £7500 (nonsurvivors not included)
HUI	Total operating costs Can\$74 371 for nocturnal, Can\$72 688 for conventional, and Can\$ 67 281 for daily HD, respectively	Daily HD and nocturnal HD groups showed an increase in QoL, whereas conventional HD patients experienced a decline. Annualized QALY 0.84 in daily, 0.70 in nocturnal, and 0.71 in conventional HD group, respectively	Small number of subjects. The study was clearly underpowered to detect costing differences of statistical significance	Total annualized cost/QALY Can\$85 442 in daily, Can\$120 903 in nocturnal, and Can\$116 753 in conventional HD group, respectively
EQ-5D	Total costs per year were £975.40 for BT patients and £145.38 for non-BT patients	HRQoL data not reported. The QALY gain over 1 year associated with the use of BT was 0.024-0.029 QALYs	No proper control group. Several assumptions required for the calculation of QALYs	Cost per QALY using EQ-5D estimated at £30 000 to £35 000, using TTO £45 000 to £55 000

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Kemler et al 2002 The Netherlands	Spinal cord stimulation (SCS)	To assess the economic aspects of treatment of chronic reflex sympathetic dystrophy (RSD)	54 patients with chronic RSD received either SCS together with physical therapy (PT, n=36) or physical therapy alone (n=18). 24 patients responded positively to trial stimulation and underwent SCS implantation	Randomized controlled trial. HRQoL assessed at baseline, 1 day prior to implantation, and 1, 3, 6, and 12 month following implantation	Societal. Routine RSD costs, SCS costs, out-of-pocket costs. Both costs and health effects discounted at a rate of 3%. Sensitivity analysis
McCrone et al 2003 United Kingdom	Intravenous immunoglobulin (IVIg) and prednisolone	To compare CE of IVIg to that of prednisolone	Double-blind multicenter randomized controlled cross-over trial. 32 chronic inflammatory demyelinating polyradiculoneuropathy patients entered into the trial. HRQoL scores available from 25 patients	HRQoL apparently assessed at baseline and at 6 week (not clearly stated)	Societal, costs of lost employment not included
Parkin et al 2000 United Kingdom	Interferon beta-1b	To evaluate CE of interferon beta-1b for relapsing - remitting multiple sclerosis (RRMS)	40 patients who had experienced a relapse in the preceding 6 month and 62 who had not	The remission group patients assessed their HRQoL at the start and end of a 6-week period. The recent relapse group patients assessed how they were currently and how they were at the worst of their relapse	Direct medical costs
Remák et al 2004 United Kingdom	Antiepileptic drugs	To assess the cost-utility of 5 adjunctive newer antiepileptic drugs	125 patients with intractable epilepsy followed in a prospective, non-randomized observational study	HRQoL assessed at baseline and at 3 and 6 month	Direct medical costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	Mean cost per patient to 1 year €9805 in the SCS+PT group and €5741 in the PT group. Mean projected cost per patient to death €171 153 in the SCS+PT group and €229 624 in the PT group. Incremental projected cost (discounted) of SCS+PT to death €-17 927	At 12 month HRQoL in SCS+PT group 0.22, in PT group 0.03. Incremental QALY gain in the SCS+PT group 0.18 per year. Incremental projected QALY gain to death 7.6 (undiscounted) or 2.33 (discounted)		SCS+PT strategy dominant over PT alone regarding cost/QALY gained
EQ-5D	Cost difference between the 2 treatments estimated to be €3754 over the 6-week treatment period	IVIg resulted in a mean relative gain of HRQoL compared with prednisolone of 0.12. 0.014 QALYs gained over 6 week through IVIg compared with prednisolone	Small number of subjects. Short follow-up. The change in HRQoL N.S.	Cost per QALY gained by IVIg €268 000 relative to prednisolone. Over 6 week IVIg appears not to be a cost-effective alternative to prednisolone
EQ-5D and in 50 patients TTO	Average costs £529 in the remission group and £2644 in the recent relapse group, giving relapse costs of £2115 per patient	Utilities: Average remission value 0.604, average relapse value 0.136; a net loss of 0.468 per relapse. For a 5 year model, 0.13 QALYs were gained	A modelling study. Some of the data used in the analysis derived from a separate study. Utility values partly converted from EDSS score values	The cost per QALY gained was £809 900; or £328 300 allowing for effect of progression over 5 years
EQ-5D	Mean total cost ranged from £919 to £1502 in the 5 study groups	Only patients treated with topiramate and vigabatrin had an increase in utility (statistically significant only for topiramate), utility score for patients treated with clobazam, gababentin and laomotrigine declined	Small sample size. Only 62% of patients on their study drug at 6 month	Vigabatrin treatment resulted in higher utility gain at lower cost, i.e. dominated adjunctive treatment with clobazam, gababentin or lamotrigine. Adjunctive treatment with topiramate had an ICER of £7869/QALY gained compared with vigabatrin

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Wonderling et al 2004 United Kingdom	Acupuncture	To evaluate CE of acupuncture compared to usual care in the management of chronic headache	136 patients in the acupuncture arm (up to 12 treatments over 3 month) and 119 in the control arm	SF-6D was used to calculate HRQoL for each patient from their responses to the SF-36 at baseline, 3 month and 12 month	Both NHS and societal perspective. Patients reported unit costs associated with non-prescription drugs and private health care visits
ONCOLOGY					
Uyl-de Groot et al 1998 The Netherlands	GM-CSF as an adjunct to intensive chemotherapy	To compare the effects and costs of GM-CSF as an adjunct to intensive chemotherapy in elderly patients with acute myeloid leukaemia	Patients randomized to receive daunomycin-cytosine arabinoside (n=161) or daunomycin-cytosine arabinoside with GM-CSF (n=157)	Prospective, randomized controlled, multicenter clinical trial. CE-analysis. HRQoL assessed at the start of the first treatment, during hospitalization, after hospitalization, and at 6, 12 and 24 month after randomization	Health service perspective (days in hospital, outpatient visits, day-care departments, laboratory services, diagnostics, radiotherapeutical and surgical procedures, blood products, medication, parenteral nutrition, and consultation). Discounting rate 5%
van Agthoven et al 2004 The Netherlands	Intensive chemotherapy followed by myeloablative chemotherapy and autologous stemcell rescue	Economic evaluation as related to the clinical outcome of a randomized study comparing experimental treatment with intensive chemotherapy alone	90 previously untreated multiple myeloma patients	Cost analysis based on a randomized study of 261 patients of whom 90 matched for age and sex, were included in final analysis. Utility values from patients responding to treatment were available for up to 24 month. Time horizon of the CUA was 36 month, utility values for these patients were assumed to be stable after 24 month	Service providers perspective. Average unit costs calculated for the most important items, reflecting full hospital costs, including overhead costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
SF-6D	Total costs during the 1 year study period were £403 for the acupuncture group and £217 for the control group	SF-6D score increased from the baseline 69.3 to 73.9 at 12 month in the acupuncture group, and from 70.6 to 70.7 in the control group. The mean health gain from acupuncture during the 1 year of the trial was 0.021 QALYs		Incremental cost-effectiveness was estimated to be £9180 per QALY gained, or less if analysis incorporated likely QALY differences for the years after the trial. Acupuncture considered relatively cost-effective compared with a number of other interventions
EQ-5D (VAS only)	Average total costs estimated at \$51 867 in the control group, and at \$58 087 in the GM-CSF group	Considering short-term HRQoL, GM-CSF resulted in more problems. With regard to survival or longterm HRQoL there were no significant differences between the groups. During the 2 year follow-up discounted quality-adjusted survival 0.800 in the control group and 0.816 in the GM-CSF group, respectively	For some patients the completion of the questionnaires appeared to be too much of a burden (high age, serious disease). Timing of administration of questionnaires may have biased the results towards overestimating HRQoL in GM-CSF patients. HRQoL assessment based on VAS data only	Daunomycin-cytosine arabinoside with GM-CSF considered not to be a cost-effective treatment strategy when compared with daunomycin-cytosine arabinoside alone
EQ-5D	Costs of intensive chemotherapy €68 802, of myeloablative treatment €81 643	Eventfree or overall survival did not differ between the groups. Utility values were 0.81 (intensive chemotherapy only) vs 0.65 (myeloablative treatment) at 6 month from randomization, 0.80 vs 0.62 (12 month), 0.81 vs 0.69 (18 month), and 0.77 vs 0.75 (24 month), respectively. Number of QALYs 3 years after randomization 1.87 (intensive chemotherapy only, discounted value 1.81) vs 1.63 (myeloablative treatment, discounted value 1.57)	Utility values assumed to be stable after 24 month. For patients not in remission utility values were estimated. Some uncertainty regarding costs. Longterm follow-up may affect conclusions as 5 years after randomization eventfree survival differed between the groups	Cost per QALY €36 793 (intensive chemotherapy) vs €50 087 (myeloablative treatment). Myeloablative treatment not considered cost-effective

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
van den Brink et al 2004 The Netherlands	Total mesorectal excision (TME) with or without preoperative radiotherapy (PRT)	To compare societal costs and quality-adjusted life expectancy of rectal cancer patients undergoing TME with or without PRT	Data on local recurrence rates, QoL, and costs were obtained from a multicentre randomized clinical trial. 1861 patients from 108 hospitals. HRQoL data was based on 1530 Dutch patients who completed a QoL questionnaire	A Markov model was used to project the clinical and economic outcomes of PRT. HRQoL before treatment, and at 3,6,12,18, and 24 month after surgery	Societal (costs included primary treatment, continuing care, and recurrence treatment, productivity losses (for both paid and unpaid labour), informal care-, travel, time- and out-of-pocket costs)
van den Hout et al 2003 The Netherlands	Radiotherapy in patients with painful bone metastases	To determine the QALE and societal costs for single- versus multiple-fraction radiotherapy. To determine from a societal perspective which radiotherapy schedule provided better value for the money	579 patients receiving single-fraction and 578 patients receiving multiple-fraction radiotherapy	Patients described their health state using EQ-5D. The exact time points when EQ-5D was completed are not described but patients were followed up for 2 years	Societal, both medical and non-medical costs included and discounted at 3%
ORTHOPAEDICS					
Dolan et al 1999 United Kingdom	Theoretical benefit of prevention of osteoporotic fractures	Establishment of QALY loss for Colles' fracture	50 consecutive female Colles' fracture patients	HRQoL at the initial Accident and Emergency visit rated retrospectively, postfracture HRQoL valued at each ambulatory visit	

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	The total estimated societal costs were \$115 000 for PRT+TME and \$105 200 for TME alone	PRT estimated to produce a mean 8 month gain in life expectancy and a mean 5 month gain in quality-adjusted life expectancy. For the base case analysis, PRT leads to a QALY gain of 0.39	Estimates partly based on modelling, not actual data	The observed QALY gain results in a CE-ratio of \$25 100 per QALY. Short-term preoperative radiotherapy in patients with rectal cancer undergoing TME considered both effective and cost-effective
EQ-5D	The estimated societal costs of radiotherapy during first 12 week after randomization were \$2438 for single-fraction and \$3311 for multiple-fraction schedule. The single-fraction treatment schedule saved \$1753 relative to the multiple-fraction schedule	On average, health was valued at approximately 0.40. Overall average life expectancy was approximately 9 month, overall average quality-adjusted life expectancy approximately 4 month. The single-fraction treatment schedule yielded 1.7 quality-adjusted week relative to the multiple-fraction schedule	Difference between single- and multiple-fraction radiotherapy N.S.	Single-fraction schedule likely to be more cost-effective than multiple-fraction schedule
EQ-5D	No cost data reported	Mean HRQoL increased from 0.539 to 0.925 over the average treatment time of 48 days. The fracture causes a mean 0.018 (SD = 0.014) (2%) loss in QALYs. The initial loss in QALYs does not last for the whole treatment period (as assumed in National Osteoporosis Foundation guidelines) but improves gradually during treatment	Prefracture HRQoL valued retrospectively. Mobility dimension omitted from the questionnaire and replaced by a calculated 0.025 loss on mobility for each patient	Not reported

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Gilbert et al 2004 United Kingdom	Early MR or CT imaging	To establish whether early imaging influences treatment and outcome of patients with low back pain	782 patients with symptomatic lumbar spine disorders randomized to early or delayed (imaging performed only if a clear indication developed) selective imaging groups	HRQoL measured at baseline (prior to random assignment to groups) and at 8 and 24 month	Health care provider. Direct health service costs
James et al 1996 United Kingdom	9 different orthopaedic operations	To produce a priority list for orthopaedics for maximization of efficiency in health care purchasing	99 patients undergoing various orthopaedic operations	HRQoL assessed preoperatively and 6 month post-discharge	Direct medical costs considered. Costs calculated by extra contractual referral tariff and patient based individually derived costs. Discounted by 6%
Laupacis et al 1994 Canada	Total hip arthroplasty (THA)	To determine CE of the Mallory Head cemented vs noncemented THA	Prospective, randomized trial. Baseline and 24-month HRQoL data available from 90 patients. Cost data from 30 patients in each arm of the trial	HRQoL determined preoperatively and at 3, 6, 12, and 24 month postoperatively	Direct medical costs and some costs borne by the patient, opportunity costs not included
Lavernia et al 1997 USA	Knee arthroplasty	To calculate cost per quality well year in knee arthroplasty surgery	30 male and 70 female patients	HRQoL assessed preoperatively and at 3 and, 6 month, and 1 and 2 years	Direct hospital costs, no discounting
Marti-Valls et al 2000 Spain	Total hip replacement (THA)	To quantify short-term benefits of THA, and to assess hospital costs of the procedure in Catalonia	332 patients, 70% with arthritis, 59% with associated pathology	Multicentre prospective study. HRQoL assessed prior to operation, and 6 month later	Direct medical costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	Mean total cost per patient for the early group was \$701 and for the delayed group \$614	EQ-5D improved more in the early group (adjusted difference 0.057, p=0.01). Early imaging provided an adjusted mean additional QALY gain of 0.041 during 24 month	Slight baseline differences between the groups in important variables	Incremental cost per QALY \$2124
Rosser-Kind and EQ-5D	Costs ranging from £132 to £5078	Mean patient assessed HRQoL change between -0.006 and 0.176 using Rosser index and between -0.048 and 0.353 using EQ-5D. Net health gain for the different interventions ranging from -0.172 to 4.841 QALYs using Rosser index and from -1.377 to 6.830 QALYs using EQ-5D	Relatively small sample size in the 9 diagnostic categories	Cost per QALY ranging from £-9761 to £6392 using Rosser Index and £-21 864 to £10 672 using EQ-5D
TTO	Cost of initial hospitalization Can\$9990, outpatient costs during first year Can\$1137	HRQoL improved from the preoperative 0.29 to about 0.84 at 3 month after surgery	HRQoL data not clearly reported, both groups combined for analysis	Cost per QALY Can\$8031 during the first 5 years
QWB	Average cost \$11 873	HRQoL preoperatively 0.56, at 3 month 0.56, at 6 month 0.57, at 1 year 0.59, at 2 years 0.59	Significant number of missing answers at the various follow-up points. Discrepancies in numbers reported between tables and abstract	Cost per a quality well year was \$30 695 at 3 month, \$17 804 at 6 month, \$11 560 at 1 year, \$6656 at 2 years postsurgery. Knee arthroplasty considered cost-effective
EQ-5D	Total cost of procedure ESP838 480 (€5039)	Mean QALY gain over 3 years 1.65	A relatively short follow-up period	Cost per QALY ESP505 500 (€3050)

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Ostendorf et al 2004 The Netherlands	Total hip arthroplasty (THA)	To determine the effect of waiting times for THA on loss in QALYs	161 patients put on waiting list for THA	HRQoL assessed when patient placed on waiting list, preoperatively, and 3 and 12 month after surgery	
Patrick et al 2001 USA	Aquatic exercise	To estimate cost and outcomes of the Arthritis Foundation 20-week aquatic exercise classes	125 osteoarthritis patients randomized to aquatic exercise and 125 to control group	HRQoL measured at baseline and at 20 week	Societal, including both travel and time costs, 3% discounting
Rome et al 2004 United Kingdom	Functional or accommodative orthosis	To evaluate clinical effectiveness and CE of 2 different types of foot orthoses used to treat plantar heel pain	48 patients randomly assigned to receive either a functional or an accommodative orthosis	HRQoL assessed at baseline and at 4- and 8-week intervals	NHS and societal perspective (including all other costs of treatments outside the NHS)
Spady 2000 Canada	Knee and hip replacement surgery	To investigate the validity of EQ-5D scores and make a conceptual analysis of the EQ-5D as a measure of HRQoL	540 hip or knee replacement surgery patients	HRQoL was measured pre- and 6 month post-surgery	
Torrance et al 2002 Canada	Hylan G-F 20	To assess CE and cost-utility of viscosupplementation with Hylan G-F 20 (AC+H) compared to appropriate standard care (AC)	255 patients, 128 randomized to receive AC and 127 to receive AC+H	Prospective, randomized, 1-year, openlabel, multicentred trial. Structured telephone interviews, HRQoL measured at baseline, and at 1, 2, 4, 6, 8, 10, and 12 month	Societal. Costs include appropriate knee OA treatment, concomitant medications, outpatient resources, hospitalization, time loss from work, out-of-pocket expenses

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	No cost data reported	EQ-5D decreased from 0.39 to 0.33 during waiting time. After surgery it had improved to 0.71 3 month, and to 0.75 12 month postoperatively. With an average waiting time of 6 month, a loss of 21 QALYs per 100 patients can be anticipated	No control group	Not reported
QWB	Annual direct medical and direct nonmedical costs for treatment group \$1945 and \$1688; for control group \$2922 and \$260, respectively	Mean baseline QWB 0.599, 0.606 and 0.581 and post-class QWB 0.599, 0.602 and 0.613 in controls, non-adherers and adherers, respectively	Adherence to exercise programme low	Cost/QALY gained \$205 186 using the QWB. Incremental cost-utility determined by QWB not considered favourable
EQ-5D	The mean total costs for the functional orthoses were £17.99 higher than those for the accommodative orthoses	Significant improvement in mean HRQoL in the functional orthosis group between baseline and 8 week, no significant improvement in the accommodative orthosis group. Functional orthosis results in a QALY gain of 0.0109	Only 8-week follow-up, number of patients quite small, baseline difference in HRQoL between the groups	Incremental cost per QALY £1650 (\$3003) for functional orthosis
EQ-5D	No cost data reported	EQ-5D index score was 0.38 and mean VAS score 54.24 presurgery. QALY gain depends on whose weights are used	Data wasn't designed for the purpose of assessing the construct validity of EQ-5D	Not reported
HUI	Total annual costs without hylan Can\$1415, with hylan Can\$2125	HRQoL improved by 0.13 in the AC+H group and by 0.03 in the AC group. AC+H group patients gained 0.071 QALYS compared to AC group patients	The study was openlabel and patients or physicians may have been biased in favour of hylan G-F 20	The ICER was Can\$ 10 000/QALY gained. The cost-utility ratio considered to be below accepted Canadian adoption threshold

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
OTORHINOLARYNGOLOGY					
Cheng et al 2000 USA	Cochlear implant	To determine the QoL and cost consequences for deaf children	78 profoundly deaf children (average age 7.5 years)	Each parent rated his/her child's health state at survey (after an average of 1.9 years of implant use), immediately before and 1 year (TTO and VAS only) before the implantation	Societal, direct and total costs determined
Harris et al 1995 USA	Cochlear implant	To quantitate changes in economic, emotional, and HRQoL after cochlear implant	9 patients between 18 and 60 years	HRQoL assessed preoperatively, and at 6 month, 1, 2, 2.5 and 3 years postoperatively	Direct cochlear implantation costs. In addition patients' increased income taken into account
Joore et al 2003 The Netherlands	Hearing aid	To determine the CE of fitting hearing aids in adult hearing impaired persons compared with that of not fitting	78 patients completed the final measurement	Markov model based on aggregate data and results from a prospective intervention study. HRQoL determined at baseline and 4 month after	Societal. Direct health care costs included in analysis but indirect cost not
Palmer et al 1999 USA	Multichannel cochlear implant	To determine the cost-utility of adult multichannel cochlear implants	62 profoundly hearing-impaired adult recipients and 24 controls	HRQoL assessed at time of enrolment, and at 6 and 12 month post surgery	Direct medical costs including patient-borne costs. Costs and benefits discounted 3%
Wong et al 2000 Hong Kong	Cochlear implantation	To estimate the net costs and net outcomes or benefits of the cochlear implantation program	13 postlingually deafened adults and 22 prelingually deafened children	Cost per QALY was calculated by costing of the cochlear implantation pathway, QoL improvement as measured by 15D. No details on when HRQoL was assessed	Direct medical costs. Service provider's perspective. 6% discount rate, calculation based on a life expectancy of 75 years

QALY Instrument	Costs	Results	Limitations	Cost/QALY
TTO, VAS and HUI	Discounted (3% rate) direct costs were \$60 228. Including indirect costs such as reduced educational expenses, the cochlear implant provided a savings of \$53 198 per child	Mean VAS scores increased by 0.27 to 0.86 at survey. In a subset of participants, TTO scores increased by 0.22 to 0.97 (n = 40) and HUI scores increased by 0.39, to 0.64 (n = 22)	Retrospective assessment of HRQoL before the implantation. HRQoL assessed by parents	Cost per QALY was \$9029 using the TTO, \$7500 using the VAS, and \$5197 using the HUI. Cost-utility of cochlear implant considered acceptable
QWB	Cost of implantation \$22 380 per patient	QWB score improved from the baseline 0.639 to 0.711 3 years after implantation (N.S.). Instead of QALYs authors report well years benefits. Cochlear implant produces 1.85 well years	Noncontrolled study. Small number of patients. The rather low cost of well year partly based on expected increase in patients' future income	Cost per well year \$31 711
EQ-5D	Average cost of fitting hearing aids was €781	Mean improvement in EQ-5D population utility estimate was 0.03 and on EQ-5D VAS 0.02. The changes are small and N.S.		EQ-5D population utility estimate yielded an average cost/QALY estimate of €15 807/QALY for the base-case. The average according to the EQ-5D VAS was €23 745/QALY
HUI	Mean total cost from time of implantation through 12-month follow-up \$36 837	HUI scores in the implant group 0.58 at baseline, 0.76 at 6 month and 0.78 at 12 month, in the control group 0.58 at baseline, 0.57 at 6 month and 0.58 at 12 month	Significant number of control group patients switched over to the intervention group	For patients with a mean 22-year life expectancy, marginal cost per QALY was \$14 670. The cost of QALY was considered favourable
15D		Total 15D score increase for adults: 0.1229, for children: 0.1266. Change in level in hearing: adults: from 5 to 2 children: from 4 to 2	Very limited amount of data in the article. The 15D questionnaire for adults was used for children	Cost/QALY \$17 195 in adults, \$23 474 in children. Cost-utility ratio considered acceptable

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
PULMONARY DISEASES					
Chakravorty et al 2002 United Kingdom	Continuous positive airway pressure (CPAP) therapy	To compare the cost-utility of CPAP therapy to that of conservative lifestyle strategies	A total of 71 sleep clinic patients with a history of snoring and excessive daytime sleepiness randomized to receive CPAP therapy or conservative lifestyle intervention	HRQoL assessment completed at baseline and after the 3-month treatment phase	
Griffiths et al 2001 United Kingdom	Pulmonary rehabilitation	To compare cost-utility of pulmonary rehabilitation to that of standard care	200 patients, mainly with chronic obstructive pulmonary disease, randomized to 6 week rehabilitation or standard medical management	HRQoL measured before randomization, at the end of the 6 week intervention period, and 12 month after entering the study	Direct health care costs plus patient costs incurred in attending the program
Grossman et al 1998 Canada	Treatment with antibiotics	To evaluate the costs and effectiveness of ciprofloxacin compared to standard antibiotic care in treatment of exacerbation of chronic bronchitis	240 chronic bronchitis patients, with a history of frequent exacerbations, presenting with an acute exacerbation	Randomized, multicentre, parallelgroup, openlabel study. HRQoL instruments completed at regular visits 3, 6, 9 and 12 month after enrolment and at follow-up visits after an acute exacerbation	Societal
Mar et al 2003 Spain	Nasal continuous positive airway pressure (nCPAP) treatment	To compare long term CE of nCPAP to conventional null treatment	Markov model based on many parameters derived from literature, no control group, 46 patients completed EQ-5D	HRQoL assessed before treatment and when the patient had used nCPAP for 3 month	Direct health care costs. Sensitivity analysis including discounting

QALY Instrument	Costs	Results	Limitations	Cost/QALY
SG and EQ-5D	No cost data reported	EQ-5D values in CPAP group 0.73 at baseline and 0.77 at 3 month. Respective SG values 0.31 to 0.35. Baseline health status improved by 23% adding 8 QALYs in the CPAP group, compared to a 4% improvement with 4.7 QALYs added in the lifestyle group when measured with SG. EQ-5D showed a much lower magnitude of change in group 1 and no change in group 2	Small groups (32 in CPAP and 21 in lifestyle after exclusions due to health reasons after randomization). A fairly large number of drop-outs especially in the lifestyle intervention group	Not reported
SF-6D	Mean incremental cost of adding rehabilitation to standard care was £ -152 (CI -881 to 577) per patient	Incremental utility of adding rehabilitation was 0.030 QALYs/patient	HRQoL assessed by the SF-36 instrument, answers converted to a single utility score	The point estimate of incremental cost/utility ratio was negative, the probability of the cost /QALY generated being below £0 was 0.64. Rehabilitation produces cost per QALY ratios within bounds considered to be cost-effective
HUI	Total costs (direct and indirect) in the ciprofloxacin treated group Can\$3194, in the usualcare group Can\$2617	A slight, statistically insignificant, improvement in all QoL measures with ciprofloxacin. Mean QALY per person receiving ciprofloxacin was 0.791±0.016 compared to 0.760±0.025 (N.S.) for patients receiving usual care	Treatment compliance not measured. Ciprofloxacin supplied free of charge to patients, usualcare patients not reimbursed. Unblinded study. Clinical differences favouring ciprofloxacin were small and N.S.	Incremental cost per QALY gained for ciprofloxacin Can\$18 588
EQ-5D	Incremental 5-year costs of nCPAP €2664, incremental lifespan costs of nCPAP €7311	Mean HRQoL gain 0.073	Most parameters of the model derived from literature, not measured in the study. No control group	ICER for base case €7861/QALY if time horizon limited to 5 years; €4938/QALY when model computed using the lifespan of the patient

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Ramsey et al 2003 USA	Lung-volume-reduction surgery	To evaluate the effectiveness of lung-volume-reduction surgery compared to medical treatment	1218 patients with severe emphysema at 17 medical centres	Multicentre, randomized controlled trial. Primary outcome measures: mortality and maximal exercise capacity. Secondary outcomes: walk test and lung function test and HRQoL assessed at screening, and after pulmonary rehabilitation (baseline, at 6 and 12 month, and yearly thereafter)	Societal, direct and indirect costs discounted by 3% per year
Schermer et al 2002 The Netherlands	Guided asthma selfmanagement	To investigate whether a family practice-based self-management program for adults with asthma provides an efficient treatment alternative in terms of health care utilization and absence from work, without asthma-control being compromised	19 Dutch family practices (49 family physicians) randomly allocated to guided selfmanagement or usual care (98 selfmanagement and 95 usual care patients)	Randomized multicentre clinical trial. Effectiveness evaluated on the basis of asthma control parameters and QoL. Preference-based utilities were assessed at baseline and half-yearly for 2 years using an interval rating scale	Societal. 3 major cost categories distinguished: program implementation, direct health care, and productivity (indirect) costs
Torrance et al 1999 Canada	Ciprofloxacin	To undertake a 1-year prospective economic evaluation of ciprofloxacin compared with usual antimicrobial care for treatment of acute exacerbations of chronic bronchitis (AECB) in adults presenting with a type I or type II AECB	222 patients, 115 in the ciprofloxacin group and 107 in the usual care group	Randomized, prospective, multicentre CE-trial over 1 year. HRQoL was measured during first AECB, at the end of each AECB and at regular visits 3, 6, 9 and 12 month after enrolment. CE-analysis based on cost per AECB-symptom day averted; CUA based on cost per QALY gained. Cost-benefit analysis based on willingness-to-pay (WTP) data	Societal viewpoint and the viewpoint of a major third-party payer (antibacterials, concomitant medications, outpatient resources, hospitalizations, out-of-pocket expenses for patients and caregivers, time loss from work or usual major activity for working patients and caregivers). No discounting

QALY Instrument	Costs	Results	Limitations	Cost/QALY
QWB	Mean total costs per person at 3 years \$98 952 in the surgery group and \$62 560 in the medical-therapy group	Mean QWB-score before randomization 0.58 in the surgery group and 0.57 in the medical therapy group. After 3 years mean number of QALYs gained higher in the surgery than medical-therapy group (1.46 vs 1.27)	All costs were not included. HRQoL data after intervention not reported in the article	CE ratio for surgery as compared with medical therapy \$190 000 per QALY at 3 and \$53 000 per QALY at 10 years. Surgery considered costly compared to medical therapy
RS	Total costs €1084 for selfmanagement and €1097 for usual care. Mean productivity cost due to limited activity days €213 lower among selfmanagement patients. When all costs were included, selfmanagement was costeffective on all outcomes	Change in rating scale scores reported only in a figure. Selfmanagement patients gained 0.039 QALYs, usual care patients 0.024 QALYs	Disadvantage of using rating scale tend to produce higher quality weights. Study did not randomize individual patients with asthma but family practices	The probability that selfmanagement was cost-effective relative to usual care in terms of QALYs was 52%. Incremental cost per QALY of selfmanagement relative to usual care €13 267
HUI	The incremental annual cost was Can\$578 from the societal viewpoint and Can\$840 for the third-party payer; the CE ratio per AECB-symptom day averted was Can \$209 from the societal viewpoint and Can \$304 for the third-party payer	Mean annual QALYs 0.79 in the ciprofloxacin group, 0.76 in the usual care group	The statistical uncertainty in the results is sufficient that the findings cannot be accepted unequivocally. A further study with a larger sample size would be useful to confirm the findings of this study	Cost/ QALY Can\$18 588 from a societal perspective and Can\$27 025 from ministry of health perspective considered reasonable

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Tousignant et al 1994 Canada	Nasal continuous positive airway pressure (nCPAP)	To assess impact of treatment with nCPAP on HRQoL and QALYs gained	19 patients with obstructive sleep apnea	Utility was measured before treatment (assessed retrospectively) and while on treatment	Service providers perspective. Costs used in analysis reflect the costs of supplies, rental and maintenance associated with the use of equipment and costs of overnight sleep study. Discounting at 5%
van den Boom et al 2001 The Netherlands	Fluticasone propionate versus placebo	To determine CE of early treatment with fluticasone propionate	74 subjects with previously undiagnosed obstructive airway disease randomized to fluticasone or placebo were included in the intention-to-treat analysis	Utilities were measured with SG at baseline, and at 6 and 12 month	Direct medical costs. Also indirect costs determined but not used in base case analysis
PSYCHIATRY					
Gater et al 1998 United Kingdom	Screening and treatment for psychiatric disorder in medical inpatients	To assess whether involvement of a psychiatrist in assessment and care of medical inpatients improves outcome and reduces costs of care	209 patients who scored over the screening threshold for probable psychiatric illness on the General Health Questionnaire randomized to assessment by psychiatrist (n=68), physician informed (n=70), or control group (n=71)	HRQoL assessed with NHP at baseline and at 6-month follow-up and with Rosser Index at 6-month follow-up and retrospectively over the 6 month of follow-up	Direct and indirect costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
SG	Annual costs of treating a patient Can\$2348, cost of an overnight sleep study estimated at Can\$500	Mean utility before treatment 0.63, while on treatment 0.87. A mean of 5.4 QALYs gained by treatment	Very small sample size. Pretreatment utility assessed retrospectively	Cost-utility ratio ranged between Can\$3397 and Can\$9792 per QALY added
SG	In base case analysis incremental cost of early fluticasone intervention \$35 100 per 100 treated subjects. Inclusion of indirect cost resulted in net savings of \$18 600 per 100 treated subjects per year. The incremental cost of early detection and intervention was \$91 500 per 100 treated subjects	Estimated mean difference in utilities at 1 year was 0.046 in favour of fluticasone. Estimated QALY gain of treatment at the end of 1 year follow-up 2.70 QALYs per 100 treated subjects	Conflicting interpretation of results: fluticasone intervention reported to have resulted in significant and relevant improvement in lung function and HRQoL although the difference in utility between the 2 groups was not significant. Small number of subjects, the study might have had insufficient power to detect any relevant improvement	CE \$13 016 per QALY
Rosser-Kind	The physician informed group had a median cost saving of £412, the assessment by psychiatrist group cost an additional £804	No significant differences between the groups at follow-up in the subjective health status. Median QALYs 0.872, 0.922 and 0.922 in assessment by psychiatrist, physician informed, or control groups, respectively	Mixed group of patients in terms of physical and psychiatric diagnoses. Large variation in costs, differences observed may be a consequence of differing employment status	Not reported

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Morgan et al 2003 United Kingdom	Cognitive behaviour therapy (CBT)	To evaluate clinical and cost impact of CBT in insomnia	Pragmatic, randomized, controlled, multicentre trial. 209 patients with chronic sleep problems having used hypnotic drugs randomized (across practices) to receive CBT or no additional treatment	HRQoL assessed at baseline and at 3 and 6 month	NHS perspective. Direct medical costs
Severens et al 2004 The Netherlands	Cognitive behaviour therapy (CBT)	To assess CE of CBT compared to guided support groups (SG), or the "natural course" (NC)	278 chronic fatigue syndrome patients randomly assigned to CBT, SG, or NC, data available from 128 patients at 14 month	HRQoL assessed before randomization and at 8- and 14-month follow-up	Health care perspective and societal perspective
Wilkinson et al 1990 United Kingdom	Patients with long-term psychiatric disorders	To examine the application of the QALYs approach to the field of mental health	14 men and 24 women assessed: 15 with schizophrenia, 13 with affective disorder, and 10 with neurosis	Patients interviewed at entry and 3 times thereafter	Health service perspective
RHEUMATOLOGY					
Hartman et al 2004 The Netherlands	Folate supplementation	To determine CE of folate supplementation in rheumatology patients on methotrexate treatment	Rheumatic arthritis patients started on methotrexate with placebo (n=137), folic acid (n=133), or folinic acid (n=141) (In the economic analysis: placebo=127, folic acid=122, folinic acid=127)	An economic evaluation performed alongside a randomized, double-blind, placebo controlled trial. HRQoL measured at intake, and after 12, 24, 36, and 48 week of follow-up	Medical and non-medical costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
SF-6D	Mean cost of counselling was £154.40 per patient	By 6 month health related utility had increased by 0.024 in the CBT group and decreased by -0.014 in the control group (+0.007 and -0.014 when withdrawals were included)	Significant number of patients refused to participate in the trial. Large number of drop-outs during the trial	At 6 month mean incremental cost per QALY by CBT was £3416 (£4819 including withdrawals). CBT delivery by counsellors is practical, effective and affordable
EQ-5D	Total costs (productivity costs excluded) for the 14-month period €2534, €2597, and €1504 in the CBT, SG, and NC groups, respectively	Mean utility score increased from the baseline 0.4859 to 0.6014 at 14 month in the CBT group, from 0.5036 to 0.5053 in the SG group, and from 0.5257 to 0.5999 in the NC group. Mean QALYs gained at 14 month were, 0.0737, -0.0018, and 0.0458 for CBT, SG and NC groups, respectively	Rather small number of patients: the clinical trial on which the economic analysis is based powered to show an effect on physical activity	Compared to NC, the baseline CE of CBT (including productivity costs) was €21 375 per QALY gained
Rosser-Kind	Overall median cost £185	Mean baseline Rosser-index value in all patients 0.94, after 12 month 0.96	Preliminary study with small sample size	Estimated costs per QALY were lowest in people with schizophrenia (around £6000), higher in those with affective disorders (around £10 000) and highest in those with neurotic disorders (£25 000)
EQ-5D	Mean total costs were comparable in the groups: placebo \$3339, folic acid \$3632, folinic acid \$3296	HRQoL - scores at baseline: 0.49 for placebo, 0.47 for folic acid, 0.52 for folinic acid groups, respectively. QALYs produced during the 48 week treatment: 0.55 for placebo, 0.55 for folic acid, and 0.58 for folinic acid	HRQoL scores after treatment not reported. The uncertainty in differences in total costs and QALYs was substantial	ICER for folic acid could not be calculated, since folic acid and placebo were equally effective. For folinic acid versus placebo the gain in QALY was 0.03 at an average saving of \$43

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Kobelt et al 2004 Sweden	Tumour necrosis factor (TNF) inhibitors	To evaluate costs, benefits and CE of TNF treatment	160 consecutive patients (data available from 116 patients) with rheumatoid arthritis started with etanercept or infliximab	An observational study. CE analysis based on changes in outcome and costs compared to the year before treatment. HRQoL apparently measured (not clearly stated in the article) at baseline and at 12 month	Direct and some indirect medical costs. Patients' out-of-pocket expenses not included
Torrance et al 2004 USA	Adalimumab plus methotrexate	To compare HRQoL in rheumatoid arthritis (RA) patients treated with adalimumab (a human anti-tumour necrosis factor monoclonal antibody) plus methotrexate or placebo plus methotrexate	HRQoL data obtained from 2 separate randomized trials: Armada with 121 patients and DE019 with 378 patients	HRQoL data obtained at baseline, at the end of the study and at 2 time points in between	
van den Hout et al 2003 The Netherlands	Multidisciplinary care in patients with rheumatoid arthritis (RA)	To assess the relative CE of clinical nurse specialist care, inpatient team care and day patient team care	Patients with RA; 71 in the clinical nurse specialist group, 71 in the inpatient team care group, and 68 in the day patient team care group	CUA alongside a prospective randomized controlled trial. HRQoL assessed at 0, 6, 12, 26, 52, and 104 week	Societal, both direct and indirect costs included. Costs discounted at 3% a year
van Tubergen et al 2002 The Netherlands	Combined spaexercise therapy in addition to standard treatment	To evaluate CE and cost-utility of a 3-week course of combined spa therapy and exercise therapy	111 ankylosing spondylitis patients randomized to group 1 (spa treatment in Bad Hofgastein), group 2 (spa treatment in Arcen) and group 3 (controls)	Patients completed EQ-5D at baseline, and 4, 16, 28, and 40 week after the start of spaexercise therapy	Societal. Direct (health care and non-health care) and indirect (non-health care) costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	Incremental cost of TNF inhibitor treatment €12 200	EQ-5D improved from the baseline 0.28 to 0.65 at 12 month	Calculations and conclusions based on a number of assumptions. No control group	Cost/QALY gained estimated to range between €36 900 and €53 600 during the first year, depending on the assumptions
HUI	No cost data reported	HUI3 mean changes from baseline scores for adalimumab-treated patients were 0.22 and 0.21 in the 2 trials, whereas placebo patients' changes were 0.04 and 0.07. The rate of QALYs gained per year in the treatment group compared with the placebo group were 0.145 in the ARMADA trial and 0.104 in the DE019 trial	2 separate studies. The analysis was performed on a defined study population, patients had more severe disease than average RA patients. Both trials were of limited duration	Not reported
SF-6D, TTO, RS (TRS)	Costs of initial treatment estimated at €200 for clinical nurse specialist, €5000 for inpatient team, and €4100 for day patient team care. Societal costs per patient significantly lower for clinical nurse specialist, by €10 876 compared with inpatient and by €5324 compared with day patient care	Over the 2 year follow-up period, patients in all 3 groups improved. Average improvements on the SF-6D, TRS, and TTO-instruments were 0.045, 0.061, and 0.046. QALY gains based on the 3 instruments ranged from 1.202 to 1.624. No significant differences between the 3 types of care	Imbalance between the groups at randomization regarding several baseline variables	Not reported. Preferred treatment from a health-economic perspective is care provided by clinical nurse specialist
EQ-5D	Mean total costs were \$3023 for group 1, \$3240 for group 2 and \$1754 for group 3	The between-group difference for EQ-5D utility-AUC was 0.17 for group 1 versus controls and 0.08 for group 2 versus controls		The cost per QALY was \$7465 for group 1 and \$18 575 for group 2. Cost-utility ratios considered favorable compared to standard treatment

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Verhoeven et al 1998 The Netherlands	Combined step-down prednisolone, methotrexate and sulphasalazine, compared to sulphasalazine alone	To assess CE and cost-utility of early intervention with the combined treatment in rheumatoid arthritis (RA) patients	76 patients in the combined treatment group and 78 patients in the sulphasalazine group	Multi-centre 56 week randomized double-blind trial with full economic analysis of direct costs and utility analysis. HRQoL assessed at baseline and at 28 and 56 week	Societal. Only direct costs included (intervention, non-protocol drugs, other outpatient care, inpatient care and direct non-medical costs)
TRANSPLANTATION SURGERY					
Greiner et al 2001 Germany	Kidney transplantation	To assess the costs, HRQoL and CE of kidney transplantation	150 patients subsequently undergoing transplantation	Prospective, socioeconomic evaluation study. HRQoL assessed while on waiting list and 14 days, 1, 3, 6, and 12 month after transplantation	Societal. Direct and indirect costs considered. Discounted at 5%
Laupacis et al 1996 Canada	Renal transplantation	To assess the cost-utility of renal transplantation compared with dialysis	A prospective cohort of 168 pre-transplant patients followed for up to 2 years after transplantation	Patients interviewed every 6 month prior to and at 1, 3, 6, 12, 18, and 24 month after transplantation. TTO values were used for the production of QALYs. TTO values were available for 136 patients after transplantation, 76 of whom were followed for 2 years or until death	Societal (including inpatient hospitalizations, outpatient visits, nephrectomy of the living related donor, the transplant program and patient-borne costs)

QALY Instrument	Costs	Results	Limitations	Cost/QALY
RS, SG	Mean total costs per patient in the first week of follow-up \$5519 for combined treatment and \$6511 for treatment with sulphasalazine alone	At week 28, rating scale scores had increased by 0.24 in the combined-treatment and 0.15 in the sulphasalazine group, and SG utilities by 0.10 and 0.06, respectively. At week 56, most of the between-group contrast in utilities had been lost. Gain in QALYs 0.06 assessed by rating scale and 0,02 assessed by SG	Trial was primarily designed to study clinical benefits and thus probably underpowered to prove cost-benefits	Combined treatment considered cost-effective due to enhanced efficacy at lower or equal costs
EQ-5D	Direct cost for dialysis DM55 200/year, cost for kidney transplant (excluding medication and indirect costs) DM59 980. The immunosuppressive drugs and laboratory tests add DM14 000/year	Patients achieved a higher HRQoL as a result of kidney transplant compared to dialysis treatment. Overall 0.76 QALYs (discounted) per patient were gained as a result of kidney transplant	Resource consumption for inpatient treatment completely assessed for only 77 patients. Great variation in the number of questionnaires completed at different time-points. Only 58 completed questionnaires one year after	Cost-utility value for dialysis was calculated to be DM147 800/QALY and for kidney transplant DM38 300/QALY
TTO	The average annual cost of dialysis was \$66 782. The average transplantation costs were \$66 290 for the first year and \$27 875 for the second year	The mean TTO score improved from 0.57 pretransplant to 0.70 2 years after transplantation.	Because TTO data were not available on all patients, extrapolations were made for some values. In order to compare the incremental CE of transplantation with dialysis, some estimates of the utilities and costs of dialysis, if the patients had not been transplanted, had to be made. It was assumed that the QoL on dialysis would have remained the same over the 2 years	For patients who progressed from dialysis to transplantation the cost/QALY was -\$59 325 for the first and -\$394 500 for the second year. Transplantation thus more effective and less expensive than dialysis

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
Longworth et al 2003 United Kingdom	Liver transplant	Economic evaluation of adult liver transplantation	347 patients: 122 with primary biliary cirrhosis (PBC), 155 with alcoholic liver disease (ALD), and 70 with primary sclerosing cholangitis (PSC). Of the patients 81, 82 and 45, for each class respectively, were transplanted	HRQoL assessed before transplantation at the time of listing, then at 3-month intervals until transplantation, and posttransplantation at 3, 6, 12, and 24 month	That of the NHS transplant centres. Incremental CE analysis. Costs discounted at 6% and QALYs at 1.5%
Polsky et al 2001 USA	Basilixi-mab	Whether basilixi-mab demonstrates economic and QoL advantages over other induction therapies	Renal transplant recipients randomized to receive basiliximab (n=70) or antithymocyte (ATG) globulin (n=65) as part of quadruple immunosuppressive regimen	Patients completed the (VAS) from EQ-5D at base-line, day 4, and week 1, 2, 4, 12, 24 and 52	Direct medical costs
TenVergert et al 1998 and van Enckevort et al 1998 (two reports) The Netherlands	Lung transplantation	To assess effectiveness and CE of lung transplantation	157 waiting list patients of whom 76 reached the transplantation phase	Effect of lung transplantation on survival assessed by means of a Cox regression model. In the QoL study a longitudinal analysis (data from 24 patients) was applied. HRQoL assessed before and 1, 4, 7, 13, and 19 month post transplantation	Direct medical and indirect costs

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	Incremental costs measured over 27 month from date of listing. Cost (including assessment), were £15 224 for PBC, £25 712 for ALD and £12 182 for PSC, respectively	Incremental QALYs measured over 27 month from date of listing were 0.54 for PBC, 0.55 for ALD and 0.58 for PSC, respectively	Costs of QALYs gained estimated only for a short time period. Data on survival and HRQoL of non-transplanted patients rests on prognostic models and assumptions. Proportion of patients with usable EQ-5D values at main time points ranged from 48% to 76%	Mean incremental cost per QALY from time of listing to 27 month for patients with PBC, ALD, and PSC were £29 000, £48 000 and £21 000, respectively. Using benchmark figure of £30 000 for what NHS can pay for additional QALY, transplantation for ALD is not cost-effective
EQ-5D VAS	First year treatment costs: basiliximab \$45 857, ATG \$54 729	Mean EQ-5D VAS scores 68.0 and 66.9 at baseline and 82.4 and 83.8 at week 52 in the ATG and basiliximab groups, respectively. Differences between groups N.S. One-year quality-adjusted survival 81.5 in the basiliximab and 81.1 in the ATG groups, respectively	Inpatient billing data in the multicentre trial collected only from 1 treatment centre but applied to all patients. QALY calculation not valid as based on EQ-5D VAS only	Not reported
EQ-5D	Cost of transplantation Hfl670 000 per patient	Average EQ-5D score 0.40 while on the waiting list, 0.70 soon after transplantation and near 0.90 2 years after transplantation. Average QALY gain as a consequence of lung transplantation estimated at 5.2		The average cost per QALY gained estimated at Hfl120 000. CE considered unfavourable compared with other Dutch transplant programmes

Study, Country of origin	Intervention	Aim of the study	Subjects	Methods	Perspective
UROLOGY					
Noble et al 2002 United Kingdom	Non-contact laser therapy	To compare CE of noncontact laser therapy with transurethral prostate resection and conservative management	Pragmatic, multi-centre, randomized trial. 108, 109 and 106 patients in the laser, resection, and conservative groups, respectively	HRQoL assessed at baseline and at 7.5-month follow-up	Medical costs from the service provider (NHS) and patient perspective including travel and out-of-pocket expenses
Reed et al 2004 Multinational	Zoledronic acid	To estimate CE of zoledronic acid	643 men with advanced stage prostate cancer randomized to receive zoledronic acid (n=214) or placebo (n=208) for 15 month	Double-blind, placebo controlled, randomized trial. HRQoL assessed at baseline at and at every 3 month	Direct medical costs, no discounting

QALY Instrument	Costs	Results	Limitations	Cost/QALY
EQ-5D	Resection £883 and laser treatment £1177 more costly than conservative treatment	Mean EQ-5D-scores at baseline 0.774, 0.772 and 0.773 and at the follow-up 0.790, 0.816 and 0.772 for the laser, resection and conservative groups, respectively. Compared to conservative treatment resection produced 0.01 and laser treatment 0.03 QA-LYs, respectively	QALY calculations based on certain assumptions. Limitations due to short follow-up and missing data	ICER £76 779 /QALY for resection and £38 291/ QALY for laser treatment, respectively
EQ-5D	Mean direct costs excluding study medication \$5365 for zoledronic acid and \$5689 for control patients. Average cost of zoledronic acid \$5677	HRQoL scores for the individual study groups not reported. Average incremental gain in quality adjusted time approximately 2 week for patients receiving zoledronic acid	Resource use data were not collected for almost 17% of patients and 9 to 16% of EQ-5D measures were missing	Nominal ICER estimated at \$159 200 i.e. higher than thresholds used by many decision makers to confer CE

ARTIKKELIT, JOITA SITEERATTU LIITTEESSÄ 2/ ARTICLES CITED IN APPENDIX 2

Campbell HE, Tait S, Buxton MJ, Sharples LD, Caine N, Schofield PM, Wallwork J, Mueller X, Bodor E, Bachet J, Large S, Svennevig J. A UK trial-based cost-utility analysis of transmyocardial laser revascularization compared to continued medical therapy for treatment of refractory angina pectoris. *Eur J Cardiothorac Surg* 2001;20:312–318.

Chakravorty I, Cayton RM, Szczepura A. Health utilities in evaluating intervention in the sleep apnoea/hypopnoea syndrome. *Eur Respir J* 2002;20:1233–8.

Cheng AK, Rubin HR, Powe NR, Mellon NK, Francis HW, Niparko JK. Cost-utility analysis of the cochlear implant in children. *JAMA* 2000;284:850–6.

Cunningham SJ, Sculpher M, Sassi F, Manca A. A cost-utility analysis of patients undergoing orthognathic treatment for the management of dentofacial disharmony. *Br J Oral Maxillofac Surg* 2003;41:32–5.

Dolan P, Torgerson D, Kakarlapudi TK. Health-related quality of life of Colles' fracture patients. *Osteoporos Int* 1999;9:196–9.

Eefting F, Nathoe H, van Dijk D, Jansen E, Lahpor J, Stella P, Suyker W, Diephuis J, Suryapranata H, Ernst S, Borst C, Buskens E, Grobbee D, de Jaegere P. Randomized comparison between stenting and off-pump bypass surgery in patients referred for angioplasty. *Circulation* 2003;108:2870–6.

Garry R, Fountain J, Brown J, Manca A, Mason S, Sculpher M, Napp V, Bridgman S, Gray J, Lilford R. EVALUATE hysterectomy trial: a multicentre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy. *Health Technol Assess* 2004;8:1–154.

Gater RA, Goldberg DP, Evanson JM, Lowson K, McGrath G, Tantam D, Million L. Detection and treatment of psychiatric illness in a general medical ward: a modified cost-benefit analysis. *J Psychosom Res* 1998;45:5:437–448.

Gilbert FJ, Grant AM, Gillan MG, Vale LD, Campbell MK, Scott NW, Knight DJ, Wardlaw D. Low back pain: influence of early MR imaging or CT on treatment and outcome—multicenter randomized trial. *Radiology* 2004;231:343–51.

Greiner W, Obermann K, Schulenburg JM. Socio-economic evaluation of kidney transplantation in Germany. *Archives of Hellenic Medicine* 2001;18:147–155.

Griffiths TL, Phillips CJ, Davies S, Burr ML, Campbell IA. Cost effectiveness of an outpatient multidisciplinary pulmonary rehabilitation programme. *Thorax* 2001;56:779–84.

Grossman R, Mukherjee J, Vaughan D, Eastwood C, Cook R, LaForge J, Lampron N. A 1-year community-based health economic study of ciprofloxacin vs usual antibiotic treatment in acute exacerbations of chronic bronchitis: the Canadian Ciprofloxacin Health Economic Study Group. *Chest* 1998;113:131–41.

Gudex CM, Hawthorne MR, Butler AG, Duffey PO. Measuring patient benefit from botulinum toxin in the treatment of dystonia. Feasibility of cost-utility analysis. *Pharmacoeconomics* 1997;12:675–84.

Hamel MB, Phillips RS, Davis RB, Teno J, Connors AF, Desbiens N, Lynn J, Dawson NV, Fulkerson W, Tsevat J. Outcomes and cost-effectiveness of ventilator support and aggressive care for patients with acute respiratory failure due to pneumonia or acute respiratory distress syndrome. *Am J Med* 2000;109:614–20.

Harris JP, Anderson JP, Novak R. An outcomes study of cochlear implants in deaf patients: Audiologic, economic, and quality-of-life changes. *Arch Otolaryngol Head Neck Surg* 1995;121:4:398–404.

Hartman M, van Ede A, Severens JL, Laan RF, van de Putte L, van der Wilt GJ. Economic evaluation of folate supplementation during methotrexate treatment in rheumatoid arthritis. *J Rheumatol* 2004;31:902–8.

James M, St Leger S, Rowsell KV. Prioritising elective care: a cost utility analysis of orthopaedics in the north west of England. *J Epidemiol Community Health* 1996;50:182–9.

Joore MA, van der Stel H, Peters HJ, Boas GM, Anteunis LJ. The cost-effectiveness of hearing-aid fitting in the Netherlands. *Arch Otolaryngol Head Neck Surg* 2003;129:297–304.

Kaplan RM, Atkins CJ, Wilson DK. The cost-utility of diet and exercise interventions in non-insulin-dependent diabetes mellitus. *Health Promot* 1988;2:5:331–340.

Kemler MA, Furnee CA. Economic evaluation of spinal cord stimulation for chronic reflex sympathetic dystrophy. *Neurology* 2002;59:1203–9.

Kennedy A, Nelson E, Reeves D, Richardson G, Roberts C, Robinson A, Rogers A, Sculpher M, Thompson D. A randomised controlled trial to assess the impact of a package comprising a patient-orientated, evidence-based self-help guidebook and patient-centred consultations on disease management and satisfaction in inflammatory bowel disease. *Health Technol Assess.* 2003;7:iii:1–113.

Kobelt G, Eberhardt K, Geborek P. TNF inhibitors in the treatment of rheumatoid arthritis in clinical practice: costs and outcomes in a follow up study of patients with RA treated with etanercept or infliximab in southern Sweden. *Ann Rheum Dis* 2004;63:4–10.

Kristiansen IS, Falch JA, Andersen L, Aursnes I. Bruk av alendronat ved osteoporose er det kostnadseffektive? [Use of alendronate in osteoporosis-is it cost-effective?] *Tidsskr Nor Laegeforen* 1997;117:2619–22.

Kroeker A, Clark WF, Heidenheim AP, Kuenzig L, Leitch R, Meyette M, Muirhead N, Ryan H, Welch R, White S, Lindsay RM. An operating cost comparison between conventional and home quotidian hemodialysis. *Am J Kidney Dis* 2003;42:49–55.

Laupacis A, Bourne R, Rorabeck C, Feeny D, Wong C, Tugwell P, Leslie K, Bullas R. Costs of elective total hip arthroplasty during the first year. Cemented versus noncemented. *J Arthroplasty* 1994;9:481–7.

Laupacis A, Keown P, Pus N, Krueger H, Ferguson B, Wong C, Muirhead N. A study of the quality of life and cost-utility of renal transplantation. *Kidney Int* 1996;50:235–242.

Lavernia CJ, Guzman JF, Gachupin-Garcia A. Cost effectiveness and quality of life in knee arthroplasty. *Clin Orthop* 1997;1:345:134–9.

Longworth L, Young T, Buxton MJ, Ratcliffe J, Neuberger J, Burroughs A, Bryan S. Midterm cost-effectiveness of the liver transplantation program of England and Wales for three disease groups. *Liver Transpl* 2003;9:1295–307.

Manca A, Sculpher MJ, Ward K, Hilton P. A cost-utility analysis of tension-free vaginal tape versus colposuspension for primary urodynamic stress incontinence. *BJOG* 2003;110:255–62.

Mar J, Rueda JR, Duran-Cantolla J, Schechter C, Chilcott J. The cost-effectiveness of nCPAP treatment in patients with moderate-to-severe obstructive sleep apnoea. *Eur Respir J* 2003;21:515–22.

Marti-Valls J, Alonso J, Lamarca R, Pinto JL, Auleda J, Girvent R, Lopez E, Miralles R, Ramon R, Salo JM. Efectividad y costes de la intervención de prótesis total de cadera en siete hospitales de Cataluña [Effectiveness and cost of total hip replacement in seven hospitals in Catalonia, Spain]. *Med Clin (Barc)* 2000;114:34–9.

McCrone P, Chisholm D, Knapp M, Hughes R, Comi G, Dalakas MC, Illa I, Kilindireas C, Nobile-Orazio E, Swan A, van den Bergh P, Willison HJ. Cost-utility analysis of intravenous immunoglobulin and prednisolone for chronic inflammatory demyelinating polyradiculoneuropathy. *Eur J Neurol* 2003;10:687–94.

Mitton CR, Rose MS, Koshman MI, Sheldon RS. A cost-utility analysis of pacemakers for the treatment of neurally mediated syncope. *Am J Cardiol* 1999;84:1356–1359.

Morgan K, Dixon S, Mathers N, Thompson J, Tomeny M. Psychological treatment for insomnia in the management of long-term hypnotic drug use: a pragmatic randomised controlled trial. *Br J Gen Pract* 2003;53:923–8.

Nathoe HM, van Dijk D, Jansen EW, Suyker WJ, Diephuis JC, van Boven WJ, de la Riviere AB, Borst C, Kalkman CJ, Grobbee DE, Buskens E, de Jaegere PP. A comparison of on-pump and off-pump coronary bypass surgery in low-risk patients. *N Engl J Med* 2003;348:394–402.

Noble SM, Coast J, Brookes S, Neal DE, Abrams P, Peters TJ, Donovan JL. Transurethral prostate resection, noncontact laser therapy or conservative management in men with symptoms of benign prostatic enlargement? An economic evaluation. *J Urol* 2002;168:2476–82.

O'Brien BJ, Goeree R, Blackhouse G, Smieja M, Loeb M. Oseltamivir for treatment of influenza in healthy adults: pooled trial evidence and cost-effectiveness model for Canada. *Value Health* 2003;6:116–25.

Oldridge N, Furlong W, Feeny D, Torrance G, Guyatt G, Crowe J, Jones N. Economic evaluation of cardiac rehabilitation soon after acute myocardial infarction. *Am J Cardiol* 1993;72:154–61.

- Ostendorf M, Buskens E, van Stel H, Schrijvers A, Marting L, Dhert W, Verbout A. Waiting for total hip arthroplasty: Avoidable loss in quality time and preventable deterioration. *J Arthroplasty* 2004;19:302–309.
- Palmer CS, Niparko JK, Wyatt JR, Rothman M, de Lissovoy G. A prospective study of the cost-utility of the multichannel cochlear implant. *Arch Otolaryngol Head Neck Surg* 1999;125:1221–8.
- Parkin D, Jacoby A, McNamee P, Miller P, Thomas S, Bates D. Treatment of multiple sclerosis with interferon beta: an appraisal of cost-effectiveness and quality of life. *J Neurol Neurosurg Psychiatry* 2000;68:144–9.
- Patrick DL, Ramsey SD, Spencer AC, Kinne S, Belza B, Topolski TD. Economic evaluation of aquatic exercise for persons with osteoarthritis. *Med Care* 2001;39:413–24.
- Polsky D, Weinfurt KP, Kaplan B, Kim J, Fastenau J, Schulman KA. An economic and quality-of-life assessment of basiliximab vs antithymocyte globulin immunoprophylaxis in renal transplantation. *Nephrol Dial Transplant* 2001;16:1028–1033.
- Ramsey SD, Berry K, Etzioni R, Kaplan RM, Sullivan SD, Wood DE. Cost effectiveness of lung-volume-reduction surgery for patients with severe emphysema. *N Engl J Med* 2003;348:2092–102.
- Rawles J, Light J. Loss of quality adjusted days as a trial endpoint: effect of early thrombolytic treatment in suspected myocardial infarction. Grampian Region Early Anistreplase Trial (GREAT). *J Epidemiol Community Health* 1993;47:377–81.
- Reed SD, Radeva JI, Glendenning GA, Saad F, Schulman KA. Cost-effectiveness of zoledronic acid for the prevention of skeletal complications in patients with prostate cancer. *J Urol* 2004;171:1537–42.
- Remak E, Hutton J, Selai CE, Trimble MR, Price M J. A cost-utility analysis of adjunctive treatment with newer antiepileptic drugs in the UK. *Journal of Drug Assessment* 2004;7:109–20.
- Ridley S, Biggam M, Stone P. A cost-utility analysis of intensive therapy. II: Quality of life in survivors. *Anaesthesia* 1994;49:192–6.
- Rome K, Gray J, Stewart F, Hannant SC, Callaghan D, Hubble J. Evaluating the clinical effectiveness and cost-effectiveness of foot orthoses in the treatment of plantar heel pain: a feasibility study. *J Am Podiatr Med Assoc* 2004;94:229–238.
- Schermer TR, Thoonen BP, van den Boom G, Akkermans RP, Grol RP, Folgering HT, van Weel C, van Schayck CP. Randomized controlled economic evaluation of asthma self-management in primary health care. *Am J Respir Crit Care Med* 2002;166:1062–72.
- Schulman KA, Buxton M, Glick H, Sculpher M, Guzman G, Kong J, Backhouse M, Mauskopf J, Bell L, Eisenberg JM. Results of the economic evaluation of the first study. A multinational prospective economic evaluation. *FIRST Investigators*.

Flolan International Randomized Survival Trial. *Int J Technol Assess Health Care* 1996;12:698–713.

Sculpher M, Manca A, Abbott J, Fountain J, Mason S, Garry R. Cost effectiveness analysis of laparoscopic hysterectomy compared with standard hysterectomy: results from a randomised trial. *BMJ* 2004;328:134.

Severens JL, Prins JB, van der Wilt GJ, van der Meer JW, Bleijenberg G. Cost-effectiveness of cognitive behaviour therapy for patients with chronic fatigue syndrome. *Q J Med* 2004;97:153–61.

Spady BL. The validity of EuroQol scores of patients with hip or knee replacement. Doctoral Dissertation, University of Alberta, Canada 2000.

TenVergert EM, van Enkevort PJ, Geertsma A, Koopmanschap MA, van der Bij W, de Boer WJ. Evaluatie van het Groningse longtransplantatieprogramma: overleving, kwaliteit van leven en kosteffectiviteit.[The evaluation of the lung transplantation program at Groningen: survival rate, quality of life, and cost-effectiveness]. *Ned Tijdschr Geneesk* 1998;142:957–62.

Torrance G, Walker V, Grossman R, Mukherjee J, Vaughan D, La Forge J, Lampron N. Economic evaluation of ciprofloxacin compared with usual antibacterial care for the treatment of acute exacerbations of chronic bronchitis in patients followed for 1 year. *Pharmacoeconomics* 1999;16:499–520.

Torrance GW, Raynauld JP, Walker V, Goldsmith CH, Bellamy N, Band PA, Schultz M, Tugwell P. A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for patients with knee osteoarthritis (Part 2 of 2): economic results. *Osteoarthritis Cartilage* 2002;10:518–27.

Torrance GW, Tugwell P, Amorosi S, Chartash E, Sengupta N. Improvement in health utility among patients with rheumatoid arthritis treated with adalimumab (a human anti-TNF monoclonal antibody) plus methotrexate. *Rheumatology (Oxford)* 2004;43:712–8.

Tousignant P, Cosio MG, Levy RD, Groome PA. Quality adjusted life years added by treatment of obstructive sleep apnea. *Sleep* 1994;17:52–60.

Uyl-de Groot CA, Löwenberg B, Vellenga E, Suciú S, Willemze R, Rutten FFH. Cost-effectiveness and quality-of-life assessment of GM-CSF as an adjunct to intensive remission induction chemotherapy in elderly patients with acute myeloid leukaemia. *Br J Haematol* 1998;100:4:629–636.

van Agthoven M, Segeren CM, Buijt I, Uyl-De Groot CA, van der Holt B, Lokhorst HM, Sonneveld P. A cost-utility analysis comparing intensive chemotherapy alone to intensive chemotherapy followed by myeloablative chemotherapy with autologous stem-cell rescue in newly diagnosed patients with stage II/III multiple myeloma; a prospective randomised phase III study. *Eur J Cancer* 2004;40:1159–69.

van den Boom G, Rutten-van Molken MP, Molema J, Tirimanna PR, van Weel C, van Schayck C.P. The cost effectiveness of early treatment with fluticasone

propionate 250 microg twice a day in subjects with obstructive airway disease. Results of the DIMCA program. *Am J Respir Crit Care Med* 2001;164:2057-66.

van den Brink M, van den Hout WB, Stiggelbout AM, Klein Kranenbarg E, Marijnen CA, van de Velde CJ, Kievit J. Cost-utility analysis of preoperative radiotherapy in patients with rectal cancer undergoing total mesorectal excision: a study of the Dutch Colorectal Cancer Group. *J Clin Oncol* 2004;22:244-53.

van den Hout WB, Tijhuis GJ, Hazes JM, Breedveld FC, Vliet Vlieland TP. Cost effectiveness and cost utility analysis of multidisciplinary care in patients with rheumatoid arthritis: a randomised comparison of clinical nurse specialist care, inpatient team care, and day patient team care. *Ann Rheum Dis* 2003; 62:308-15.

van den Hout WB, van der Linden YM, Steenland E, Wiggeraad RG, Kievit J, de Haes H, Leer JW. Single- versus multiple-fraction radiotherapy in patients with painful bone metastases: cost-utility analysis based on a randomized trial. *J Natl Cancer Inst* 2003;95:222-9.

van Enckevort PJ, TenVergert EM, Bonsel GJ, Geertsma A, van der Bij W, de Boer WJ, Koopmanschap MA, Al MJ, Rutten FFH. Technology assessment of the Dutch Lung Transplantation Program. *Int J Technol Assess Health Care* 1998;14:344-56.

van Gemert WG, Adang EM, Kop M, Vos G, Greve JW, Soeters PB. A prospective cost-effectiveness analysis of vertical banded gastroplasty for the treatment of morbid obesity. *Obes Surg* 1999;9:484-91.

van Tubergen A, Boonen A, Landewe R, Rutten-van Molken M, van der Heijde D, Hidding A, van der Linden S. Cost effectiveness of combined spa-exercise therapy in ankylosing spondylitis: a randomized controlled trial. *Arthritis Rheum* 2002;47:459-67.

Verhoeven AC, Bibo JC, Boers M, Engel GL, van der Linden S. Cost-effectiveness and cost-utility of combination therapy in early rheumatoid arthritis: randomized comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone. *Br J Rheumatol* 1998;37:1102-1109.

Wilkinson G, Croft-Jeffreys C, Krekorian H, McLees S, Falloon I. QALYs in psychiatric care? *Psychiatr Bull R Coll Psychiatr* 1990;14:582-585.

Wonderling D, Vickers AJ, Grieve R, McCarney R. Cost effectiveness analysis of a randomised trial of acupuncture for chronic headache in primary care. *BMJ* 2004;328:747.

Wong BYK, Hui Y, Au D, Wei W. Economic evaluation of cochlear implantation. *Adv Otorhinolaryngol* 2000;57:377-81.

SIDONNAISUUDET/CONFLICT OF INTERESTS

Eija Roine	Ei sidonnaisuuksia.
Risto P. Roine	Ei sidonnaisuuksia.
Olli-Pekka Ryytänen	Wisane Oy:n hallituksen jäsen.
Pirjo Räsänen	Ei sidonnaisuuksia.
Virpi Semberg-Konttinen	Ei sidonnaisuuksia.
Harri Sintonen	30 prosenttia yliopiston palkasta lahjoituksena Lääketeollisuudelta. Toimii useiden lääkeyritysten ja lääkeyritysten järjestöjen (Lääketeollisuus ry ja Lääketietokeskus) koulutuksissa ja symposiumeissa luennoitsijana. Osallistunut yhdelle Oy Eli Lilly Finland Ab:n rahoittamalle kongressimatkalle. Mukana useiden eri lääkeyhtiöiden tutkimushankkeissa. Suomen MSD Oy:n tieteellisen neuvottelukunnan jäsen. HERCO Ltd:n osakas ja hallituksen puheenjohtaja. Tytär työskentelee hinta- ja korvattavuusasioissa Pfizer Oy:ssä.

Eija Roine	None declared.
Risto P. Roine	None declared.
Olli-Pekka Ryytänen	Board member, Wisane Inc.
Pirjo Räsänen	None declared.
Virpi Semberg-Konttinen	None declared.
Harri Sintonen	30 per cent of university salary donated by medical industry. Has acted as a speaker in several educational symposia organized by medical industry. Has participated in one congress trip sponsored by Eli Lilly, Finland. Involved in several research projects supported by medical industry. Member of the Scientific Advisory Board of MSD, Finland. Shareholder and chairman of HERCO Ltd. Daughter works with pricing- and reimbursement matters in Pfizer, Finland.