

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

N. Egger^{1*†}, B. Wild^{2†}, S. Zipfel³, F. Junne³, A. Konnopka¹, U. Schmidt⁴, M. de Zwaan⁵, S. Herpertz⁶, A. Zeeck⁷, B. Löwe⁸, J. von Wietersheim⁹, S. Tagay¹⁰, M. Burgmer¹¹, A. Dinkel¹², W. Herzog^{2‡} and H.-H. König^{1‡}

¹Department of Health Economics and Health Services Research, Hamburg Center for Health Economics (HCHE), University Medical Center Hamburg-Eppendorf, Hamburg, Germany; ²Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Heidelberg, Germany; ³Department of Psychosomatic Medicine and Psychotherapy, University Hospital Tübingen, Tübingen, Germany; ⁴Section of Eating Disorders, Department of Psychological Medicine, King's College London, London, UK; ⁵Department of Psychosomatic Medicine and Psychotherapy, Hannover Medical School, Hannover, Germany; ⁶Department of Psychosomatic Medicine and Psychotherapy, LWL-University Clinic Bochum, Ruhr-University Bochum, Bochum, Germany; ⁷Department of Psychosomatic Medicine and Psychotherapy, University Hospital Freiburg, Freiburg, Germany; ⁸Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, and Schön Klinik Hamburg-Eilbek, Hamburg, Germany; ⁹Department of Psychosomatic Medicine and Psychotherapy, University Hospital of Ulm, Ulm, Germany; ¹⁰Clinic for Psychosomatic Medicine and Psychotherapy, LVR Hospital Essen, University of Duisburg-Essen, Essen, Germany; ¹¹Department of Psychosomatics and Psychotherapy, University Hospital Münster, Münster, Germany; ¹²Department of Psychosomatic Medicine and Psychotherapy, Klinikum rechts der Isar, Technische Universität München, Munich, Germany

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 out-patient focal psychodynamic psychotherapy (FPT), enhanced cognitive-behavioural therapy (CBT-E), and optimized treatment as usual (TAU-O) in the treatment of adult women with AN.

Method. The analysis was conducted alongside the randomized controlled Anorexia Nervosa Treatment of OutPatients (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 in-patient treatment and the percentage of patients who required in-patient treatment were considerably lower in both intervention groups. The unadjusted ICERs indicated FPT and CBT-E to be dominant compared with TAU-O. Moreover, FPT was dominant compared with CBT-E. CEACs showed that the probability for cost-effectiveness of FPT compared with TAU-O and CBT-E was $\geq 95\%$ if the WTP per recovery was $\geq €9825$ and $\geq €24\,550$, 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.

Received 2 February 2016; Revised 23 June 2016; Accepted 19 July 2016; First published online 9 September 2016

Key words: Anorexia nervosa, cognitive therapy, cost effectiveness, psychodynamic psychotherapy.

Introduction

Anorexia nervosa (AN) is a serious illness with substantial morbidity and mortality (Zipfel *et al.* 2014; Zerwas

et al. 2015). 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 hospitalizations and lost productivity generate substantial personal and societal costs (Stuhldreher *et al.* 2012).

Whilst out-patient psychotherapy is the recommended first-line treatment for most adults with AN, available evidence from randomized controlled trials (RCTs) investigating different therapies has not

* Address for correspondence: N. Egger, Department of Health Economics and Health Services Research, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany.
(Email: n.egger@uke.de)

† N. Egger and B. Wild share the first authorship of this paper.

‡ W. Herzog and H.-H. König are the senior authors of the paper.

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 (FPT), 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 out-patient treatment was more effective and less costly than both alternatives (Byford *et al.* 2007).

Here we present a cost-effectiveness analysis of the Anorexia Nervosa Treatment of OutPatients (ANTOP) study, the world-wide largest ($n=242$) multi-centre RCT comparing the efficacy of FPT, enhanced cognitive-behavioural therapy (CBT-E) and optimized 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. hospitalizations, out-patient 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.

Method

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/m² 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 randomized to either FPT ($n=80$), CBT-E ($n=80$) or TAU-O ($n=82$). Data on clinical outcomes, health care utilization 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 out-patient 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 (Legenbauer & Vocks, 2005; Fairburn, 2008; Friederich *et al.* 2014; 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 analysed at the principal investigators' centres (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 out-patient 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 1-year follow-up were compared between the three groups. In addition, full recovery was defined as having a BMI $>17.5 \text{ kg/m}^2$ and a score on the psychiatric status rating (PSR) scale 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 & 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 AN) to 6 (patient has severe symptoms of AN that require admission). A score of 5 indicates that all Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria for AN have been fulfilled. At baseline, all patients were assessed with a PSR score of 4 (subsyndromal AN, 45%) or 5 (full syndromal AN, 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 three-level EuroQol Five Dimensions Questionnaire (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 1 year in perfect health corresponds to 1 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 in-patient treatment for AN was retrieved from monitoring data and hospital records to ensure that all in-patient stays within the observation period of 22 months were included.

Further health care utilization 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.* (2000) and adapted to the specific requirements in this study. The questionnaire covered out-patient 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 health care system (Krauth *et al.* 2005; Deutsche Krankenhausgesellschaft, 2009; Rote Liste Service GmbH, 2008; Statistisches Bundesamt, 2008, 2009a, 2009b, 2009c). Informal care was valued using the opportunity cost approach, i.e. hours with informal care were considered as lost leisure time (Statistisches Bundesamt, 2010). 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 km according to the tax-deductible rate for work-related travel in Germany.

The costs of out-patient 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, 2009c). Costs of presenteeism, i.e. costs of reduced productivity at the workplace, were calculated according to 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, 2010).

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 & Groothuis-Oudshoorn, 2011). At the 1-year follow-up, 73 participants (30%) had dropped out and the maximum percentage of missings per variable was 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 a 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 1-year follow-up were available ($n=156$). Additionally, we conducted intention-to-treat (ITT) analyses including all randomized patients ($n=242$) in order to investigate the robustness of the results.

Since the majority of costs in AN patients results from hospitalizations, 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 a 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 (Zethraeus *et al.* 2003; Hoch *et al.* 2006) 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 with TAU-O. A linear contrast was calculated to analyse the cost-effectiveness of FPT compared with CBT-E. Age, illness duration (≤ 6 years *v.* >6 years), co-morbid 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 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 & 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., USA). IVEware (Imputation and Variance Estimation Software), a SAS-based application developed by Raghunathan *et al.* (2001), was employed to perform MICE.

Table 1. Baseline characteristics

	Base case (n = 156)	FPT (n = 54)	CBT-E (n = 62)	TAU-O (n = 40)	ITT (n = 242)
Mean age, years (s.d.)	27.0 (7.8)	28.5 (8.7)	25.9 (7.1)	26.9 (7.2)	27.2 (8.2)
Employed, n (%)	76 (49)	28 (52)	29 (47)	19 (48)	118 (49)
Health insurance, n (%)					
Statutory	120 (77)	41 (76)	48 (78)	31 (78)	189 (79)
Private	34 (22)	13 (24)	14 (22)	7 (18)	39 (16)
Mean BMI, kg/m ² (s.d.)	16.7 (0.99)	16.6 (0.96)	16.8 (1.02)	16.8 (1.00)	16.7 (0.99)
Subtype, n (%)					
Binge/purge	71 (46)	26 (48)	27 (44)	18 (45)	111 (46)
Restrictive	85 (54)	28 (52)	35 (56)	22 (55)	131 (54)
Illness duration >6 years, n (%)	57 (37)	22 (41)	19 (31)	16 (40)	92 (38)
Specific co-morbid diagnoses, n (%)					
Affective disorder	35 (22)	10 (19)	19 (31)	6 (15)	57 (24)
Anxiety disorder	36 (23)	9 (17)	16 (26)	11 (28)	59 (25)
Somatization disorder	3 (2)	1 (2)	2 (3)	0 (0)	5 (2)
Mean EQ-5D index score (s.d.)	0.81 (0.24)	0.83 (0.19)	0.80 (0.22)	0.77 (0.22)	0.65 (0.28)

FPT, Focal psychodynamic psychotherapy; CBT-E, enhanced cognitive-behavioural therapy; TAU-O, optimized treatment as usual; ITT, intention to treat; s.d., standard deviation; BMI, body mass index; EQ-5D, EuroQol Five Dimensions Questionnaire.

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 drop-out rates differed significantly between the treatment groups, with the highest drop-out rate observed in TAU-O (51%), and the lowest in CBT-E (22%). This led to significant differences between groups regarding illness duration, co-morbid affective disorder and co-morbid anxiety disorder. Therefore, we controlled for illness duration and co-morbid conditions in the adjusted analyses.

Costs

The unadjusted mean costs 3 months prior to baseline and the unadjusted mean total costs for the observation period of 22 months are reported in Table 2. Mean intervention costs accrued to €2613 (s.d. = 98) in FPT and €2494 (s.d. = 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 hospitalizations accounted for 45, 55 and 70% of the direct costs in FPT, CBT-E and TAU-O, respectively. This

corresponds to a mean duration of admission of 18.4 (s.d. = 52.9) days in FPT, 26.5 (s.d. = 53.9) days in CBT-E and 32.6 (s.d. = 57.3) days in TAU-O, respectively. In FPT 19% of the patients required in-patient treatment, 29% in CBT-E, and 40% in TAU-O. However, none of the differences was significant.

Effectiveness

At the end of the observation period, 35.2% of the patients in FPT were defined as recovered compared with 21% in CBT-E and 12.5% in TAU-O. The difference between FPT and TAU-O was significant ($p = 0.036$). The mean BMI was 18.2 (s.e. = 0.24) kg/m² in FPT, 18.1 (s.e. = 0.23) kg/m² in CBT-E and 17.9 (s.e. = 0.26) kg/m² in TAU-O. With regard to the secondary outcome, quality of life, mean QALYs were 1.53 (s.e. = 0.41) in FPT, 1.48 (s.e. = 0.54) in CBT-E and 1.44 (s.e. = 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 with TAU-O and CBT-E, i.e. it was more effective at lower costs. CBT-E also was dominant compared with TAU-O. Unadjusted ICERs are presented in Table 3.

Table 2. Unadjusted mean costs in Euro (year 2008 values) 3 months prior to baseline and total costs after 22 months (n = 156)

Cost category	Baseline			Total costs after 22 months		
	FPT	CBT-E	TAU-O	FPT	CBT-E	TAU-O
In-patient treatment for AN	2502 (908)	1999 (651)	3927 (1097)	4620 (1814)	6663 (1723)	8220 (2281)
Out-patient mental health care ^a	82 (27)	42 (12)	40 (17)	3231 (188)	3334 (207)	3339 (368)
Other direct costs	442 (166)	233 (40)	270 (100)	2201 (438)	1829 (359)	1890 (376)
All direct costs	3026 (999)	2274 (658)	4238 (1113)	10 052 (2047)	11 826 (1850)	13 448 (2351)
Absenteeism	942 (396)	1100 (390)	1861 (638)	2780 (1032)	2402 (950)	3286 (1470)
Presenteeism	905 (284)	1263 (357)	634 (258)	8680 (2474)	10 462 (2737)	8092 (2824)
Direct + indirect costs	4873 (1328)	4636 (980)	6733 (1512)	21 512 (3625)	24 690 (3538)	24 827 (4697)

Data are given as mean (standard deviation).

FPT, Focal psychodynamic psychotherapy; CBT-E, enhanced cognitive-behavioural therapy; TAU-O, optimized treatment as usual; AN, anorexia nervosa; S.D., standard deviation.

^a Including mean intervention costs of €2613 (s.d. = 98) in FPT and €2494 (s.d. = 102) in CBT-E.

Table 3. Unadjusted point estimates of the incremental cost-effectiveness ratios for the base and ITT samples and different outcomes in Euro at the 12-month follow-up

Cost/effect	FPT v. TAU-O	CBT-E v. TAU-O	FPT v. CBT-E
Base case (n = 156)			
Direct/recovery	Dominant	Dominant	Dominant
Direct/QALYs	Dominant	Dominant	Dominant
Direct/BMI	Dominant	Dominant	€11 600 ^a
(Direct + indirect)/recovery	Dominant	Dominant	Dominant
ITT analyses (n = 242)			
Direct/recovery	Dominant	€23 704	Dominant
Direct/QALYs	Dominant	€11 223	Dominant
Direct/BMI	€11 464	Dominated	€30 056 ^a
(Direct + indirect)/recovery	Dominant	€179 407	Dominant

ITT, Intention to treat; FPT, focal psychodynamic psychotherapy; TAU-O, optimized treatment as usual; CBT-E, enhanced cognitive-behavioural therapy; QALYs, quality-adjusted life years; BMI, body mass index.

^a FPT was associated with lower mean costs and lower mean BMI; thus the interpretation is more straightforward if the relationship is presented inversely, i.e. CBT was associated with additional costs of €11 600 or €30 056 per gained BMI point compared with FPT.

Due to the differences in baseline costs and covariates, the adjusted results from the net-benefit regressions are more informative. The CEACs in Fig. 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 $\geq €9825$ per recovered patient, FPT was certainly cost-effective, i.e. the probability was above the probability threshold of 95%. Comparing FPT with CBT-E, the required WTP was $\geq €24 550$. In contrast, CBT-E could not be considered cost-effective compared with TAU-O; the probability remained below 90% for all considered WTPs.

Secondary analyses

When QALYs were employed as an alternative outcome, using the base case sample, the unadjusted results were similar (Table 3): FPT was dominant compared with TAU-O and CBT-E; CBT-E was dominant compared with TAU-O. In the adjusted analyses, the probability of being cost-effective was 82% for FPT compared with TAU-O at the common WTP of €50 000 per QALY, and 75% compared with CBT-E (Fig. 2). For CBT-E compared with TAU-O the respective probability was 62%.

BMI at the end of the observation period was also employed as a measure of effect in a secondary

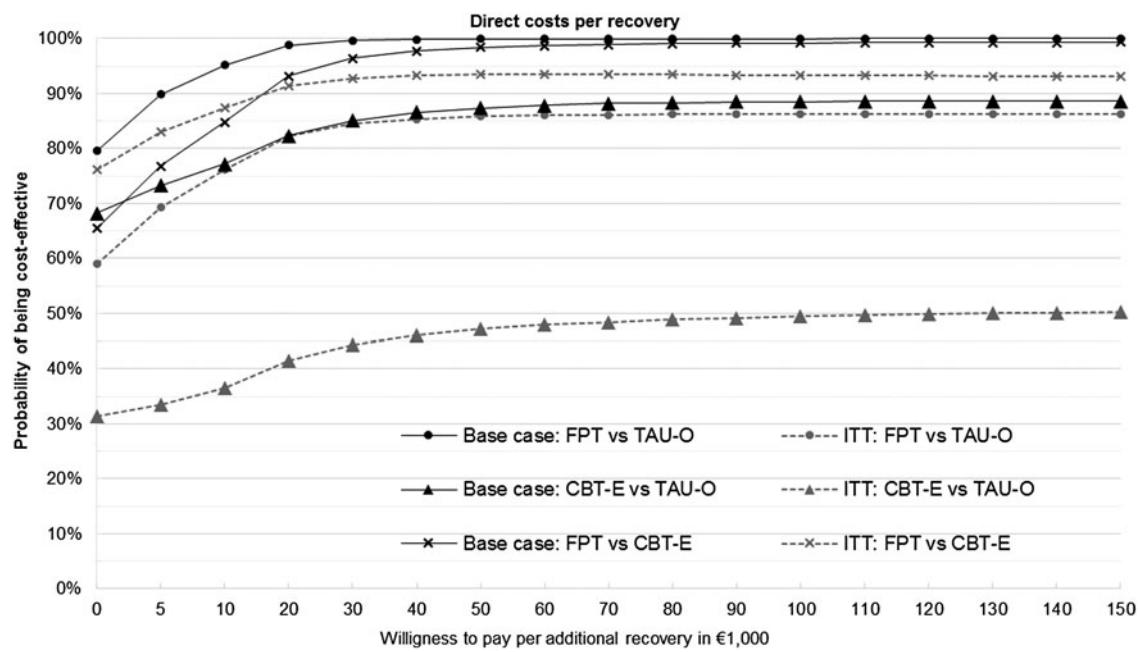


Fig. 1. Cost-effectiveness acceptability curves based on direct costs per recovery after 22 months for the base case sample ($n=156$) and the intention-to-treat (ITT) sample ($n=242$). FPT, Focal psychodynamic psychotherapy; TAU-O, optimized treatment as usual; CBT-E, enhanced cognitive-behavioural therapy.

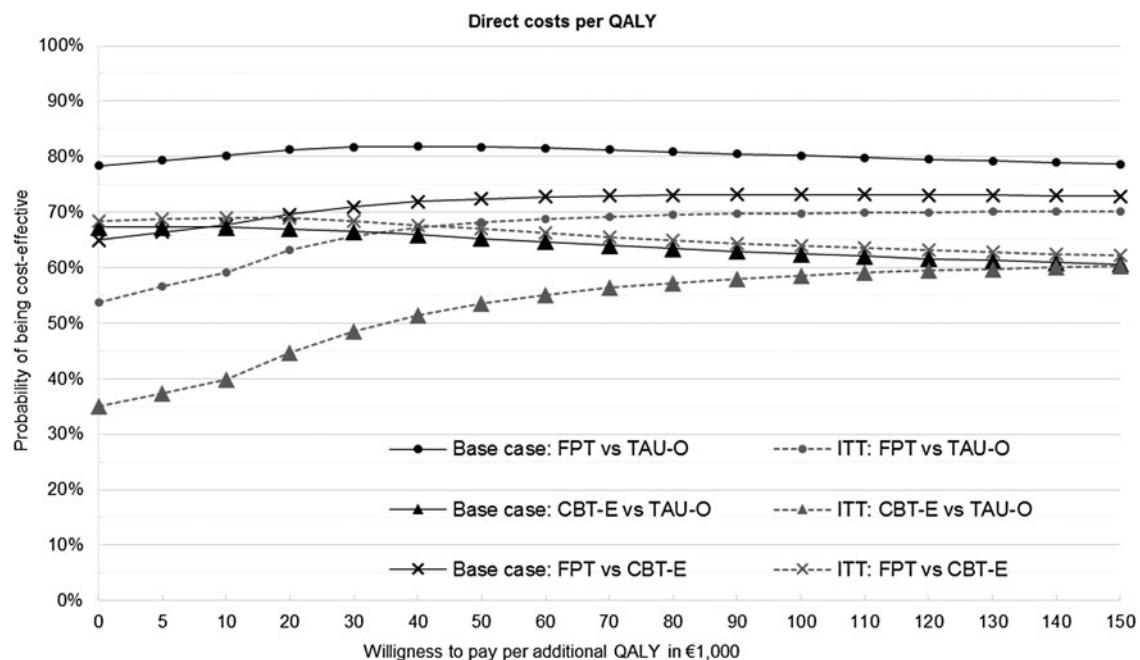


Fig. 2. Cost-effectiveness acceptability curves based on direct costs per quality-adjusted life year (QALY) after 22 months for the base case sample ($n=156$) and the intention-to-treat (ITT) sample ($n=242$). FPT, Focal psychodynamic psychotherapy; TAU-O, optimized treatment as usual; CBT-E, enhanced cognitive-behavioural therapy.

analysis using $n=169$ patients with available data at the end of the observation period. Here, in the unadjusted ICERs (Table 3) FPT and CBT-E were dominant compared with TAU-O. FPT was associated with

lower mean costs and lower mean effects compared with 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

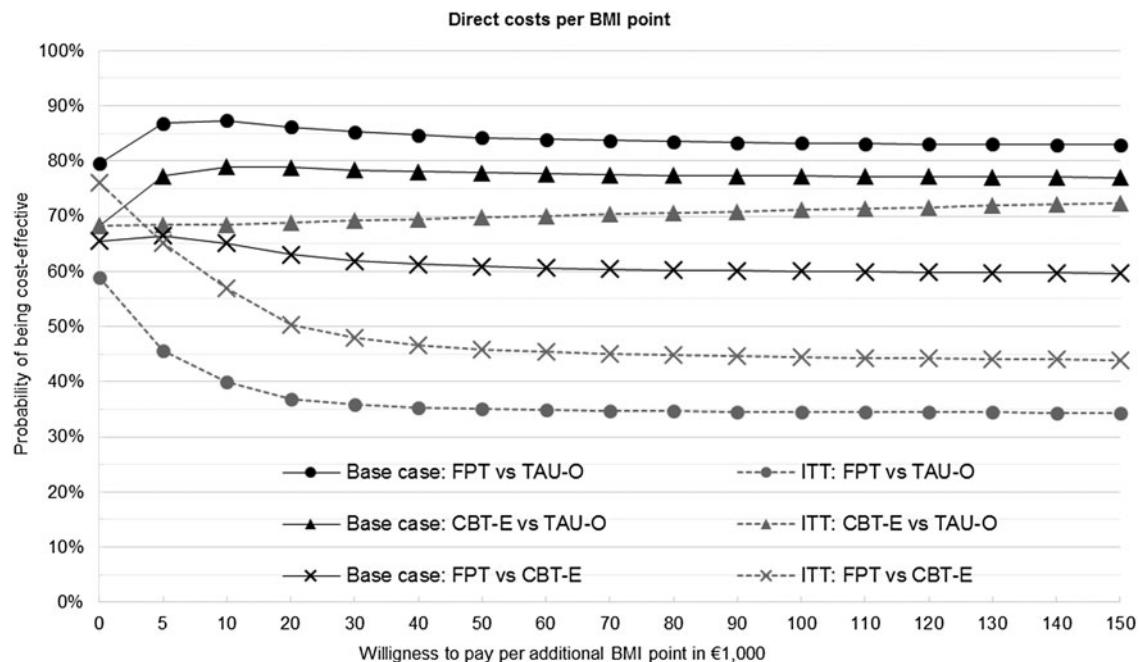


Fig. 3. Cost-effectiveness acceptability curves based on direct costs per body mass index (BMI) point after 22 months for the base case sample ($n=169$) and the intention-to-treat (ITT) sample ($n=242$). FPT, Focal psychodynamic psychotherapy; TAU-O, optimized treatment as usual; CBT-E, enhanced cognitive-behavioural therapy.

per gained BMI point compared with FPT. In the adjusted analyses, at a WTP of €5000 for instance, the probability of being cost-effective was 87% for FPT, 77% for CBT-E, in comparison with TAU-O. The respective probability of FPT compared with CBT-E was 67%. **Fig. 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.

In a further analysis, we included indirect costs due to productivity loss. Regarding the unadjusted ICERs, both interventions remained dominant compared with TAU-O and FPT was dominant compared with CBT-E (**Table 3**). The adjusted CEACs (data not shown) of FPT and CBT-E compared with TAU-O were similar to the base case, but increased more slowly. In contrast, the probability of FPT being cost-effective compared with 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 with CBT-E and TAU-O. CBT-E was associated with additional costs of €23 704 per

recovery compared with TAU-O (**Table 3**). In the adjusted analyses (**Fig. 1**), FPT's probability of being cost-effective compared with TAU-O was $<90\%$ for all WTPs. For FPT compared with CBT-E, the probability remained $<95\%$ for all WTPs. At WTPs $<€20\,000$, FPT's probability of being cost-effective compared with 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 out-patient 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 of follow-up. Although differences in costs were not significant, the absolute costs of in-patient treatment and the percentage of patients who required in-patient 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 with TAU-O as well as in comparison with CBT-E, depending on a decision maker's maximum WTP. CBT-E compared with TAU-O had a lower probability of being cost-effective. The analysis including direct and indirect costs showed a similar trend compared with 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 with TAU-O. 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 with CBT-E and TAU-O in the treatment of adult out-patients with AN. The cost-effectiveness of CBT-E compared with 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.* (2007) reported that specialist out-patient treatment based on individual CBT and parental counselling was cost-effective compared with in-patient treatment, as well as compared with TAU provided by community child and adolescent mental health services. Here, the differences in costs also resulted primarily from differences in in-patient costs. Participants randomized to TAU required almost as much in-patient treatment as participants initially allocated to in-patient treatment. These results are not comparable with our findings because of different settings and study samples; however, they support the cost-effectiveness of specialist out-patient 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 has included any economic evaluation. Therefore, the particular characteristics of FPT, which led to higher recovery rates and reduced hospitalizations, are still unknown. Systematic research is warranted to investigate the specific impact of each intervention on health care utilization.

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 in-patients with severe AN (Abbate-Daga *et al.* 2014). In this study, patients' quality of life improved during hospitalization 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 1-year follow-up, the index scores had decreased below the respective baseline values in each group. This might indicate that improvements in HrQoL 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 utilization covered only 3 months prior to each measurement to avoid any recall bias. Moreover, we observed missing data in single variables of health care utilization at the different measurement time points as well as complete drop-outs 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 drop-out 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 1-year follow-up, were available. This limits the generalizability of our results. Moreover, if reasons for drop-out 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 manualized treatments and particularly FPT seem to provide a cost-effective strategy for improving weight and eating disorder pathology.

Acknowledgements

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

Declaration 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**, 285–291.

Ackard DM, Richter S, Egan A, Engel S, Cronemeyer CL (2014). The meaning of (quality of) life in patients with eating disorders: a comparison of generic and disease-specific measures across diagnosis and outcome. *International Journal of Eating Disorders* **47**, 259–267.

Azur MJ, Stuart EA, Frangakis C, Leaf PJ (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research* **20**, 40–49.

Brooks A, Hagen SE, Sathyaranayanan S, Schultz AB, Edington DW (2010). Presenteeism: critical issues. *Journal of Occupational and Environmental Medicine* **52**, 1055–1067.

Byford S, Barrett B, Roberts C, Clark A, Edwards V, Smethurst N, Gowers SG (2007). Economic evaluation of a randomised controlled trial for anorexia nervosa in adolescents. *British Journal of Psychiatry* **191**, 436–440.

Chisholm D, Knapp MR, Knudsen HC, 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–s33.

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

de Jong P, Heller GZ (2008). *Generalized Linear Models for Insurance Data*. Cambridge University Press: Cambridge, UK.

Dejong H, Broadbent H, Schmidt U (2012). A systematic review of dropout from treatment in outpatients with anorexia nervosa. *International Journal of Eating Disorders* **45**, 635–647.

Deutsche Krankenhausgesellschaft (2009). *Bestandsaufnahme zur Krankenhausplanung und Investitionsfinanzierung in den Bundesländern*. Deutsche Krankenhausgesellschaft: Berlin.

Dolan P (1997). Modeling valuations for EuroQol health states. *Medical Care* **35**, 1095–1108.

Fairburn CG (2008). *Cognitive Behavior Therapy and Eating Disorders*. The Guilford Press: New York and London.

Fenwick E, O'Brien BJ, Briggs A (2004). Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. *Health Economics* **13**, 405–415.

Fichter M, Quadflieg N (2001). The Structured Interview for Anorexic and Bulimic Disorders for DSM-IV and ICD-10 (SIAB-EX): reliability and validity. *European Psychiatry* **16**, 38–48.

Friederich HC, Herzog W, Wild B, Zipfel S, Schauenburg H (2014). *Anorexia Nervosa. Fokale psychodynamische Psychotherapie*. Hogrefe: Göttingen.

Graham JW, Olchowski AE, Gilreath TD (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science* **8**, 206–213.

Gray AM (2011). *Applied Methods of Cost-Effectiveness Analysis in Health Care*. Oxford University Press: Oxford.

Hoch JS, Briggs AH, Willan AR (2002). Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Economics* **11**, 415–430.

Hoch JS, Rockx MA, Krahn AD (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**, 68.

Kessler RC, Ames M, Hymel PA, Loeppke R, McKenas DK, Richling DE, Stang PE, Ustun TB (2004). Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *Journal of Occupational and Environmental Medicine* **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, 736–746.

Legenbauer T, Vocks S (2005). *Manual der kognitiven Verhaltenstherapie bei Anorexie und Bulimie*. Springer: Berlin.

Nan L, Johnson JA, Shaw JW, Coons SJ (2007). A comparison of EQ-5D index scores derived from the US and UK population-based scoring functions. *Medical Decision Making* 27, 321–326.

National Collaborating Centre for Mental Health (2004). National Institute for Health and Clinical Excellence: guidance. In *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders*, pp. 64–65. British Psychological Society: UK, The British Psychological Society & The Royal College of Psychiatrists: Leicester.

Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger P (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology* 27, 85–95.

Rote Liste Service GmbH (2008). *Rote Liste 2008. Arzneimittelverzeichnis für Deutschland*. Rote Liste Service GmbH: Frankfurt am Main.

Statistisches Bundesamt (2008). *Pflegestatistik 2007*. Statistisches Bundesamt: Wiesbaden.

Statistisches Bundesamt (2009a). *Kostennachweis der Krankenhäuser 2008*. Statistisches Bundesamt: Wiesbaden.

Statistisches Bundesamt (2009b). *Grunddaten der Krankenhäuser 2008*. Statistisches Bundesamt: Wiesbaden.

Statistisches Bundesamt (2009c). *Verdienste und Arbeitskosten*. Statistisches Bundesamt: Wiesbaden.

Statistisches Bundesamt (2010). *Volkswirtschaftliche Gesamtrechnungen 2009*. Statistisches Bundesamt: Wiesbaden.

Stuhldreher N, Konnopka A, Wild B, Herzog W, Zipfel S, Lowe B, Konig HH (2012). Cost-of-illness studies and cost-effectiveness analyses in eating disorders: a systematic review. *International Journal of Eating Disorders* 45, 476–491.

van Buuren S, Groothuis-Oudshoorn K (2011). mice: Multiple imputation by chained equations in R. *Journal of Statistical Software* 45, 1–67.

Wild B, Friederich HC, Gross G, Teufel M, Herzog W, Giel KE, 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 JE, Kearns B, Relton C, Smith C, Cooper CL (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. *European Journal of Health Economics* 16, 141–151.

Zerwas S, Larsen JT, Petersen L, Thornton LM, Mortensen PB, Bulik CM (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 KE, Bulik CM, Hay P, Schmidt U (2015). Anorexia nervosa: aetiology, assessment, and treatment. *Lancet Psychiatry* 2, 1099–1111.

Zipfel S, Wild B, Gross G, Friederich HC, Teufel M, Schellberg D, Giel KE, de Zwaan M, Dinkel A, Herpertz S, Burgmer M, Löwe B, Tagay S, von Wietersheim J, Zeeck A, Schade-Brittinger C, Schauenburg H, Herzog W; ANTOP study group (2014). Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study): randomised controlled trial. *Lancet* 383, 127–137.