DOI:
10.1017/S0033291716002002

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

Cost-effectiveness of focal psychodynamic therapy and enhanced cognitive behavioural therapy in outpatients with anorexia nervosa

Nina Egger[1$], Beate Wild[2$], Stephan Zipfel[3], Florian Junne[3], Alexander Konnopka[1], Ulrike Schmidt[4], Martina de Zwaan[5], Stephan Herpertz[6], Almut Zeeck[7], Bernd Löwe[8], Jörn von Wietersheim[9], Sefik Tagay[10], Markus Burgmer[11], Andreas Dinkel[12], Wolfgang Herzog[2]*, and Hans-Helmut König[1]*

[1] Department of Health Economics and Health Services Research, Hamburg Center for Health Economics (HCHE), University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[2] Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Heidelberg, Germany

[3] Department of Psychosomatic Medicine and Psychotherapy, University Hospital Tübingen, Tübingen, Germany

[4] Section of Eating Disorders, Department of Psychological Medicine, King's College London, London, UK

[5] Department of Psychosomatic Medicine and Psychotherapy, Hannover Medical School, Hannover, Germany

[6] Department of Psychosomatic Medicine and Psychotherapy, LWL-University Clinic Bochum, Ruhr-University Bochum, Bochum, Germany

[7] Department of Psychosomatic Medicine and Psychotherapy, University Hospital Freiburg, Freiburg, Germany

[8] Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, and Schön Klinik Hamburg-Eilbek, Hamburg, Germany

[9] Department of Psychosomatic Medicine and Psychotherapy, University Hospital of Ulm, Ulm, Germany

[10] Clinic for Psychosomatic Medicine and Psychotherapy, LVR Hospital Essen, University of Duisburg-Essen, Essen, Germany

[11] Department of Psychosomatics and Psychotherapy, University Hospital Münster, Münster, Germany

[12] Department of Psychosomatic Medicine and Psychotherapy, Klinikum rechts der Isar, Technische Universität München, Munich, Germany

[$] Nina Egger and Beate Wild share the first authorship in this manuscript

*Wolfgang Herzog and Hans-Helmut König are the senior authors of the manuscript

Short title: Cost-effectiveness of FPT and CBT-E in AN

This study was funded by a grant from the German Ministry of Education and Research (grant number 01GV1002).

Word count: 4110

Corresponding author:
Nina Egger
Department of Health Economics and Health Services Research
University Medical Center Hamburg-Eppendorf
Martinistraße 52
20246 Hamburg, Germany.
E-mail: n.egger@uke.de
phone: +49 40 7410 54180
fax: +49 40 7410 40261

Abstract

Background: Anorexia nervosa (AN) is a serious illness leading to substantial morbidity and mortality. The treatment of AN very often is protracted; repeated hospitalizations and lost productivity generate substantial economic costs in the health care system. Therefore, this study aimed to determine the differential cost-effectiveness of outpatient focal psychodynamic psychotherapy (FPT), enhanced cognitive behavioural therapy (CBT-E), and optimised treatment as usual (TAU-O) in the treatment of adult women with AN.

Method: The analysis was conducted alongside the randomised controlled ANTOP study. Cost-effectiveness was determined using direct costs per recovery at 22 months post-randomization (n=156). Unadjusted incremental cost-effectiveness ratios (ICERs) were calculated. To derive cost-effectiveness acceptability curves (CEACs) adjusted net-benefit regressions were applied assuming different values for the maximum willingness to pay (WTP) per additional recovery. Cost-utility and assumptions underlying the base case were investigated in exploratory analyses.

Results: Costs of inpatient treatment and the percentage of patients who required inpatient treatment were considerably lower in both intervention groups. The unadjusted ICERs indicated FPT and CBT-E to be dominant compared to TAU-O. Moreover, FPT was dominant compared to CBT-E. CEACs showed that the probability for cost-effectiveness of FTP compared to TAU-O and CBT-E was ≥95% if the willingness to pay per recovery was ≥€9,825 and ≥€24,550, respectively. Comparing CBT-E with TAU-O, the probability of being cost-effective remained <90% for all WTPs. The exploratory analyses showed similar but less pronounced trends.

Conclusions: Depending on the WTP, FPT proved cost-effective in the treatment of adult AN.

Keywords: cost effectiveness; anorexia nervosa; cognitive therapy; psychodynamic psychotherapy

Introduction

Anorexia nervosa (AN) is a serious illness with substantial morbidity and mortality (Zerwas et al., 2015, Zipfel et al., 2014). About half the patients fully recover from AN, one third improves and 20% remain chronically ill (Zipfel et al., 2015). Especially in adults, who typically have a more enduring form of the illness, treatment is often protracted; repeated hospitalisations and lost productivity generate substantial personal and societal costs (Stuhldreher et al., 2012).

Whilst outpatient psychotherapy is the recommended first-line treatment for most adults with AN, available evidence from randomised controlled trials (RCTs) investigating different therapies has not identified any such treatment as clearly more effective than others (Zipfel et al., 2015). Thus, evidence-based guidelines (National Collaborating Centre for Mental Health, 2004) suggest that different psychotherapies, including inter alia cognitive behavioural treatment or focal psychodynamic psychotherapy, may be used in these patients depending on availability and preference. Therefore, information regarding the cost-effectiveness of different interventions could be of interest to policy makers to help determine optimal allocation of limited resources within a given health care system (Crow, 2014).

To date, no study has investigated the cost-effectiveness of adult AN treatments (Crow, 2014). The only published study compared three different care pathways for adolescents with AN, and in this study specialist outpatient treatment was more effective and less costly than both alternatives (Byford et al., 2007).

Here we present a cost-effectiveness analysis of the ANTOP study, the world-wide largest (n=242) multi-centre RCT comparing the efficacy of focal psychodynamic therapy (FPT), enhanced cognitive-behavioural therapy (CBT-E) and optimised treatment as usual (TAU-O) in adult outpatients with AN (Wild et al., 2009, Zipfel et al., 2014). The efficacy analysis showed no differences between groups regarding the main outcome, body mass index (BMI) at the end of treatment. However, the FPT group had a significantly higher recovery rate compared with TAU-O. The economic analysis aims to provide in-depth information

regarding direct and indirect costs of the ANTOP patients and their association with the different treatment arms.

Direct costs refer to expenses directly resulting from treatment, e.g. hospitalisations, outpatient treatment or medications. (Gray, 2011). Indirect costs describe costs due to reduced or lost productivity (Gray, 2011). In principle, the cost-effectiveness of an intervention can be evaluated from several perspectives. (Gray, 2011). The most comprehensive perspective is the societal perspective, which includes both direct and indirect costs.

In a cost-effectiveness analysis, treatment alternatives are assessed regarding costs and outcome by the means of an incremental analysis. As a result, the incremental analysis renders the expenses that are required to gain an additional health effect. If one treatment clearly emerges to be more effective and less costly, the decision which intervention should be adopted is straightforward. However, if one intervention is more effective at higher costs, the decision depends on the acceptable trade-off between costs and effects, i.e. the maximum willingness to pay (WTP) per additional health gain (Gray, 2011).

The aim of this analysis was to determine the cost-effectiveness of FPT, CBT-E, and TAU-O, all of which are widely used approaches in the treatment of adult women with AN.

Methods

Study design and participants

Detailed information regarding study design (Wild et al., 2009) and clinical outcomes (Zipfel et al., 2014) has been published previously. The study took place between May 2007 and May 2011 at 10 University Departments of Psychosomatic Medicine and Psychotherapy across Germany. Participants were adult women with a BMI between 15 and 18.5 kg/m^2 and a primary diagnosis of AN or subsyndromal AN. All participants gave written informed consent. Women with current substance abuse, medication with neuroleptics, suicidal ideation, psychotic or bipolar disorder, ongoing psychotherapy, pregnancy, a primary somatic disease, and women who were medically unstable were excluded. Initially, 242 participants were randomised to either FPT (n=80), CBT-E (n=80), or TAU-O (n=82). Data on clinical outcomes, health care utilisation and productivity loss were collected at baseline, at end of treatment and at two follow-ups after 3, and after 12 months (total study period: 22 months). The local research ethics committees of each study site approved the protocol.

Treatments

Treatments for both the FPT and CBT-E groups were individual outpatient therapies based on standardized treatment manuals. Each intervention comprised up to 40 sessions delivered over 10 months. Details on the structure and content of both interventions are elsewhere (Fairburn, 2008, Friederich et al., 2014, Legenbauer and Vocks, 2005, Zipfel et al., 2014).

To avoid contamination between groups, treatment was provided by different therapists, who were skilled at the respective therapeutic approach. Therapists received initial 2-day training from experts in FPT or CBT-E, followed by annual training updates. At every fourth session, experienced experts in FPT, or CBT-E respectively, supervised the therapists' work. Four full–length sessions were recorded and analyzed at the PI centers (Tübingen, Germany for CBT-E and Heidelberg, Germany for FPT). Based on the audio tapes, therapists received a

prompt, but brief, structured email of feedback regarding their adherence to the therapy manual (Zipfel et al., 2014). The implementation of the two treatment approaches was equally facilitated and ensured for both intervention groups.

Participants in the TAU-O group were given details of psychotherapists in their area who might be able to offer outpatient treatment for AN according to the German psychotherapy guidelines. Additionally, their family doctors, who measured weight, took blood tests, and were asked to inform the respective study centre if a patient's condition worsened, monitored them regularly. There were no rules regarding dosage and type of therapy in this group.

Outcome measures

In the ANTOP study, patients' BMI at the end of the treatment was the primary outcome. In the main outcome analysis, BMI at the end of treatment and BMI at the one-year follow-up was compared between the three groups. In addition, full recovery was defined as having a BMI >17.5 kg/m^2 and a score on the psychiatric status rating scale (PSR) of 1 or 2 (indicating no symptoms of AN) (Zipfel et al., 2014). The PSR score was rated by masked assessors based on the full structured interview for anorexic and bulimic syndromes (SIAB-EX) (Fichter and Quadflieg, 2001).The PSR scale is used to measure the general severity of the anorexic disorder. PSR scores range from 1 (patient has no symptoms of anorexia nervosa) to 6 (patient has severe symptoms of anorexia nervosa that require admission). A score of 5 indicates that all DSM-IV criteria for anorexia nervosa have been fulfi lled. At baseline, all patients were assessed with a PSR score of 4 (subsyndromal anorexia nervosa, 45%) or 5 (full syndromal anorexia nervosa, 55%). Thus, at baseline, no patient met the definition of "full recovery".

For the present analysis, we used "recovery" at the end of the observation period of 22 months as the main outcome measure because, for clinicians and policy makers, a

recovered patient is of utmost importance. In further analyses, we also investigated the cost-effectiveness of the treatments regarding BMI and quality of life.

Health-related quality of life (HrQol) was measured using the EQ-5D-3L (Dolan, 1997). The EQ-5D-3L is a generic measure of HrQol covering five dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) with three levels each ("No problems", "Some problems", and "Extreme problems"); thus generating $3^5 = 243$ theoretically possible health states. To calculate quality adjusted life years (QALYs), which are the recommended measure of effect in economic evaluations, utility weights were assigned to the patient's health state. The utility weights reflect the respective health state over the relevant time period and range from 0 (death) to 1 (perfect health). In this study, we used utility weights which were derived from a large UK sample using the time trade-off method and which were shown to be valid for other cultures (Nan et al., 2007). Thus, living one year in perfect health corresponds to one QALY. We calculated QALYs for the observation period of 22 months using linear interpolation between the measurements. Thus, a patient could attain between 0 and 1.83 QALYs. However, since the EQ-5D has not been validated for patients with AN, QALYs were employed in a secondary analysis only.

Cost assessment

Information on inpatient treatment for AN was retrieved from monitoring data and hospital records to ensure that all inpatient stays within the observation period of 22 months were included.

Further health care utilisation and productivity loss were collected for 3 months retrospectively using a questionnaire. The instrument was based on the Client Socio-Demographic and Service Receipt Inventory, developed and validated by Chisholm et al. (Chisholm et al., 2000) and adapted to the specific requirements in this study. The questionnaire covered outpatient treatment by physicians, psychologists, and other

therapists, pharmacotherapy as well as formal nursing care and informal care. Moreover, costs of travel as well as travel distances were collected. Productivity losses were assessed using the number of sick leave days and hours with health service use. Reduced productivity at the work place was assessed for the remaining work days by means of a visual analogue scale from 0 (no reduction in productivity) to 10 (completely impaired).

To calculate direct costs, specific unit costs were employed reflecting average prices within the German healthcare system (Krauth et al., 2005, Rote Liste Service GmbH, 2008, Statistisches Bundesamt, 2008, Statistisches Bundesamt, 2009d, Deutsche Krankenhausgesellschaft, 2009, Statistisches Bundesamt, 2009a, Statistisches Bundesamt, 2009b). Informal care was valued using the opportunity cost approach, i.e. hours with informal care were considered as lost leisure time (Statistisches Bundesamt, 2010b). Thus, the mean net income was applied for valuation. For public transport the reported costs were charged; travel distances by car were valued with €0.30 per kilometre according to the tax-deductible rate for work-related travel in Germany.

The costs of outpatient mental health care reflect the number of therapy sessions each patient received during the observation period multiplied by the costs per session. This information was retrieved from monitoring data. These costs correspond to the fee that was reimbursed by the German statutory health insurances per session of individual psychotherapy in 2008 (€75.54). Intervention costs for FPT and CBT-E correspond to the number of sessions during the 10-month treatment.

The calculation of indirect costs was restricted to gainfully employed participants. The monetary valuation was based on the human capital method, i.e. all productivity loss during the observation period was considered. Costs of absenteeism, i.e. costs of sick leave and costs of productivity loss due to health service use, were valued using the average gross wage for women in Germany including non-wage labour costs (Statistisches Bundesamt, 2009d). Costs of presenteeism, i.e. costs of reduced productivity at the workplace, were calculated according to Kessler et al. (Kessler et al., 2004). To this end, the subjective rating

of reduced productivity was translated into days of lost productivity by multiplying the days at work by the proportion of impairment, e.g. if productivity was reduced by 50% and the participant had 10 working days, this resulted in 5 additional days of lost productivity. However, the valuation of presenteeism involves several restrictions regarding the conversion of reduced productivity into lost time and its translation into costs (Brooks et al., 2010). Therefore, these costs were only considered in a sensitivity analysis.

All costs were calculated in Euro for the year 2008. If unit costs were not available for 2008, previous data were inflated to 2008 price levels (Statistisches Bundesamt, 2010b).

Statistical analysis

We applied multiple imputation by chained equations (MICE) to account for missing information due to drop out and missings on single variables of resource use and outcomes (Azur et al., 2011, van Buuren and Groothuis-Oudshoorn, 2011). At the one-year follow-up, 73 participants (30%) had dropped out and the maximum percentage of missings per variable were 54%. We included all available data from all participants at each measurement point in the imputation process. In order to reduce the number of items for which models have to be fitted and to improve the fit for the remaining imputation models, we calculated costs per category of care and imputed missings at this superordinate level. Intervention group was included as covariate. To minimize the power-falloff, given the fraction of missing information, 40 imputations were created (Graham et al., 2007).

For the main - *base case* - analysis, we employed the imputed datasets but restricted the sample to participants for whom data regarding the global outcome (recovered/ not recovered) at the one-year follow-up was available (n=156). Additionally, we conducted intention-to-treat (ITT) analyses including all randomised patients (n=242) in order to investigate the robustness of the results.

Since the majority of costs in AN patients results from hospitalisations, in the base case analysis only direct costs were considered. Recovery rates after 22 months were defined as the measure of health effects and direct costs during the total observation period of 22 months were defined as measure of costs. Linear interpolation was used to estimate costs between measurements and to calculate the total sum.

Cost-effectiveness was analysed in two steps. First, we calculated the unadjusted incremental cost-effectiveness ratios (ICERs) for all three pairwise comparisons (Hoch et al., 2002). The ICER renders the costs associated with one intervention to achieve an additional health effect in relation to the respective comparator.

The ICER corresponds to an unadjusted point estimate. Therefore, in the second step we applied the net-benefit approach and calculated the net-monetary-benefit (NMB) (Hoch et al., 2006, Zethraeus et al., 2003) to determine the statistical uncertainty of these point estimates and to adjust for covariates and baseline differences.

Multivariate linear regression models were fitted to determine whether FPT and CBT-E are cost-effective compared to TAU-O. A linear contrast was calculated to analyse the cost-effectiveness of FPT compared to CBT-E. Age, illness duration (≤ 6 years vs. >6 years), comorbid affective and anxiety disorders, employment status as well as BMI at baseline, and baseline costs were entered as covariates. To adjust for the multi-centre design and potential regional differences, study sites were additionally included as fixed effects. Since the maximum willingness to pay (WTP) per recovered patient is unknown, the regression analyses were re-run, and the WTPs were varied between €0 and €150,000 in steps of €10,000.

Differences in costs and outcomes between the intervention groups and TAU-O after 22 months were investigated using generalized linear models with respective distributions and link functions(de Jong and Heller, 2008). Covariates correspond to those described before and included the respective baseline values.

The results are presented in graphs as cost-effectiveness acceptability curves (CEACs). Given the data, a CEAC renders the probability that an intervention is cost-effective for a range of WTPs (Fenwick et al., 2004). However, there is no unique probability threshold above which an intervention is considered cost-effective with certainty; a common threshold is >95%.

We used SAS software for all statistical analyses (Version 9.3 of the SAS System for Windows. Copyright © 2002-2010 SAS Institute Inc., Cary, NC, USA). IVEware (Imputation and Variance Estimation Software), a SAS-based application developed by Raghunathan et al. (Raghunathan et al., 2001) was employed to perform MICE.

Results

The characteristics of the base case sample (n=156 patients) are presented in Table 1. A description of the patient flow can be found in the main study (Zipfel et al., 2014). The base case and the ITT sample were comparable regarding all baseline characteristics. However, at the end of the observation period dropout rates differed significantly between the treatment groups with the highest dropout rate observed in TAU-O (51%), the lowest in CBT-E (22%). This led to significant differences between groups regarding illness duration, comorbid affective, and comorbid anxiety disorder. Therefore, we controlled for illness duration and comorbid conditions in the adjusted analyses.

Please insert Table1 about here.

Costs

The unadjusted mean costs 3 months prior baseline and the unadjusted mean total costs for the observation period of 22 months are reported in Table 2. Mean (SD) intervention costs accrued to €2,613 (98) in FPT and €2,494 (102) in CBT-E. Baseline costs were comparable in FPT and CBT-E but the TAU-O group showed higher costs in nearly all categories. Therefore, the unadjusted cost estimates should be interpreted with caution. After adjusting for baseline values and covariates total costs for 22 months did not differ significantly between groups.

At the end of the observation period, costs of hospitalisations accounted for 45%, 55%, and 70% of the direct costs in FPT, CBT-E and TAU-O, respectively. This corresponds to a mean (SD) duration of admissions of 18.4 (52.9) days in FPT, 26.5 (53.9) in CBT-E, and 32.6 (57.3) in TAU-O, respectively. In FPT 19% of the patients required inpatient treatment, 29% in CBT-E, and 40% in TAU-O. However, none of the differences were significant.

Please insert Table2 about here.

Effectiveness

At the end of the observation period, 35.2% of the patients in FPT were defined as recovered compared to 21% in CBT-E and 12.5% in TAU-O. The difference between FPT and TAU-O was significant (p=0.036). The mean (SE) BMI was 18.2 (0.24) in FPT, 18.1 (0.23) in CBT-E, and 17.9 (0.26) in TAU-O. With regard to the secondary outcome, quality of life, mean (SE) QALYs were 1.53 (0.41) in FPT, 1.48 (0.54) in CBT-E, and 1.44 (0.47) in TAU-O. However, as quality of life also differed between groups at baseline, the unadjusted QALYs should be considered preliminary. In the adjusted analyses, no statistically significant difference between groups was found for BMI and QALYs.

Cost-effectiveness

In the unadjusted base case analysis using direct costs as the measure of costs and recovery rates as the measure of effect, FPT was found to be dominant compared to TAU-O and CBT-E, i.e. it was more effective at lower costs. CBT-E also was dominant compared to TAU-O. Unadjusted ICERs are presented in Table3.

Please insert Table3 about here.

Due to the differences in baseline costs and covariates, the adjusted results from the net-benefit regressions are more informative. The CEACs in Figure 1 show the probability of each intervention being cost-effective at different WTPs. If higher WTPs were assumed, the probability of being cost-effective increased for all comparisons. At WTPs ≥€9,825 per

recovered patient, FPT was certainly cost-effective, i.e. the probability was above the probability threshold of 95%. Comparing FPT to CBT-E, the required WTP was ≥€24,550. In contrast, CBT-E could not be considered cost-effective compared to TAU-O; the probability remained below 90% for all considered WTPs.

Please insert Figure1 about here.

Secondary analyses

When QALYs were employed as an alternative outcome, using the base case sample, the unadjusted results were similar (Table3): FPT was dominant compared to TAU-O and CBT-E; CBT-E was dominant compared to TAU-O. In the adjusted analyses, the probability of being cost-effective was 82% for FPT compared to TAU-O at the common WTP of €50,000 per QALY, and 75% compared to CBT-E (Figure 2). For CBT-E compared to TAU-O the respective probability was 62%.

Please insert Figure 2 about here.

BMI at the end of the observation period was also employed as measure of effect in a secondary analysis using n=169 patients with available data at the end of the observation period. Here, in the unadjusted ICERs (Table3) FPT and CBT-E were dominant compared to TAU-O. FPT was associated with lower mean costs and lower mean effects compared to CBT-E. To ease the interpretation of the ICER, the relationship is presented inversely here, i.e. CBT-E is associated with additional costs of €11,600 per gained BMI point compared to FPT. In the adjusted analyses, at a WTP of €5,000 for instance, the probability of being cost-effective was 87% for FPT, 77% for CBT-E, in comparison to TAU-O. The respective

probability of FPT compared to CBT-E was 67%. Figure 3 reveals that the probabilities for both interventions increase if higher WTPs can be assumed. However, in respect to BMI, FPT and CBT-E seem to be rather similar.

Please insert Figure 3 about here.

In a further analysis, we included indirect costs due to productivity loss. Regarding the unadjusted ICERs, both interventions remained dominant compared to TAU-O and FPT was dominant compared to CBT-E (Table3). The adjusted cost-effectiveness acceptability curves (data not shown) of FPT and CBT-E compared to TAU-O were similar to the base case, but increased more slowly. In contrast, the probability of FPT being cost-effective compared to CBT-E increased more quickly. At a WTP of €0, FPT already had a probability of being cost-effective of 80%.

To investigate the robustness of our results we examined the cost-effectiveness of the interventions based on the ITT sample (n=242), and repeated the analyses for each outcome (recovery, QALYs, and BMI). The unadjusted results are reported in Table 3. The adjusted results are presented in the respective figures together with the results from the base case sample. With regard to the primary outcome, recovery, FPT remained dominant compared to CBT-E and TAU-O. CBT-E was associated with additional costs of €23,704 per recovery compared to TAU-O (Table3). In the adjusted analyses (Figure 1), FPT's probability of being cost-effective compared to TAU-O was <90% for all WTPs. For FPT compared to CBT-E, the probability remained <95% for all WTPs. At WTPs <€20,000, FPT's probability of being cost-effective compared to CBT-E was higher than in the base case analysis. CBT-E could not be considered cost-effective, i.e. the probability of being cost-effective remained >50% for all WTPs.

Discussion

This is the first study to assess the cost-effectiveness of outpatient psychological therapies in the treatment of adult women with AN. Costs were determined for the observation period of 22 months, including 10 months of treatment and 12 months follow-up. Although differences in costs were not significant, the absolute costs of inpatient treatment and the percentage of patients who required inpatient treatment were considerably lower in both intervention groups. With regard to recovery, results from the adjusted analyses indicate that FPT is cost-effective in comparison to TAU-O as well as in comparison to CBT-E, depending on a decision maker's maximum willingness to pay. CBT-E compared to TAU-O had a lower probability of being cost-effective. The analysis including direct and indirect costs showed a similar trend compared to the base case analysis, with slightly changed WTPs for the various treatment groups. In the investigation of the cost-utility using QALYs the unadjusted estimates showed a similar trend as the base case analysis. However, the adjusted estimates indicated that regarding QALYs no intervention could be considered cost-effective with certainty. The results of the secondary analysis using BMI as a measure of effect showed smaller differences between FPT and CBT-E but both interventions were likely to be cost-effective compared to TAU. The analyses based on the ITT sample, using recovery or QALYs, pointed in the same direction as the respective base case analyses, with less pronounced findings. However, with regard to BMI cost-effectiveness remained ambiguous. We therefore conclude that there is at least a trend for FPT being cost-effective compared to CBT-E and TAU-O in the treatment of adult outpatients with AN. The cost-effectiveness of CBT-E compared to TAU-O is subject to greater uncertainty.

Putting our findings into context, to date, there is no evidence regarding cost-effectiveness of treatments for adults with AN. In adolescent patients with AN, Byford et al. (Byford et al., 2007) reported that specialist outpatient treatment based on individual CBT and parental counselling was cost-effective compared to inpatient treatment, as well as compared to TAU provided by community child and adolescent mental health services. Here, the differences in

costs also resulted primarily from differences in inpatient costs. Participants randomised to TAU required almost as much inpatient treatment as participants initially allocated to inpatient treatment. These results are not comparable to our findings because of different settings and study samples; however, they support the cost-effectiveness of specialist outpatient treatment in general.

Comparing the two interventions with each other, FPT clearly emerged as advantageous. However, studies, which have compared both approaches directly, are scarce and none have included any economic evaluation. Therefore, the particular characteristics of FPT, which led to higher recovery rates and reduced hospitalisations, are still unknown. Systematic research is warranted to investigate the specific impact of each intervention on health care utilisation.

The interventions' effects on indirect costs due to absenteeism and presenteeism are less clear and should be interpreted with caution: no other studies are available and our cost estimates were comparable across all groups. We observed lower employment rates than in the German general population in this age group but we were not able to determine whether unemployment was related to AN, and whether the employment status changed following the interventions because this information was obtained at baseline only. Thus, the development of indirect costs in AN patients remains to be evaluated in more detail in further studies together with the interventions' potential to improve patients' ability to work.

With regard to cost-utility, our results were subject to considerable uncertainty. As the validity of the EQ-5D-3L in eating disorders has not been determined we decided to use QALYs only in a sensitivity analysis. However, a recent study used the EQ-5D-3L in inpatients with severe AN (Abbate-Daga et al., 2014). In this study, patients' quality of life improved during hospitalisation with a mean index score of 0.70 at hospital discharge (Abbate-Daga et al., 2014). In our study, we also observed increases towards the end of treatment, but at the end of the one-year follow-up, the index scores had decreased below the respective baseline values in each group. This might indicate that improvements in health related quality of life

are short-term only and may reflect patients' ambivalence towards recovery, i.e. quality of life worsens in spite of clinical improvements (Abbate-Daga et al., 2013). However, the scores in our sample were clearly higher than those reported for other patient populations with chronic conditions, e.g. 0.57 in anxiety, and 0.53 in depression (Wu et al., 2015). This probably reflects that AN patients tend to misperceive the severity of their condition (Abbate-Daga et al., 2014, Ackard et al., 2014). Therefore, disease-specific measures currently seem superior to generic instruments in assessing quality of life in AN (Ackard et al., 2014).

Despite a strong study design, a large sample, and careful monitoring the applied instrument to assess health care utilisation covered only 3 months prior to each measurement to avoid any recall bias. Moreover, we observed missing data in single variables of health care utilisation at the different measurement time points as well as complete dropouts during follow-up. However, we applied several strategies to account for this missing information appropriately.

First, we tried to retrieve missing information from other sources, i.e. hospital costs were derived from clinical records, which covered the complete observation period. Second, we applied MICE to impute further missing data. Third, with regard to service use and productivity loss between the measurements we used linear interpolation to calculate costs for the complete observation period. This involves the assumption that service use is linear over time. However, it is unclear whether this produced higher or lower cost estimates and it is very unlikely that this assumption changed the results with regard to group differences.

In AN dropout from treatment poses a particular problem. Although there is evidence that patients who dropped out have negative long-term outcomes, the actual course of patients lost to follow-up is unknown (Dejong et al., 2012). We decided to restrict our base case analysis to patients for whom data regarding recovery, at one-year follow-up were available. This limits the generalisability of our results. Moreover, if reasons for dropout differed systematically between the three groups, this also might have led to an overestimation of the

interventions' cost-effectiveness. Results of the ITT analysis largely corresponded to the base case, but reflected greater uncertainty due to missing information.

To conclude, treatment of adult patients with AN remains challenging. Nonetheless, from a health care perspective focal psychodynamic therapy seems to provide a cost-effective strategy for improving weight and eating disorder pathology.

Financial support

This study was funded by a grant from the German Ministry of Education and Research (grant number 01GV1002). US receives salary support from the National Institute of Health Research (NIHR) Biomedical Research Centre (BRC) for Mental Health at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the National Health Service, the NIHR or the Department of Health.

Conflict of interest

None.

References

ABBATE-DAGA, G., AMIANTO, F., DELSEDIME, N., DE-BACCO, C. & FASSINO, S. 2013. Resistance to treatment and change in anorexia nervosa [corrected]: a clinical overview. *BMC Psychiatry,* 13**,** 294.

ABBATE-DAGA, G., FACCHINI, F., MARZOLA, E., DELSEDIME, N., GIOVANNONE, C., AMIANTO, F. & FASSINO, S. 2014. Health-related quality of life in adult inpatients affected by anorexia nervosa. *European eating disorders review,* 22**,** 285-91.

ACKARD, D. M., RICHTER, S., EGAN, A., ENGEL, S. & CRONEMEYER, C. L. 2014. The meaning of (quality of) life in patients with eating disorders: a comparison of generic and disease-specific measures across diagnosis and outcome. *The International journal of eating disorders,* 47**,** 259-67.

AZUR, M. J., STUART, E. A., FRANGAKIS, C. & LEAF, P. J. 2011. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res,* 20**,** 40-9.

BROOKS, A., HAGEN, S. E., SATHYANARAYANAN, S., SCHULTZ, A. B. & EDINGTON, D. W. 2010. Presenteeism: Critical Issues. *J Occup Environ Med,* 52**,** 1055-1067.

BYFORD, S., BARRETT, B., ROBERTS, C., CLARK, A., EDWARDS, V., SMETHURST, N. & GOWERS, S. G. 2007. Economic evaluation of a randomised controlled trial for anorexia nervosa in adolescents. *British Journal of Psychiatry,* 191**,** 436-40.

CHISHOLM, D., KNAPP, M. R., KNUDSEN, H. C., AMADDEO, F., GAITE, L. & VAN WIJNGAARDEN, B. 2000. Client Socio-Demographic and Service Receipt Inventory--European Version: development of an instrument for international research. EPSILON Study 5. European Psychiatric Services: Inputs Linked to Outcome Domains and Needs. *British Journal of Psychiatry,* 39**,** s28-33.

CROW, S. 2014. The economics of eating disorder treatment. *Current Psychiatry Reports,* 16**,** 454.

DE JONG, P. & HELLER, G. Z. 2008. *Generalized Linear Models for Insurance Data*, Cambridge University Press.

DEJONG, H., BROADBENT, H. & SCHMIDT, U. 2012. A systematic review of dropout from treatment in outpatients with anorexia nervosa. *The International Journal of Eating Disorders,* 45**,** 635-47.

DEUTSCHE KRANKENHAUSGESELLSCHAFT 2009. Bestandsaufnahme zur Krankenhausplanung und Investitionsfinanzierung in den Bundesländern. Berlin: Deutsche Krankenhausgesellschaft.

DOLAN, P. 1997. Modeling valuations for EuroQol health states. *Medical Care,* 35**,** 1095-108.

FAIRBURN, C. G. 2008. *Cognitive behavior therapy and eating disorders.,* New York London, The Guilford Press.

FENWICK, E., O'BRIEN, B. J. & BRIGGS, A. 2004. Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions. *Health Economics,* 13**,** 405-15.

FICHTER, M. & QUADFLIEG, N. 2001. The structured interview for anorexic and bulimic disorders for DSM-IV and ICD-10 (SIAB-EX): reliability and validity. *Eur Psychiatry,* 16**,** 38-48.

FRIEDERICH, H. C., HERZOG, W., WILD, B., ZIPFEL, S. & SCHAUENBURG, H. 2014. *Anorexia Nervosa. Fokale psychodynamische Psychotherapie.,* Göttingen, Hogrefe.

GRAHAM, J. W., OLCHOWSKI, A. E. & GILREATH, T. D. 2007. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev Sci,* 8**,** 206-13.

GRAY, A. M. 2011. *Applied methods of cost-effectiveness analysis in health care,* Oxford [u.a.], Oxford Univ. Press.

HOCH, J. S., BRIGGS, A. H. & WILLAN, A. R. 2002. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Economics,* 11**,** 415-30.

HOCH, J. S., ROCKX, M. A. & KRAHN, A. D. 2006. Using the net benefit regression framework to construct cost-effectiveness acceptability curves: an example using data from a trial of external loop recorders versus Holter monitoring for ambulatory monitoring of "community acquired" syncope. *BMC Health Services Research,* 6.

KESSLER, R. C., AMES, M., HYMEL, P. A., LOEPPKE, R., MCKENAS, D. K., RICHLING, D. E., STANG, P. E. & USTUN, T. B. 2004. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med,* 46**,** S23-S37.

KRAUTH, C., HESSEL, F., HANSMEIER, T., WASEM, J., SEITZ, R. & SCHWEIKERT, B. 2005. Empirical standard costs for health economic evaluation in Germany - a proposal by the working group methods in health economic evaluation. *Gesundheitswesen,* 67**,** 736-46.

LEGENBAUER, T. & VOCKS, S. 2005. *Manual der kognitiven Verhaltenstherapie bei Anorexie und Bulimie.,* Berlin, Springer.

NAN, L., JOHNSON, J. A., SHAW, J. W. & COONS, S. J. 2007. A comparison of EQ-5D index scores derived from the US and UK population-based scoring functions. *Med Decis Making,* 27**,** 321-6.

NATIONAL COLLABORATING CENTRE FOR MENTAL HEALTH 2004. National Institute for Health and Clinical Excellence: Guidance. *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders.* Leicester (UK): British Psychological Society (UK), The British Psychological Society & The Royal College of Psychiatrists.

RAGHUNATHAN, T. E., LEPKOWSKI, J. M., VAN HOEWYK, J. & SOLENBERGER, P. 2001. A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models. *SurveyMethodology,* 27.

ROTE LISTE SERVICE GMBH 2008. Rote Liste 2008. Arzneimittelverzeichnis für Deutschland. Frankfurt am Main: Rote Liste Service GmbH.

STATISTISCHES BUNDESAMT 2008. Pflegestatistik 2007. Wiesbaden: Statistisches Bundsamt.

STATISTISCHES BUNDESAMT 2009a. Kostennachweis der Krankenhäuser 2008. Wiesbaden: Statistisches Bundesamt.

STATISTISCHES BUNDESAMT 2009b. Grunddaten der Krankenhäuser 2008. Wiesbaden: Statistisches Bundesamt.

STATISTISCHES BUNDESAMT 2009d. Verdienste und Arbeitskosten. Wiesbaden: Statistisches Bundesamt.

STATISTISCHES BUNDESAMT 2010b. Volkswirtschaftliche Gesamtrechnungen 2009. Wiesbaden: Statistisches Bundesamt.

STUHLDREHER, N., KONNOPKA, A., WILD, B., HERZOG, W., ZIPFEL, S., LOWE, B. & KONIG, H. H. 2012. Cost-of-illness studies and cost-effectiveness analyses in eating disorders: a systematic review. *The International journal of eating disorders,* 45**,** 476-91.

VAN BUUREN, S. & GROOTHUIS-OUDSHOORN, K. 2011. MICE: Multiple Imputation by Chained Equations in R. *J Stat Softw,* 45.

WILD, B., FRIEDERICH, H. C., GROSS, G., TEUFEL, M., HERZOG, W., GIEL, K. E., DE ZWAAN, M., SCHAUENBURG, H., SCHADE-BRITTINGER, C., SCHAFER, H. & ZIPFEL, S. 2009. The ANTOP study: focal psychodynamic psychotherapy, cognitive-behavioural therapy, and treatment-as-usual in outpatients with anorexia nervosa--a randomized controlled trial. *Trials,* 10**,** 23.

WU, M., BRAZIER, J. E., KEARNS, B., RELTON, C., SMITH, C. & COOPER, C. L. 2015. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. *Eur J Health Econ,* 16**,** 141-51.

ZERWAS, S., LARSEN, J. T., PETERSEN, L., THORNTON, L. M., MORTENSEN, P. B. & BULIK, C. M. 2015. The incidence of eating disorders in a Danish register study: Associations with suicide risk and mortality. *Journal of Psychiatric Research,* 65**,** 16-22.

ZETHRAEUS, N., JOHANNESSON, M., JONSSON, B., LOTHGREN, M. & TAMBOUR, M. 2003. Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. *Pharmacoeconomics,* 21**,** 39-48.

ZIPFEL, S., GIEL, K. E., BULIK, C. M., HAY, P. & SCHMIDT, U. 2015. Anorexia nervosa: aetiology, assessment, and treatment. *The Lancet Psychiatry,* doi: 10.1016/S2215-0366(15)00356-9. [Epub ahead of print].

ZIPFEL, S., WILD, B., GROSS, G., FRIEDERICH, H. C., TEUFEL, M., SCHELLBERG, D., GIEL, K. E., DE ZWAAN, M., DINKEL, A., HERPERTZ, S., BURGMER, M., LOWE, B., TAGAY, S., VON WIETERSHEIM, J., ZEECK, A., SCHADE-BRITTINGER, C., SCHAUENBURG, H. & HERZOG, W. 2014. Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): randomised controlled trial. *Lancet,* 383**,** 127-37.