

Assessing the order of magnitude of outcomes in single-arm cohorts through systematic comparison with corresponding cohorts: An example from the AMOS study

Harald J Hamre^{1§}, Anja Glockmann¹, Wilfried Tröger², Gunver S Kienle¹, Helmut Kiene¹

¹Institute for Applied Epistemology and Medical Methodology, Freiburg, Germany

²Director, Clinical Research Dr Tröger, Freiburg, Germany

[§]Corresponding author

Email addresses:

HJH: harald.hamre@ifaemm.de

AG: anja.glockmann@ifaemm.de

WT: troeger@crdt.de

GSK: gunver.kienle@ifaemm.de

HK: helmut.kiene@ifaemm.de

Abstract

Background

When a therapy has been evaluated in the first clinical study, the outcome is often compared to outcomes in corresponding cohorts receiving other treatments. Such comparisons are often limited to selected studies, and often mix different outcomes and follow-up periods. We here give an example of a systematic comparison to all cohorts with identical outcomes and follow-up periods.

Methods

The therapy to be compared (anthroposophic medicine, a complementary therapy system) had been evaluated in one single-arm cohort study: the Anthroposophic Medicine Outcomes Study (AMOS). The five largest AMOS diagnosis groups (A-cohorts: asthma, depression, low back pain, migraine, neck pain) were compared to all retrievable corresponding cohorts receiving other therapies (C-cohorts) with SF-36 scales or summary measures evaluable after three, six or 12-months. Between-group differences (pre-post difference in A-cohort minus pre-post difference in respective C-cohort) were divided with the standard deviation (SD) of the baseline score of the A-cohort.

Results

A-cohorts (n = 392 patients) were similar to C-cohorts (n = 16,432 patients) regarding age, disease duration, baseline affection and follow-up rates. A-cohorts had ≥ 0.50 SD larger improvements than C-cohorts in 13.5% (70/517) of comparisons, small/minimal differences (-0.49 to 0.49 SD) were found in 79.7% of comparisons, and C-cohorts had ≥ 0.50 SD larger improvements than A-cohorts in 6.8% of comparisons.

Conclusions

The results suggest that anthroposophic therapy can be associated with SF-36 improvements largely of the same order of magnitude as improvements following other treatments. The analysis also demonstrates the value of a systematic approach when comparing a therapy cohort to corresponding therapy cohorts.

Background

In the early phase of the clinical evaluation of a therapy, when first study results are published, it can be desirable to assess therapy outcomes, relative to outcomes of other treatments for the disease in question. At this stage, a systematic review of controlled studies of the therapy in question vs. other treatments will give limited information, because few such studies will be available. An alternative is to compare the therapy cohort (or cohorts) to all corresponding cohorts, i. e. to single-arm cohorts and therapy arms in controlled studies, receiving other treatments. Although such 'all corresponding cohorts comparisons' cannot assess comparative effectiveness, they nevertheless yield information about the order of magnitude of treatment outcomes. For therapies which have been evaluated exclusively in single-arm studies (e. g. many drugs [1-4], surgery [5-8], other procedures [9,10]), 'corresponding cohorts comparisons' remain the only possibility.

Brief comparisons with corresponding cohorts are often presented in discussion sections of papers (e. g. [11,12]), and are often unsystematic (limited to selected studies) and unnecessarily imprecise (mixing different outcomes and follow-up periods). We here give an example of a systematic comparative review, restricted to cohorts with identical outcomes and comparable follow-up periods. Since the cohorts compared are derived from different studies, the comparisons are necessarily explorative and the analyses descriptive: results are not pooled but ordered in increasing magnitude and presented graphically.

Methods

Occasion for the comparison

The therapy to be compared (anthroposophic medicine, a physician-provided complementary therapy system including counselling, medication, art and movement exercises, and massage) had been evaluated in a large single-arm cohort study: the Anthroposophic Medicine Outcomes Study (AMOS) [13]. AMOS was conducted 1998-2005 in 155 German medical practices. Patients with chronic disorders were enrolled before starting anthroposophic therapy and followed up for four years [14-17].

For the five largest AMOS diagnosis groups in adult patients (asthma, depression [18], low back pain, migraine, and neck pain) AMOS is so far the only outpatient study of anthroposophic therapy for the respective diagnosis [19]. In all five groups, changes in health status had been evaluated with the SF-36 Health Survey, which is widely used [20], enabling comparison to other cohorts. We conducted a systematic review, comparing these five cohorts (A-cohorts, patients enrolled up to 31 December 2004) to all retrievable patient cohorts (C-cohorts) with corresponding diagnosis, outcome measure (SF-36), and follow-up periods.

Objective

The objective of this systematic review was to assess the comparative order of magnitude of pre-post changes of health status in adult patients receiving anthroposophic therapy for one of five chronic diseases.

Eligible comparison studies

For comparison to A-cohorts, we considered prospective studies from any setting in any country with any therapeutic intervention including treatment-as-usual, and with a cohort of at least 20 evaluable patients, published in Danish, English, German, French, Italian, Norwegian, Russian, Spanish, or Swedish language (the languages understood by the authors).

Studies were eligible if $\geq 80\%$ of participants of the study or of a defined subgroup had one of the following five diagnoses, occurring in at least 20 adult AMOS patients: asthma, depression, low back pain, migraine, and neck pain. No requirements of diagnostic criteria were made. Low back pain cohorts with more than 25% patients with congenital spinal malformations, spinal infectious or malignant disease, ankylosing spondylitis, Behcet's Syndrome, Reiter's Syndrome, osteoporosis with vertebral fracture, spinal stenosis, spondylolysis, spondylolisthesis, fibromyalgia, traumatic vertebral fracture, or previous spinal operations were excluded from the analysis (these diagnoses were excluded from the corresponding A-cohort). Studies with persons aged ≥ 60 years only were also excluded (only 12% of A-patients were aged ≥ 60 years).

Studies were required to have at least one of the following ten outcomes from the SF-36 Health Survey, four-week version: eight SF-36 scales (Physical Function, Role Physical, Role Emotional, Social Functioning, Mental Health, Bodily Pain, Vitality, General Health), SF-36 Physical Component Summary Measure (PCS), or SF-36 Mental Component Summary Measure (MCS). The outcome was included if the arithmetic mean was presented with a number or could be estimated from a figure (a) before commencement of any study intervention and (b) after three, six, or 12 months ($\pm 20\%$).

Literature search

We searched Ovid MEDLINE, Ovid MEDLINE In-Process & Other Non-Indexed Citations, Journals@Ovid full text, BIOSIS Previews, Cochrane Database of Systematic Review, American College of Physicians Journal Club, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, Psychlit, the online SF-36 database (www.sf-36.org), literature references of retrieved articles and our own literature archive. Articles published up to December 2005 were considered.

The general search strategy was: “SF-36 keyword” in title, abstract or keyword AND “disease keyword” in title. SF-36 keywords were “SF-36” OR “Short-Form” OR “Medical Outcomes” OR “Quality of Life” OR “Disability” OR “Outcome Assessment (Health Care)”. Disease keywords were “Asthma” OR “Depress*” OR “Dysthymic disorder” OR “Low back pain” OR “Spinal diseas*” OR “Migraine” OR “Neck Pain” OR [“Spinal Diseas*” AND “Neck”].

Articles were read and assessed for provisional inclusion by one reviewer. All provisionally included articles were re-assessed for fulfilment of eligibility criteria by a second reviewer. Disagreements about fulfilment of eligibility criteria were resolved by discussion. Data of included articles were extracted and entered into Microsoft Excel data files by one person; all entered data were subject to source document verification by another person.

Analysis

Unit of analysis were individual SF-36 outcomes of the smallest statistically independent cohorts; data from cohorts which were not statistically independent (e. g. patients randomised to drug therapy with or without information booklet, outcomes not differing between the two treatment groups) were pooled prior to analysis.

For each C-cohort, the following study characteristics were extracted and tabulated: diagnosis, publication year, evaluable SF-36 outcomes, sample size at baseline, last evaluable follow-up, follow-up rates, design, setting, country, age, gender, disease duration, and study treatment.

The statistical analysis (SPSS 14.0) was descriptive. For each evaluable SF-36 outcome of a C-cohort, the pre-post difference (mean score at last evaluable follow-up minus mean baseline score) was subtracted from the corresponding difference of the respective A-cohort, yielding a mean outcome difference ($\text{Mean}_{A-FU} - \text{Mean}_{A-0} - (\text{Mean}_{C-FU} - \text{Mean}_{C-0})$). For each of the ten SF-36 outcomes, all mean outcome differences were aggregated with summary statistics of distribution of the differences and graphical presentation of the differences in increasing magnitude. In order to aggregate all differences of all outcomes, the differences were also expressed as between-group effect sizes through division by the standard deviation (SD) of the baseline score of the A-cohort ($(\text{Mean}_{A-FU} - \text{Mean}_{A-0}) - (\text{Mean}_{C-FU} - \text{Mean}_{C-0}) / \text{SD}_{A-0}$). The baseline SD of the A-cohort was used instead of the SD of the C-cohort or a

pooled SD from A- and C-cohorts because for many C-cohorts the SD was not available. To avoid redundancy when aggregating differences, comparisons of SF-36 PCS and MCS were not included if, for the C-cohort in question, all the eight SF-36 scales were evaluable for comparison. Effect sizes were classified as large (≥ 0.80), medium (0.50–0.79), small (0.20–0.49), and minimal (0.00–0.19) [21,22].

Results

Excluded publications

A total of 528 publications were excluded from this review for the following reasons: diagnosis not fulfilling eligibility criteria (n = 192 publications), no follow-up data (n = 129), no SF-36 data (n = 55), multiple publications (n = 43), SF-36 data presented without means (n = 30), cohort with patients ≥ 60 years only (n = 27), duration of follow-up differing $> 20\%$ from three, six, and 12 months, respectively (n = 20), no baseline SF-36 data (n = 12), cohort with < 20 patients (n = 8), SF-36 acute form only (n = 3), modified SF-36 (n = 3), language not fulfilling eligibility criteria (n = 2), no trial (n = 1), other (n = 3). A table of excluded publications with reasons for exclusion is provided in Appendix 1.

Description of AMOS subgroups and included comparison cohorts

The five A-cohorts with a total of 392 patients were compared to 85 C-cohorts with 16,432 patients. These 85 C-cohorts were presented in 64 publications (Table 1, Appendix 2). Diagnoses are displayed in Table 1. Seven of the 85 C-cohorts were published 1994–1996, 15 in the period 1997–1999, 21 in 2000–2002 and 42 were published in 2002–2005. Evaluable outcomes of C-cohorts were: all eight SF-36 scales (n = 41 of 85 C-cohorts), MCS or PCS or both (n = 11), all eight SF-36 scales plus MCS or PCS or both (n = 20), less than all eight SF-36 scales (n = 13). Median sample size per cohort was 56 patients (interquartile range (IQR) 41–127) for A-cohorts and 137 patients (IQR 65–244) for C-cohorts (Table 1). The last evaluable follow-up ensued after three months in 24 of 85 C-cohorts, after six months in 28 C-cohorts and after 12 months in 26 C-cohorts. Three-month-follow-up rates were 87.5% and 82.2% in A- and C-cohorts, respectively; six-month rates were 82.1% and 79.1%; and 12-month rates were 78.8% and 72.2%.

Study designs of C-cohorts were randomised controlled trials (n = 40 of 85 C-cohorts), non-randomised comparative studies (n = 9), and single-arm cohort studies (n = 36). Study settings of A-patients were primary care practice (85.5%, 335/392 A-patients), referral practice (3.8%), and outpatient clinic (10.7%). Study settings of C-cohorts were primary care or health maintenance organisation (n = 28 C-cohorts, 34.5% (5673/16432) of C-patients), non-academic hospital or outpatient clinic (n = 28 C-cohorts, 27.1% of C-patients), academic hospital or outpatient clinic (n = 19, 22.0%), and other or not specified (n = 10, 13.4%).

The 85 C-cohorts came from the USA (n = 38), Germany (n = 13), UK (n = 11), Canada (n = 4), Australia (n = 3), Japan (n = 3), Italy (n = 2), from eight further different countries (each C-cohort: n = 1) and from more than one country (n = 1). Mean age, weighted for sample size, was 44.4 years (SD 11.6) in A-cohorts and 44.4 years in C-cohorts (evaluable in 77/85 C-cohorts). The percentage of women was 83.2% (326/392) in A-cohorts and 60.8% (9243/15192) in C-cohorts (evaluable in 78 C-cohorts). Mean disease duration, weighted for sample size, was 10.5 years (SD 10.8) in A-cohorts and 11.8 years in C-cohorts (evaluable for 15/85 C-cohorts).

Main anthroposophic treatment modalities in A-patients were: eurythmy therapy (45.4%, 178 of 392 A-patients), art therapy (26.5%), rhythmical massage therapy (11.5%), and physician-provided anthroposophic therapy (16.6%). Study treatments in C-cohorts were: drugs (n = 33 of 85 C-cohorts, 48.2%, 7922 of 16432 C-patients), treatment-as-usual (n = 19 C-cohorts, 17.3% of C-patients), surgery (n = 9, 10.6%), physiotherapy (n = 5, 10.2%), other physical therapy (n = 5, 1.4%), educational intervention (n = 5, 10.6%), and mixed or other therapy (n = 7, 7.9%).

Comparisons between AMOS subgroups and comparison cohorts

A total of 560 comparisons between A-cohorts and corresponding C-cohorts were possible (Fig. 1). After exclusion of PCS and MCS comparisons of cohorts for which all eight SF-36 scales were evaluable for comparison, 517 comparisons remained for aggregation:

At baseline, A-cohorts were slightly more severely affected than C-cohorts (median difference 0.22 SD, IQR -0.13 to +0.53): Baseline differences between A- and C-cohorts were minimal or small (-0.49 SD to 0.49 SD) in 64.6% (334/517) of the comparisons, medium-to-large (≥ 0.50 SD) with A-cohorts more severely affected than C-cohorts in 27.5% of the comparisons, and medium-to-large with C-cohorts more severely affected in 7.9% of the comparisons.

At follow-up (Fig. 2, All diagnoses), the between-group effect sizes (pre-post improvements of A-cohorts minus pre-post improvements of C-cohorts) were median 0.11 (IQR -0.11 to 0.35).

- Effect sizes were positive, i. e. showing larger (≥ 0.20) improvements of A-cohorts than of C-cohorts in 41.0% (212/517) of the comparisons. These positive effect sizes were large (≥ 0.80) in 3.3% of the comparisons, medium (0.50-0.79) in 10.3%, and small (0.20-0.49) in 27.5% of the comparisons.
- Effect sizes showed minimal differences (-0.19 to 0.19) between A- and C-cohorts in 40.6% of the comparisons.
- Effect sizes were negative, i. e. showing larger improvements of C-cohorts than of A-cohorts in 18.4% of the comparisons. These negative effect sizes were large (≥ 0.80) in 2.5% of the comparisons, medium (0.50-0.79) in 4.3%, and small (0.20-0.49) in 11.6% of the comparisons.

Similar distributions of differences between the improvements in A- and C-groups were found in individual SF-36 outcomes (Fig. 1) and in individual diagnoses (Fig. 2).

Discussion

We have presented a systematic comparative review of SF-36 outcomes in five chronic conditions (asthma, depression, low back pain, migraine, neck pain). The review arose from the situation of having results from the first study of a given therapy (the AMOS study of anthroposophic medicine in outpatients [13]). The objective was to assess the order of magnitude of AMOS outcomes, relative to outcomes of other therapies. For this purpose we compared AMOS diagnostic subgroups (A-cohorts) to all retrievable patient cohorts (C-cohorts) with corresponding diagnoses, outcome measures and follow-up periods. More than 500 comparisons of ten different SF-36 scales showed improvements largely of the same order of magnitude in corresponding A- and C-cohorts, with small differences in 80%

of the comparisons, and medium-to-large differences favouring A- and C-groups in 14% and 7%, respectively.

This systematic review has five characteristic features: 1) we compared one therapy to all other treatments for the respective indications, 2) each comparison was of corresponding cohorts from different studies, 3) comparisons were restricted to cohorts with identical outcome measure and comparable follow-up periods, 4) analyses were descriptive, with results ordered in increasing magnitude instead of being pooled, 5) different outcomes were converted to a common metric, allowing for data synthesis into one variable. We are not aware of other systematic reviews combining these five features.

The present type of review can be regarded as a systematic extension and upgrading of the common ‘discussion-reviews’ in publications presenting new therapies, where results of the first therapy study are compared descriptively to results of other treatments for the given indication. In contrast to such narrative reviews, the present review has the strengths of systematic, criteria-based literature selection and analysis. For complementary and other complex therapy systems in widespread use regardless of evidence from randomised trials, it has been argued that the conventional drug research strategy – starting with studies of biological mechanisms and moving through Phase I, II and III clinical trials – should be replaced by a more appropriate strategy, moving from descriptive studies (‘Phase 1’) towards comparative studies of the whole system and its parts, and ending with studies of biological mechanisms (‘Phase 5’) [23]. In the context of this reversed strategy, the present review would represent an intermediate step between Phases 1-2 (studies of paradigms, utilization, perceived benefit, safety) and Phase 3 (comparative effectiveness studies).

Notably, the present review is limited to comparative order of magnitude. For our review, data from one single-arm study was available. Accordingly, each comparison was of two cohorts derived from different studies conducted in different settings. The only predefined criteria were comparable diagnosis and follow-up period, and identical outcome measure. A- and C-cohorts were found to be similar regarding age, disease duration, baseline affection and follow-up rates, and different regarding sample size and gender. Because of the obvious heterogeneity in these non-concurrent comparisons, the assessment could not be aimed at statistical precision. The analyses were purely descriptive, without attempting to pool data or to adjust for baseline differences.

Our search strategy was limited to ten online databases, thus some eligible studies may have been missed. Furthermore, only two cohorts with neck pain were available for comparison. Otherwise a range of patient settings (primary care, clinic, academic hospital) and treatments (as-usual, drugs, physical therapies, educational intervention, surgery) were represented in the C-cohorts.

Finally, the present review was restricted to a generic instrument (SF-36). Disease-specific comparisons might be more sensitive to relevant differences undetected by this review. On the other hand, anthroposophic therapy aims to improve a broad range of symptoms and functional limitations rather than only disease-specific symptoms [24], an aim which may be better reflected with broad instruments like the SF-36.

Conclusions

Within the limits of non-concurrent comparisons, this review suggests that anthroposophic therapy for chronic asthma, back or neck pain, depression and migraine can be associated with improvements of SF-36 scales of largely the same order of magnitude as improvements following other treatments. This implication

may sound trivial but is not: If our analyses had shown mostly small improvements compared to other treatments, one might have concluded that further studies of SF-36 as outcome of anthroposophic therapy are not worthwhile. Had our analysis shown large differences favouring anthroposophic treatment, results would have appeared more sensational. The present results suggest that anthroposophic therapy can be associated with clinically meaningful health improvements, and that SF-36 can be a useful outcome measure for further investigations of this therapy. The analysis also demonstrates the value of a systematic approach to non-concurrent comparisons of a therapy to corresponding cohorts, which is particularly relevant for therapies evaluated exclusively or predominantly in single-arm studies.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

HJH, GSK and HK designed the review. HJH and WT wrote the analysis plan. AG analysed data. HJH performed literature search, re-assessed provisionally included articles, was principal author of the paper, had full access to all data, and is guarantor. All authors contributed to manuscript drafting and revision and approved the final manuscript.

Acknowledgements

This review was funded by the Software-AG Stiftung and the Innungskrankenkasse Hamburg, with supplementary grants from the Deutsche BKK, the Betriebskrankenkasse des Bundesverkehrsministeriums, the Dr. Hauschka Stiftung, the Förderstiftung Anthroposophische Medizin, the Mahle Stiftung, and the Zukunftsstiftung Gesundheit. The sponsors had no influence on study design or planning; on collection, analysis, or interpretation of data; on the writing of the manuscript; or on the decision to submit the manuscript for publication. We thank P. Siemers for technical assistance.

References

1. Claassen J, Hirsch LJ, Emerson RG, Mayer SA: **Treatment of refractory status epilepticus with pentobarbital, propofol, or midazolam: a systematic review.** *Epilepsia* 2002, **43**: 146-153.
2. Sibai BM: **Diagnosis and management of gestational hypertension and preeclampsia.** *Obstet Gynecol* 2003, **102**: 181-192.
3. Srisurapanont M, Kittiratanapaiboon P, Jarusuraisin N: **Treatment for amphetamine psychosis.** *Cochrane Database Syst Rev* 2001, CD003026.
4. Critchley P, Plach N, Grantham M, Marshall D, Taniguchi A, Latimer E *et al.*: **Efficacy of haloperidol in the treatment of nausea and vomiting in the palliative patient: a systematic review.** *J Pain Symptom Manage* 2001, **22**: 631-634.

5. McCahill L, Ferrell B: **Palliative surgery for cancer pain.** *West J Med* 2002, **176**: 107-110.
6. Esiashvili N, Landry J, Matthews RH: **Carcinoma of the anus: strategies in management.** *Oncologist* 2002, **7**: 188-199.
7. Huisstede BM, Miedema HS, van Opstal T, de Ronde MT, Kuiper JI, Verhaar JA *et al.*: **Interventions for treating the posterior interosseus nerve syndrome: a systematic review of observational studies.** *J Peripher Nerv Syst* 2006, **11**: 101-110.
8. Ismail-Khan R, Robinson LA, Williams CC, Jr., Garrett CR, Bepler G, Simon GR: **Malignant pleural mesothelioma: a comprehensive review.** *Cancer Control* 2006, **13**: 255-263.
9. Gill AL, Bell CN: **Hyperbaric oxygen: its uses, mechanisms of action and outcomes.** *QMJ* 2004, **97**: 385-395.
10. Whitlock EP, Garlitz BA, Harris EL, Beil TL, Smith PR: **Screening for hereditary hemochromatosis: a systematic review for the U.S. Preventive Services Task Force.** *Ann Intern Med* 2006, **145**: 209-223.
11. Caffrey SL, Willoughby PJ, Pepe PE, Becker LB: **Public use of automated external defibrillators.** *N Engl J Med* 2002, **347**: 1242-1247.
12. Hoeper MM, Halank M, Marx C, Hoeffken G, Seyfarth HJ, Schauer J *et al.*: **Bosentan therapy for portopulmonary hypertension.** *Eur Respir J* 2005, **25**: 502-508.
13. Hamre HJ, Becker-Witt C, Glockmann A, Ziegler R, Willich SN, Kiene H: **Anthroposophic therapies in chronic disease: The Anthroposophic Medicine Outcomes Study (AMOS).** *Eur J Med Res* 2004, **9**: 351-360.
14. Hamre HJ, Witt CM, Glockmann A, Ziegler R, Willich SN, Kiene H: **Anthroposophic medical therapy in chronic disease: a four-year prospective cohort study.** *BMC Complement Altern Med* 2007, **7**: doi:10.1186/1472-6882-7-10.
15. Hamre HJ, Witt CM, Glockmann A, Ziegler R, Willich SN, Kiene H: **Eurythmy therapy in chronic disease: a four-year prospective cohort study.** *BMC Public Health* 2007, **7**: doi:10.1186/1471-2458-7-61.
16. Hamre HJ, Witt CM, Glockmann A, Ziegler R, Willich SN, Kiene H: **Anthroposophic art therapy in chronic disease: a four-year prospective cohort study.** *Explore* 2007, **3**: 365-371.
17. Hamre HJ, Witt CM, Glockmann A, Ziegler R, Willich SN, Kiene H: **Rhythmical massage therapy in chronic disease: a four-year prospective cohort study.** *J Altern Complement Med.* 2007.

18. Hamre HJ, Witt CM, Glockmann A, Ziegler R, Willich SN, Kiene H: **Anthroposophic therapy for chronic depression: a four-year prospective cohort study.** *BMC Psychiatry* 2006, **6**: doi:10.1186/1471-244X-6-57.
19. Kienle GS, Kiene H, Albonico HU: *Anthroposophic medicine: effectiveness, utility, costs, safety.* Stuttgart, New York: Schattauer Verlag; 2006.
20. Ware JE: **SF-36 Health Survey Update.** *Spine* 2000, **25**: 3130-3139.
21. Cohen J: **A power primer.** *Psychological Bulletin* 1992, **112**: 155-159.
22. McDowell I, Newell C: *Measuring health. A guide to rating scales and questionnaires*, 2. edn. New York - Oxford: Oxford University Press; 1996.
23. Fonnebo V, Grimsgaard S, Walach H, Ritenbaugh C, Norheim AJ, MacPherson H *et al.*: **Researching complementary and alternative treatments--the gatekeepers are not at home.** *BMC Med Res Methodol* 2007, **7**: 7.
24. Ritchie J, Wilkinson J, Gantley M, Feder G, Carter Y, Formby J: *A model of integrated primary care: anthroposophic medicine.* London: Department of General Practice and Primary Care, St Bartholomew's and the Royal London School of Medicine, Queen Mary, University of London; 2001.

Figures

Figure 1 - Outcome differences: individual SF-36 scales

Differences between pre-post improvements of AMOS cohorts and comparison cohorts for the eight SF-36 scales (0-100) and the SF-36 Physical and Mental Component Summary measures, ordered in increasing magnitude for each scale (altogether n = 560 comparisons). Positive differences indicate larger pre-post improvement in AMOS cohort than in corresponding comparison cohort.

Figure 2 - Outcome differences: diagnoses

Differences between pre-post improvements of AMOS cohorts and comparison cohorts for all SF-36 scales and summary measures, expressed in effect sizes and ordered in increasing magnitude: for all diagnoses and for individual diagnoses (altogether n = 517 comparisons). Positive effect sizes indicate larger pre-post improvement in AMOS cohort than in corresponding comparison cohort.

...

Tables

Table 1 - Overview of cohorts, patients and diagnoses

Diagnosis	AMOS cohorts				Comparison cohorts				
	Cohorts	Patients		Patients per cohort	Publications	Cohorts	Patients		Patients per cohort
	N	N	Percent	Mean	N	N	N	Percent	Mean
Asthma	1	56	14.3%	56	11	12	2,030	14.1%	169
Depression	1	174	44.4%	174	14	20	4,705	23.5%	235
Low back pain	1	80	20.4%	80	24	30	4,701	35.3%	157
Migraine	1	42	10.7%	42	13	21	4,240	24.7%	202
Neck pain	1	40	10.2%	40	2	2	756	2.4%	378
Total	5	392	100.0%	78	64	85	16,432	100.0%	194

...

Additional files

Additional file 1 – Excluded publications

Separate file: HAMRE_OUTCOME_Appendix_1.doc

Additional file 2 – Included publications

Separate file: HAMRE_OUTCOME_Appendix_2.doc

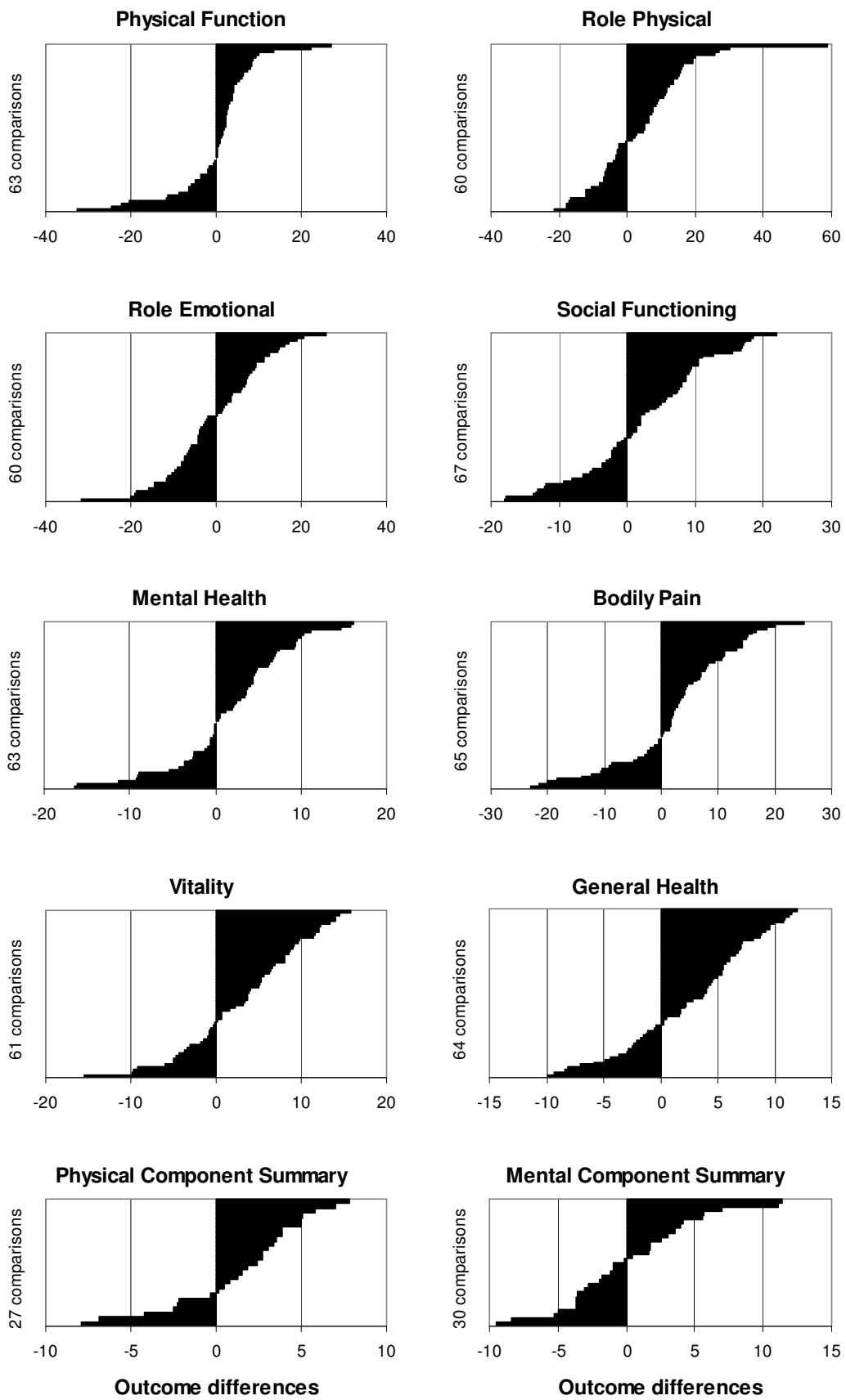


Figure 1

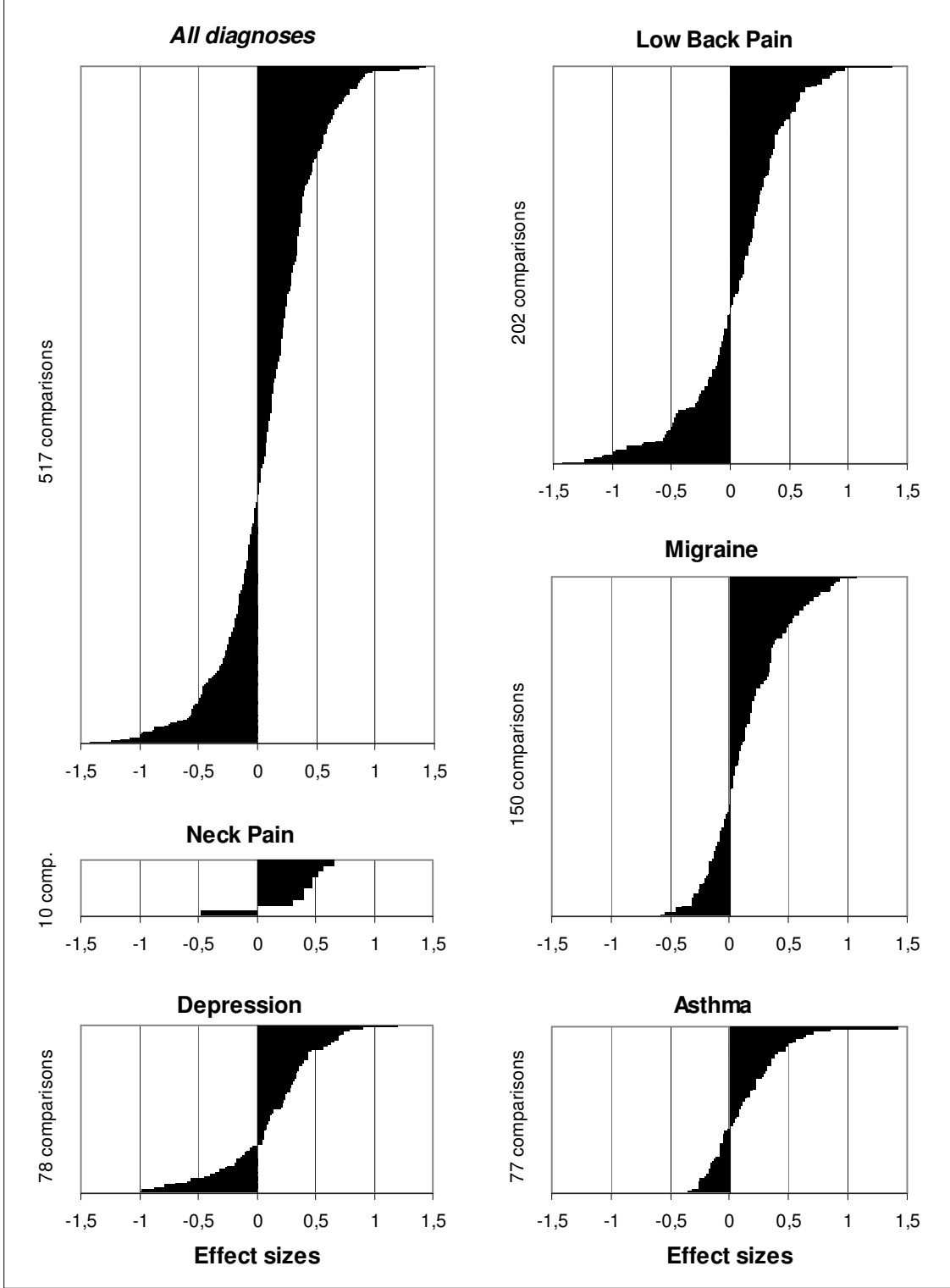


Figure 2

Additional files provided with this submission:

Additional file 1: hamre_outcome_appendix_1.doc, 737K

<http://www.biomedcentral.com/imedia/1450991494151355/supp1.doc>

Additional file 2: hamre_outcome_appendix_2.doc, 56K

<http://www.biomedcentral.com/imedia/7644268215135558/supp2.doc>