Cost-Effectiveness of Short-term Inpatient Psychotherapy based on Transactional Analysis in Patients with Personality Disorders

Running title: Cost-effectiveness of Short-Term Inpatient Psychotherapy in PD patients

Voorstel auteurs(volgorde):

Eva K. Horn [a, b] Roel Verheul [c, d] Moniek Thunnissen [e, f] Jos Delimon [c] Mirjam Soons [g] Anke M. M. A. Meerman [h] Uli M. Ziegler [i] Bert V. Rossum [j] Theo Stijnen [k] Paul M. G. Emmelkamp [d, l] Jan J. V. Busschbach [a, b]

[a] Viersprong Institute for Studies on Personality Disorders (VISPD), Bergen op Zoom; [b] Department of Psychiatry, section Medical Psychology and Psychotherapy, Erasmus MC, Rotterdam; [c] De Viersprong, Netherlands Institute for Personality Disorders, Halsteren; [d] Department of Clinical Psychology, University of Amsterdam, Amsterdam; [e] GGZ WNB, Bergen op Zoom; [f] Private Practice, Bergen op Zoom; [g] PuntP, Amsterdam; [h] Centre of Psychotherapy Pro Persona, Lunteren; [i] Zaans Medical Centre, Zaandam; [j] Altrecht, Zeist; [k] Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, The Netherlands, [l] The Center for Social and Humanities Research, King Abdulaziz University, Jeddah, Saudi Arabia

Summary (169 words)

Context

Short-term inpatient psychotherapy with transactional analyses (STIP-TA) in patients with personality disorders (PD) showed a higher effectiveness than other psychotherapies (OP).

Objective

To assess the cost-effectiveness of STIP-TA compared to OP in PD patients over three years follow-up.

Methods

Matched controlled study on the basis of the propensity score. Cost-effectiveness analyses were carried out from a societal perspective. Healthcare costs and lost production costs were measured using the TiC-P. Cost-effectiveness was represented by costs per quality adjusted life years (QALYs) as measured by the EQ-5D. Uncertainty was assessed using bootstrapping.

Results

Mean three year costs were €59,834 for STIP-TA and €69,337 for OP, a difference of €-9,503 (95% CI -32,561 to 15,726). QALYs were 2.29 for STIP-TA and 2.05 for OP, a difference of .24 (95% CI .05 to .44). The STIP-TA intervention was dominant, thus less costly and more effective than OP. Sensitivity analyses displayed similar results.

Conclusions

In terms of cost per QALY and compared to OP, STIP-TA is a cost-effective treatment in PD patients.

Declaration of Interest

EKH, RV, MT, JD, and JVB have worked or are still working for the Viersprong, Netherlands Institute for Personality Disorders who offers STIP-TA. MT is a trainer and supervisor of TA. MT, JD, MS, AMMAM, UMZ, and BVR have been or are still involved with the control conditions.

Introduction

Personality disordered (PD) patients are a highly prevalent group with a high individual and societal burden [1]. PD patients report a quality of life compared to patients with a serious somatic illness and the presence of a PD is associated with high costs for society [2-4]. Several psychotherapeutic treatments have proven their effectiveness in the treatment of PD patients and psychotherapy is considered to be the first choice in treatment [5-8]. State-of-the-art cost-effectiveness analyses (CEAs) on psychotherapeutic treatments in PD populations revealed differences in cost-effectiveness between treatments [9-15]. The evidence of cost-effective treatments in this population is still scarce given the large number of different modalities of psychotherapy involved in these studies. Furthermore, more evidence is essential to make deliberate decisions for reimbursed treatments in restricted healthcare budgets [16]. Soeteman et al. showed that short-term inpatient treatments in general were most cost-effective compared to other treatment modalities in a cluster C PD population [14]. Short-Term Inpatient Psychotherapy based on Transactional Analysis (STIP-TA) was included in that study, and was the shortest treatment of the short-term inpatient modality. An earlier trial compared different types of aftercare following a STIP-TA treatment and found a substantial decrease of symptoms at the end of the STIP-TA treatment [17]. In a matched controlled effectiveness study, STIP-TA was compared to other specialized psychotherapies (OP) in PD patients and proved to be an effective treatment option over a three year follow-up [18]. The current study extends this matched controlled effectiveness study by a cost-effectiveness analysis, with the outcome expressed in costs per quality-adjusted life year (QALY). As the more effective STIP-TA is also likely to be more costly then less intensive alternatives, a cost-effectiveness study on this comparison was warranted. Earlier research revealed that the higher medical costs of short-term inpatient therapies are offset by savings in other part of the health care system [14]. However, all short-term inpatient treatments were combined and the cost-effectiveness of STIP-TA could not be determined. In this study we set out to evaluate the cost-effectiveness of STIP-TA from a societal perspective using a matched controlled design.

Methods

Study population and design

The current study is part of the SCEPTRE study, a large naturalistic study about the cost-effectiveness of psychotherapeutical treatments in PD patients [19]. Six mental health centres took part in the SCEPTRE study of which all offered psychotherapeutic treatments for PD patients. They offered a wide range of various types of treatments concerning duration (i.e. 3 to 36 months), setting (i.e. outpatient, day hospital, or inpatient treatment), format (i.e. individual or group meetings), as well as theoretical orientation (e.g. cognitive-behavioural, psychodynamic). Eight hundred thirty-seven patients gave informed consent for the study and were enrolled of which 143 patients could not be followed due to several reasons (refusal, loss to follow-up). This left 694 patients of which 71 received STIP-TA at De Viersprong, Netherlands Institute for Personality Disorders (one of the six mental health centres), and 623 patients who received other specialized psychotherapies (OP) at one of the six mental health centres. It was not possible to find a good match for four STIP-TA patients, therefore 67 patients were matched 1-by-1 by the logit of the propensity score (PS), which left 134 patients for this trial (see figure 1). The PS is defined as the conditional probability of assignment to one of two treatment groups given a set of observed pre-treatment variables [20, 21]. To estimate the PS, we fitted one logistic regression model with relevant confounders, and medical and productivity costs at baseline as independent variables, and group membership (STIP-TA or OP) as outcome variable [22]. Nearest neighbourhood matching using a caliper distance of .2 of the standard deviation of the logit of the PS was used [23]. The construction and characteristics of this dataset were described in more detail in the article describing the effectiveness of STIP-TA [18].

--insert figure 1 about here–

In this study, STIP-TA is compared to other psychotherapy (OP). STIP-TA is based on the ideas of transactional analysis (e.g. [24, 25]) and has integrated these ideas in a short (13 weeks) inpatient treatment [18, 26, 27]. STIP-TA includes psychotherapy, psychomotor and art therapy, sociotherapy and milieu therapy. This program is designed for patients with personality disorders, mainly cluster C PD or personality disorder not otherwise specified (PDNOS).

The matched OP group (see above) consisted of six patients (9%) following short-term (up to six months) outpatient therapy and 18 patients (27%) following long-term (longer than six months) outpatient therapy; seven patients (10%) following short-term day hospital therapy and 11 patients (16%) following long-term day hospital therapy; and nine patients (13%) following short-term inpatient therapy and 16 patients (24%) following long-term inpatient therapy. Mean length of short-term treatments was 5.8 (SD=.70) months, long-term treatments lasted 13.4 (4.7) months on average. Twenty-two percent of treatments had a cognitive-behavioural orientation, 30% had a psychodynamic orientation, and 42% consisted of integrative treatments (e.g. cognitive-behavioural/psychodynamic). The orientation of the remaining 5% was not specified.

Handling data

All outcome measures were assessed at baseline and several follow-up points up to five years after baseline. Due to logistic reasons, three treatment centres had their assessments at baseline, end of the treatment, and six, and 12 months after the end of treatment, and at 36 and 60 months after baseline, whereas three other centres had their assessments at baseline, at 12, 24, 36, and 60 months after baseline. Since the period between the last two follow-up points was rather long (two years), and a rising number of missing data appeared, we limited the analyses to a time horizon of three years after baseline.

The expectation-maximization algorithm was used to impute missing quantitative baseline data. For missing categorical data, multinomial logistic regression was used. Since the distribution of the costs was skewed, multiple imputation with ten imputations was used in the case of intermittent missing data [28]. In this multiple imputation, cost data are considered semi-continuous data, which is characterized by a mixture of a considerable proportion of zero values and a skewed distribution of positive values. Using these assumptions, the predictive mean matching (PMM) approach was used to impute the missing values. PMM can account for the semi-continuous nature and therefore skewed distribution of costs and ensures that imputed values are plausible [29]. In PMM, the imputed variable takes on the value of one of a set of nearest observed values in the dataset [30]. The imputed datasets were further analyzed conforming to the rules established by Rubin [31, 32]. When missing cost data was due to drop-out, estimates were based on last observation carried forward (LOCF).

The analyses were performed using IBM SPSS 20.0 (IBM Corporation, USA) for data preparation, STATA 12.1 (StataCorp LP, USA) for imputing missing data, and SAS 9.2 (SAS Institute Inc, USA) for bootstrapping.

Assessments

Costs

The economic evaluation was undertaken from a societal perspective, including costs related to medical resource utilization, and costs due to productivity losses [33]. To collect data on costs in other part of health care than the initial intervention and to collect cost related to work (productivity losses), the Trimbos and Institute for Medical Technology Assessment (iMTA) Questionnaire on Costs Associated with Psychiatric Illness (TiC-P) was used [34]. The TiC-P consists of two sections:

The first section covers the medical costs, which were split up in somatic and psychiatric healthcare costs, medication costs and other costs (e.g. alternative healers, domiciliary care). The recall period for the utilization of medical resources was four weeks prior to completion of the questionnaire. For mean somatic and psychiatric healthcare costs, the total number of contacts with medical and psychological/psychiatric healthcare providers (e.g. outpatient visits, length of stay in hospital) multiplied by unit costs of the corresponding health care services [35, 36]. The costs of medication were calculated as the cost price per medication multiplied by the daily dose multiplied by the number of prescription days, plus the pharmacist's dispensing costs of €7,50 per prescription. Other costs were valued according to prices reported in the Dutch manual for cost research [35].

The second section includes the short form Health and Labour questionnaire (SF-HLQ) for collecting data on productivity losses for the two weeks prior to completion of the questionnaire. The SF-HLQ consists of three modules, i.e. (1) absence from work (absenteeism), (2) reduced efficiency at work and (3) the take-over of domestic help [37]. First, concerning absenteeism, the actual costs of hours missed at work and the number of days absent from work due to health-related problems were valued according to the average value added per employee by age and gender per hour and per day, respectively [35]. The number of days absent from work was furthermore divided into short-term and long-term absence. Short-term absence was calculated taking into account the number of days and hours of paid employment of the participant per week. To value long-term absence from work, we

applied the friction-cost method, which takes into account the fact that a formerly unemployed person may replace a person who becomes disabled [38]. The period needed to replace a worker (the so-called friction period) in 2011 was estimated to be 115 days; maximum productivity costs to society were therefore limited to productivity losses during a period of 115 days. This 'friction cost method' is recommended by the Dutch guidelines for cost research [33]. Second, concerning reduced efficiency at work, the number of hours of reduced efficiency at work while being sick was valued by multiplying the number of inefficient days by the number of hours work per day, by a self-reported inefficiency score. This inefficiency score ranged from zero (as efficient as in good health) to one (absolutely inefficient). Third, patients can be unable to perform their domestic tasks because of illness. These domestic tasks can be taken over by paid or unpaid help. Paid work is valued at €26.28 (standard price of domiciliary care), unpaid work at €12.95 (standard price of cleaning help) [35].

Furthermore, patients were not able to work during the hours or days being in treatment. These costs (sick leave) had to be added on top of the regular productivity costs. For unemployed patients these costs were set to zero, since they cannot produce any productivity costs for society. When employed patients reported oneself ill before the start of the treatment and stayed ill during the whole treatment period, the sick leave costs due to the intervention were set to zero since these costs were already accounted for in the regular productivity costs. Concerning the remaining employed patients in day hospital or inpatient psychotherapy, days in treatment were multiplied by the number of hours worked per day multiplied by the net income per hour. For patients in outpatient psychotherapy, first the number of hours worked per day were divided by eight since a regular working day consists of eight hours. These hours were then multiplied by the hours in treatment multiplied by the net income per hour.

Since the assessment at the end of treatment did not consist of the TiC-P, the six months after start of treatment was missing for all patients. As it was not possible to impute this time point, a different method had to be used.

First, medical costs of the first six months were set to the level of the 12 months measure. It was assumed that patients in day hospital and inpatient psychotherapy did not generate any additional medical costs during their psychotherapeutic treatments since all medical costs were already

accounted for in the treatment costs. Therefore we set the additional medical costs during day hospital and inpatient treatment to zero. Since we did not expect any alterations in medical costs during the waiting list period compared to baseline, medical costs at the waiting list period were set to the baseline level.

Second, costs related to absenteeism at six months were 'filled in' by the costs of sick leave because of the treatment under study (see above). When later questionnaires indicated that e.g. patients were absent from work from baseline until three years later and at baseline they were already absent from work 115 days or longer, all intermediary costs related to work absence were set to zero in order to avoid double counting of costs. Additionally if patients were absent e.g. at baseline and 12 months after baseline (>115 days) and absentness at 6 months was missing, costs of 6 months' absence was set to the level of the 12 months' costs and 12 months absence costs were set to zero following the line of reasoning of the friction costs method. Costs related to reduced efficiency at work and domestic help were set to the level of the 12 months measure.

Reference unit prices of health care services were applied and adjusted to the year of the study (2011) according to the consumer price index [39]. For medical and productivity costs, we assumed that the recall period of the TiC-P was representative of the six months prior to the assessment. Costs were discounted at an annual rate of 4%, effectiveness as measured by QALYs were discounted at a rate of 1.5%, as recommended by the Dutch guidelines for cost research [35]. The uncertainty around the mean costs and QALY scores was assessed with bootstrap simulations with 1000 replications for the imputed datasets. The uncertainty interval was represented by the 2.5^{th} and 97.5^{th} percentiles. The results were presented in a cost-effectiveness plane and an acceptability curve [40, 41].

Effects

For the economic evaluation, the effects were measured in health-related quality of life years (QALYs) using the EQ-5D-3L [42]. The EQ-5D is a standardised instrument and has shown to be sensitive to change in PD patients [43]. Five health states were measured: 'mobility', 'self-care', 'usual activities', 'pain/discomfort', and 'anxiety/depression'. Possible response levels were 'no problems', 'some or moderate problems' and 'extreme problems or complete inability'. This resulted in 243

different possible health states, which were weighted to obtain a single index score between –0.33 (worst imaginable health state) and 1.00 (best imaginable health state). QALYs were further calculated for the three year follow-up by multiplying the QALY weights by the duration of time spent in that health state (area under the curve). To calculate the mean EQ-5D index values, the Dutch norm scores were used [42, 44, 45].

Sensitivity analyses

To assess the robustness of results under different assumptions, four sensitivity analyses were done. In a first sensitivity analysis, we studied the impact of applying a 3.5% discount rate for both costs and health outcomes as recommended by the UK guideline on health technology assessment of the National Institute for Health and Clinical Excellence (NICE) [46]. A second sensitivity analysis was undertaken using the payer's perspective (insurance perspective): only the medical costs were taken into account. Third, as there were some differences in baseline costs, sensitivity analyses were done by adjusting the cost-effectiveness for differences in baseline costs. Four different types of regression-based methods were carried out (standard, split, trimmed, and replacement regression) with total costs as dependent variable and costs at baseline and treatment group as independent variables [47]. Fourth, to study the influence of missing data, an analysis was done on the complete follow-up data only.

Results

Sample

The sample consisted mainly of patients with a cluster C PD or a PDNOS diagnosis (see table 1). The majority of patients (93% of STIP-TA and 82% of OP) had a history of outpatient treatment and over half (61% of STIP-TA and 51% of OP) had a history of psychotropic drug treatment. Eighty-two percent of STIP-TA and 85% of OP patients reported that they suffered from psychological problems since more than five years. Educational level was high (European Qualifications Framework [EQF]≥6 [48]) for two-thirds of both groups, and about 60% of both groups worked or was studying. The matching on the basis of the propensity score was successful in reducing all differences between background variables as presented in table 1, except for the overall costs due to production losses (see below).

--insert table 1 about here--

Costs and effects

Total costs at baseline were €24,183 for STIP-TA and €21,283 for OP, a difference of €2,900. High somatic and psychiatric costs were primarily due to hospitalizations prior to completion of the questionnaires. Costs were not distributed normally (Kolmogorov–Smirnov test on total costs: df=134, p<.000, see table 1). The overall medical and total societal costs at baseline were not significantly different between the two conditions (one-sample sign test). The overall costs due to production losses, however, differed significantly between both groups (p=.021).

After bootstrap, mean three year costs were €59,834 for STIP-TA and €69,337 for OP, resulting in a non-significant difference in costs of €-9,503 (95% CI -32,212 to 15,692; see table 2). There was a trend that STIP-TA patients generated lower costs associated with healthcare and costs associated with production losses. Within the healthcare costs, we found a trend that STIP-TA patients generated lower costs in psychiatric and other care but more costs in somatic healthcare compared to OP. High somatic and psychiatric costs were mainly due to hospitalizations. In terms of production losses, there was a trend that STIP-TA patients reported being more often absent from work, while OP patients reported less efficiency during work and more problems concerning domestic tasks. The

STIP-TA treatment was (non-significantly) cheaper (mean €19,420, range €4,034 to €32,810) compared to the OP treatments (mean €25,459, range €4,204 to €89,088). Figure 2 shows the distribution of cost categories in STIP-TA and OP patients.

--insert table 2 about here –

--insert figure 2 about here–

Mean three year bootstrapped QALYs were 2.29 for STIP-TA and 2.05 for OP meaning a significant difference in QALYs of .24 (95% CI .05 to .43). Figure 3 shows the course of the QALYs over 36 months.

--insert figure 3 about here–

Cost-Effectiveness

The trend towards lower costs and the significant better effects of STIP-TA, made that STIP-TA was a 'dominant intervention' compared to OP: less costly and more effective. The incremental cost-effectiveness ratio (ICER) for STIP-TA compared with OP was €-41,491, which means that by gaining one additional QALY by delivering STIP-TA instead of OP, one saves about €41,500. To display the impact of uncertainty around the estimated mean costs and QALYs, we made use of a cost-effectiveness plane with the bootstrapped results. Eighty-three percent of the cost-effectiveness pairs lie in the south-east quadrant of the plane, implying lower costs and higher effectiveness, and 17% in the north-east quadrant, implying both higher costs as well as higher effectiveness (see figure 4).

--insert figure 4 about here --

The cost-effectiveness acceptability curve (CEAC) indicates that the probability that STIP-TA is cost-effective is between 80% and 90% until a threshold of approximately €20,000/QALY and achieves 93% with a ceiling ratio of 40,000/QALY (see figure 5).

--insert figure 5 about here--

Sensitivity analyses

Sensitivity analyses displayed minor differences, the dominance of STIP-TA above OP remained (see table 3). Costs for STIP-TA ranged from 37,750 (payer perspective) to 61,732 (replacement regression). In OP patients, costs ranged from 44,908 (payer perspective) to 72,298 (replacement regression). QALYs ranged from 1.97 (OP, 3,5% discount rate) to 2.36 (STIP-TA, complete data). ICERs ranged from -23,717 (trimmed regression) to -49,146 (standard regression).

--insert table 3 about here --

Discussion

In the current study the cost-effectiveness of an intensive short- inpatient treatment was compared to other specialized psychotherapies in patients with PD over a three-year time horizon. The economic evaluation was performed from the societal perspective. Results showed that the STIP-TA treatment is more cost-effective compared to other specialized psychotherapies in the treatment of PDs. STIP-TA patients generated lower costs over a three year time horizon compared to OP, and STIP-TA was significantly more effective than OP in terms of QALYs gained. With a ceiling ratio of 40,000/QALY, the chance that STIP-TA is cost-effective compared to OP is 93%. Sensitivity analyses on different discount rates, analyses from the payer's perspective, corrections for baseline cost differences, and missing data showed similar results. In both groups, about one third of all costs over 36 months were due to the intervention itself, about one quarter were due to absenteeism. Although STIP-TA is primarily a rather expensive treatment, when the costs of treatment were compared to the costs of OP for matched patients, we found that the mean costs of STIP-TA were generally lower. This was probably due to two-thirds of OP patients following (rather expensive) long-term treatments of more than a year on average.

Methodological considerations

A strength of this study is its representativeness and therefore high external validity due to the naturalistic nature and a minimal set of exclusion criteria. The OP group was found to have a great variation in treatments which supports the assumption that OP is representative for psychotherapeutical treatments for PD patients in the Netherlands. Patients were not randomized but were assigned to treatment by clinical knowledge. However, a comparison of such different dosages of treatment by means of a RCT is difficult or even impossible due to patient preferences fo one of the treatments. An earlier attempt to compare STIP-TA to outpatient treatment in cluster C PD patients in a RCT has failed since patients were not willing to participate in a random assignment to outpatient or inpatient treatment (PSILO trial [49]). Furthermore, in economic evaluations, nonrandomization of patients can even be considered an advantage, since the external validity of these studies is higher than that of a RCT [14]. However, generalizability to a PD population at large is limited due to the inclusion of mainly cluster C and PDNOS diagnoses; almost all patients having a history of outpatient treatment; and patients being treatment seeking. These limitations were already discussed in detail elsewhere [18].

We made use of a long-term follow-up of three years. However, measurement points varied over time and between institutions, which led to a high amount of missing data. This made assumptions about the missing healthcare costs and production losses necessary and led to uncertainty about the results. However, we handled missing data with sophisticated analyses (multiple imputation) and carried out a secondary analysis on the complete data to confirm the robustness of results.

Third, we used the TiC-P to estimate costs which has a recall period of two to four weeks and extrapolated these to estimate the costs of patients for six months. The assumption that these weeks are representative for the six months prior to the questionnaire can be questioned and could lead to an over-estimation of costs. A hospitalisation of a number of days prior to the completion of the questionnaire, for example, will lead to huge costs and an over-estimation of these costs, while the same event just after the completion of the questionnaire will not be taken into account and will lead to an under-estimation of these costs. We assumed that the over- as well as the under-estimation of costs will balance each other out on average. A study about the feasibility, reliability, and validity of the

TiC-P in patients with a psychiatric disorder showed the questionnaire to be a feasible and a reliable instrument in this patient group [50].

Implications

The current study showed that STIP-TA is not only a very effective but also a cost-effective treatment option in mainly cluster C PD and PDNOS patients. The results of our study are consistent with several studies showing that psychological treatments tailored to PD patients are generally very effective and that treatments differ in their cost-effectiveness [8, 16]. To make this treatment available to more patients, additional research is recommended. Soeteman et al. had studied the value of further research of cost-effective psychotherapies in cluster B and C PD patients in the Netherlands and concluded that the societal value of additional research is considerable [51].

In this study we compared STIP-TA to a variety of psychotherapies (OP). We found that the OP group used in this study was a rather heterogeneous group which found expression in a wide spread of treatment dosages and therefore treatment costs. It would be of additional value to study whether the dominance of STIP-TA holds when compared to one specific evidenced based (outpatient) psychotherapy for this patient group. An earlier attempt was done to directly compare STIP-TA with outpatient psychotherapy in a RCT. However, patients were not willing to participate in this study due to the random assignment to one of these treatments [49].

It could be worthwhile to extend the study and to analysis whether it is possible to identify specific subgroups of patients in whom STIP-TA is a more cost-effective treatment compared to other (outpatient) treatments in order to be able to give a strong and evidence-based advice. Since STIP-TA is known as a high pressure and destabilizing treatment, rather healthy PD patients are admitted to this treatment. However, a study of van Manen et al (in preparation) showed that destabilizing treatments seemed more effective than stabilizing treatments in terms of symptoms and interpersonal functioning in psychological strong as well as psychological rather weak patients [52]. It is of practical relevance to study whether more seriously disordered patients could also profit from this intensive but short-term kind of treatment and to expand the target population to (relatively mild) cluster A and B PD patients.

Moreover, more knowledge about the efficacious ingredients of this specific treatment is important. The ‘Dodo bird verdict’ states that primarily common factors such as ‘the belief in the

treatment' or 'the therapeutic alliance' account for the effectiveness of psychotherapy [53]. This should lead to equally effective treatments when the common factors were set equal. However, this cannot explain the superiority of STIP-TA above OP and points to other factors which constitute to the positive effects. With the knowledge of these factors, it could be possible to add some of these to other, less expensive and less radical outpatient treatments which is in line with the deinstitutionalisation of mental health care and the emphasis on patient empowerment. The question remains whether it is possible to keep the high pressure used in the inpatient treatment and which seems to be one of the efficacious ingredients in a treatment of lower dosage [52].

Mental health care for PD patients mainly consists of long-term outpatient psychotherapy. Inpatient treatments for rather mild PDs are mostly not the first choice of treatment but are mostly used for crisis intervention and stabilization, especially in BPD patients. A hospitalization of three months, although a relatively short period of time, can be an obstacle for patients with e.g. a paid job or parental responsibility [54], and patients have to be willing to be hospitalized for three months. Furthermore, regarding the current economic climate, less radical and (on the short-term) cheaper outpatient treatments are preferred. Indeed, presently at the Viersprong, the Netherlands Institute for Personality Disorders, STIP-TA is reduced in an attempt to reduce the cost price per treated patient. The present study points to the savings elsewhere in society and even when considering only health care costs, our estimates are that STIP-TA is less costly. Obviously, if the costs of OP are reduced, such results may change. However, as studies like the present one showed that investments in expensive treatments can be earned back on the long-term, these treatments should be continued in spite of the high initial expenses. Cost-effective inpatient treatments should be preserved for at least part of PD patients as an alternative treatment for outpatient psychotherapies and should be adopted in clinical guidelines. The present study invites both clinicians, patients and payers to reconsider the direct medical costs of treatment in the light of favorable results and future savings.

References

1. Coid, J., et al., *Prevalence and correlates of personality disorder in Great Britain.* British Journal of Psychiatry, 2006. **188**: p. 423-31.
2. Soeteman, D.I., et al., *The economic burden of personality disorders in mental health care.* Journal of Clinical Psychiatry, 2008. **69**(2): p. 259-65.
3. Soeteman, D.I., R. Verheul, and J.J. Busschbach, *The burden of disease in personality disorders: diagnosis-specific quality of life.* Journal of Personality Disorders, 2008. **22**(3): p. 259-68.
4. Maclean, J.C., et al., *Mental Health and High-Cost Health Care Utilization: New Evidence from Axis II Disorders.* Health Services Research, 2013: p. n/a-n/a.
5. *Multidisciplinaire Richtlijn Persoonlijkheidsstoornissen*. 2008: Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ.
6. Excellence, N.I.f.H.a.C., *Antisocial Personality disorder. Treatment, management and prevention*, in *NICE clinical guideline*, N.C.C.f.M. health, Editor. 2009, National Institute for Health and Clinical Excellence (NICE): London.
7. Excellence, N.I.f.H.a.C., *Borderline Personality disorder: treatment and management*, in *NICE clinical guideline*, N.C.C.f.M. health, Editor. 2009, National Institute for Health and Clinical Excellence (NICE): London.
8. Dixon-Gordon, K.L., B.J. Turner, and A.L. Chapman, *Psychotherapy for personality disorders.* International Review of Psychiatry, 2011. **23**(3): p. 282-302.
9. Berghout, C.C., J. Zevalkink, and L. Hakkaart-van Roijen, *A Cost-utility Analysis of Psychoanalysis versus Psychoanalytic Psychotherapy.* International journal of technology assessment in health care, 2010. **26**(1): p. 3-10.
10. Palmer, S., et al., *The cost-effectiveness of cognitive behavior therapy for borderline personality disorder: results from the BOSCOT trial.* Journal of Personality Disorders, 2006. **20**(5): p. 466-81.
11. Priebe, S., et al., *Effectiveness and cost-effectiveness of dialectical behaviour therapy for self-harming patients with personality disorder: a pragmatic randomised controlled trial.* Psychotherapy and Psychosomatics, 2012. **81**(6): p. 356-65.
12. Soeteman, D.I., et al., *Cost-effectiveness of psychotherapy for cluster B personality disorders.* British Journal of Psychiatry, 2010. **196**(5): p. 396-403.
13. van Asselt, A.D.I., et al., *Out-patient psychotherapy for borderline personality disorder: cost-effectiveness of schema-focused therapy v. transference-focused psychotherapy.* British Journal of Psychiatry, 2008. **192**(6): p. 450-457.
14. Soeteman, D.I., et al., *Cost-effectiveness of psychotherapy for cluster C personality disorders: a decision-analytic model in the Netherlands.* Journal of Clinical Psychiatry, 2011. **72**(1): p. 51-9.
15. Bamelis, L.L.M., et al., *Economic evaluation for personality disorders: a multicentered randomized trial.* submitted.
16. Soeteman, D.I. and J.J. Kim, *Cost-effectiveness of psychotherapy for personality disorders: treatment recommendations and implementation.* Expert review of pharmacoeconomics & outcomes research, 2013. **13**(1): p. 73-81.

17. Thunnissen, M., et al., *A randomized clinical trial on the effectiveness of a reintegration training program versus booster sessions after short-term inpatient psychotherapy.* Journal of Personality Disorders, 2008. **22**(5): p. 483-495.
18. Horn, E.K., et al., *Effectiveness of Short-Term In-patient Psychotherapy using Transactional Analysis in Patients with Personality Disorders: A Matched Control Study using Propensity Score.* Journal of Personality Disorders, submitted.
19. Bartak, A., et al., *Effectiveness of Different Modalities of Psychotherapeutic Treatment for Patients with Cluster C Personality Disorders: Results of a Large Prospective Multicentre Study.* Psychotherapy and Psychosomatics, 2010. **79**(1): p. 20-30.
20. Rosenbaum, P.R. and D.B. Rubin, *The Central Role of the Propensity Score in Observational Studies for Causal Effects.* Biometrika, 1983. **70**(1): p. 41-55.
21. Bartak, A., et al., *The Use of Propensity Score Methods in Psychotherapy Research. A Practical Application.* Psychotherapy and Psychosomatics, 2009. **78**(1): p. 26-34.
22. Brookhart, M.A., et al., *Variable selection for propensity score models.* American Journal of Epidemiology, 2006. **163**(12): p. 1149-56.
23. Austin, P.C., *Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies.* Pharmaceutical Statistics, 2011. **10**(2): p. 150-61.
24. Berne, E., *Principles of transactional analysis.* Indian Journal of Psychiatry, 1996. **38**(3): p. 154-9.
25. Barnes, G., *Not Without the Couch: Eric Berne on Basic Differences between Transactional Analysis and Psychoanalysis.* Transactional Analysis Journal, 2007. **37**(1): p. 41-50.
26. Delimon, J., *[Short inpatient psychotherapy]*, in *Handboek Groepspsychotherapie*, T.J.C. Berk, et al., Editors. 1999, Bohn Stafleu van Loghum: Houten. p. F5.1-f5.31.
27. Thunnissen, M., *Long-term prognosis and aftercare in short-term inpatient psychotherapy of personality disorders: A randomised clinical trial of two methods of aftercare*. 2007, Erasmus MedicalCenter: Rotterdam, the Netherlands.
28. Rubin, D.B., *Multiple imputation after 18+ years.* Journal of the American Statistical Association, 1996. **91**(434): p. 473-489.
29. Yu, L.M., A. Burton, and O. Rivero-Arias, *Evaluation of software for multiple imputation of semi-continuous data.* Statistical Methods in Medical Research, 2007. **16**(3): p. 243-258.
30. Horton, N.J. and K.P. Kleinman, *Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models.* The American Statistician, 2007. **61**(1): p. 79-90.
31. Rubin, D.B., *Multiple imputation for nonresponse in surveys*. 1987, New York: John Wiley & Sons.
32. Wayman, J.C., *Multiple Imputation For Missing Data: What is it and How Can I Use it?* , C. Paper presented at the 2003 Annual Meeting of the American Educational Research Association, IL., Editor. 2003.
33. Drummond, M.F., et al., *Methods for the Economic Evaluation of Health Care Programmes, third Edition*. 2005, Oxford, UK: Oxford University Press. 400.

34. Hakkaart-van Roijen, L., *Handleiding Trimbos/iMTA Questionnaire for Costs Associated with Psychiatric Illness (TiC-P)*. 2002, Institute for Medical Technology Assessment, Erasmus Universiteit Rotterdam: Rotterdam.
35. Hakkaart-van Roijen, L., S.S. Tan, and C.A.M. Bouwmans, *Manual for Cost Research: Methods and Unit-prices for Economic Evaluations in Health Care [in Dutch]*. 2010, The Health Care Insurance Board (CVZ).
36. (CVZ), T.H.C.I.B., *Pharmaceutical Compass [in Dutch]*. 2005, CVZ.
37. Roijen, L.v., et al., *Labor and health status in economic evaluation of health care: the Health and Labor Questionnaire.* International Journal of Technology Assessment in Health Care, 1996. **12**: p. 405-15.
38. Koopmanschap, M. and F. Rutten, *A practical guide for calculating indirect costs of disease.* Pharmacoeconomics, 1996. **10**: p. 460-6.
39. Netherlands, S. *Consumer prince index*. 2010; Available from: http://statline.cbs.nl.
40. Van Hout, B.A., et al., *Costs, effects and C/E-ratios alongside a clinical trial.* Health Economics, 1994. **3**(5): p. 309-319.
41. Briggs, A.H., D.E. Wonderling, and C.Z. Mooney, *Pulling cost-effectiveness analysis up by its bootstraps: A non-parametric approach to confidence interval estimation.* Health Economics, 1997. **6**(4): p. 327-340.
42. EuroQolGroup, *EQ-5D user guide*. 1995, Sanders Institute EUR: Rotterdam, NL.
43. Papaioannou, D., J. Brazier, and G. Parry, *How to measure quality of life for cost-effectiveness analyses of personality disorders: a systematic review.* J Pers Disord, 2013. **27**(3): p. 383-401.
44. Lamers, L.M., et al., *Kwaliteit van leven meten in economische evaluaties: het Nederlands EQ-5D-tarief.* Nederlands Tijdschrift voor Geneeskunde, 2005. **149**(28): p. 1574-8.
45. Brooks, R., R. Rabin, and F. de Charro, *The Measurement and Valuation of Health Status using EQ-5D: A European Perspective.* Kluwer Academic Publishers, 2003.
46. Excellence, N.I.f.H.a.C., *Guide to the methods of technology appraisal*. 2008, National Institute for Health and Clinical Excellence: London.
47. van Asselt, A.D., et al., *How to deal with cost differences at baseline.* Pharmacoeconomics, 2009. **27**(6): p. 519-28.
48. van der Sanden, K., W. Smit, and M. Dashorst, *The referencing document of The Dutch National Qualification Framework to the European Qualification Framework.* 2012.
49. *A randomized comparison of the efficacy and cost-effectiveness of Psychotherapy Short-term Inpatient versus Long-term Outpatient (PSILO trial)*. [cited 2014 06-01-2014]; Available from: http://www.narcis.nl/research/RecordID/OND1324951.
50. Bouwmans, C., et al., *Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P).* BMC Health Serv Res, 2013. **13**: p. 217.
51. Soeteman, D.I., et al., *Cost-effective psychotherapy for personality disorders in the Netherlands: the value of further research and active implementation.* Value in Health, 2011. **14**(2): p. 229-39.
52. van Manen, J.G., et al., *Tailoring psychotherapy in patients with personality disorders: matching the level of psychological strengths to the level of stabilizing versus destabilizing psychotherapy.* submitted.

53. Wampold, B.E., et al., *A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially, "all must have prizes.".* Psychological Bulletin, 1997. **122**(3): p. 203-215.
54. van Manen, J.G., et al., *Relationship between patient characteristics and treatment allocation for patients with personality disorders.* Journal of Personality Disorders, 2011. **25**(5): p. 656-67.

Tables and figures

Tables and figures.

Table 1. Characteristics of patients allocated to STIP-TA and other psychotherapies (costs in Euros).

	STIP-TA N=67		OP N=67	
Age, years [*mean* (SD)*]*	39.4 (9.8)		39.3 (10.2)	
Male gender	22 (33%)		18 (27%)	
Education				
High (EQF[1]≥6)	26 (39%)		24 (36%)	
Medium (EQF[1] 3-5)	27 (40%)		31 (46%)	
Low (EQF[1]≤2)	14 (21%)		12 (18%)	
Mode of employment				
Paid work/study	39 (58%)		38 (57%)	
Unemployed/other	28 (42%)		29 (43%)	
Presence PD[2]				
Cluster A	1 (2%)		2 (3%)	
Cluster B	8 (12%)		7 (11%)	
Cluster C	33 (49%)		26 (39%)	
PD NOS	28 (42%)		33 (49%)	
	Mean (SD)	Median	Mean (SD)	Median
Healthcare Costs				
Somatic care	5,620 (24,647)	758	1,962 (8,403)	379
Psychiatric care	3,606 (8,499)	2,303	8,272 (18,897)	2,336
Other care	868 (1,670)	377	1,280 (3,787)	377
Medication	303 (332)	209	331 (326)	195
Total direct costs	10,398 (30,260)	4,471	11,845 (20,427)	5,430
Lost Production Costs				
Absenteeism	9,750 (16,435)	2,869	7,164 (13,163)	0
Reduced efficiency at work	3.868 (10,116)	0	1,595 (5,096)	0
Domestic help	167 (686)	0	680 (4,720)	0
Total indirect costs[3]	13,785 (19,229)	6,698	9,438 (14,384)	104
Total costs[3]	24,183 (37,283)	12,724	21,283 (23,283)	14,055
EQ-5D	.54 (.27)	.69	.53 (.26)	.65

[1]European Qualifications Framework.
[2]Since it is possible to have more than one diagnosis, the sum of the prevalence is higher than 100%.
[3]One-sample sign test on medians significant between groups.

Table 2. Costs and QALYs over three years follow-up.

	STIP-TA, €/$		OP, €/$		Incremental costs, €/$	2.5-97.5 percentile
Healthcare costs						
Somatic care, mean(SE)	10,377	(6,516)	5,909	(3,715)	4,628	
Psychiatric care, mean(SE)	5,899	(2,029)	11,399	(2,913)	-5,423	
Intervention, mean(SD)	19,420	(3,048)	25,459	(20,395)		
Other care, mean(SE)	1,150	(250)	1,440	(329)	-303	
Medication, mean(SE)	703	(105)	701	(111)	-9	
Subtotal, mean, mean(SE)	37,550	(7,923)	44,908	(5,954)		
Bootstrapped subtotal, mean(SD)	37,760	(7,885)	44,885	(5,839)	-7,125	-24,115 to 13,698
Lost production costs						
Absenteeism, mean(SE)	15,347	(1,847)	18,117	(3,223)	-2,686	
Reduced efficiency at work, mean(SE)	5,548	(1,238)	4,741	(1,203)	865	
Domestic help, mean(SE)	1,103	(586)	1,664	(607)	-557	
Subtotal, mean, mean(SE)	21,998	(2,620)	24,523	(3,818)		
Bootstrapped subtotal, mean(SD)	22,074	(2,597)	24,452	(3,753)	-2,378	-11,870 to 6,277
Total societal costs, mean(SE)	59,548	(9,116)	69,430	(7,842)		
Bootstrapped total costs, mean(SD)	59,834	(9,092)	69,337	(7,724)	-9,503	-32,561 to 15,726
QALY, mean(SE)	2.29 (.06)		2.05 (.07)			
Bootstrapped QALY, mean(SD)	2.29 (.06)		2.05 (.07)		.24	.05 to .44

Table 3. Sensitivity analyses on costs and QALYs.

			Costs	QALYs	ICER (costs per qaly)
3.5% discount rate		STIP-TA	60,182	2.20	
		OP	69,725	1.97	-43,130
Payer perspective		STIP-TA	37,550	2.29	
		OP	44,908	2.05	-29,688
Correction baseline differences	Standard regression	STIP-TA	58,529	2.29	
		OP	70,324	2.05	-49,146
	Split regression	STIP-TA	58,499	2.29	
		OP	69,689	2.05	-46,625
	Trimmed regression	STIP-TA	59,441	2.29	
		OP	65,133	2.05	-23,717
	Replacement regression	STIP-TA	61,792	2.29	
		OP	72,298	2.05	-43,775
Complete data only		STIP-TA	52,216	2.36	
		OP	60,420	2.05	-25,326

Figure 1. Patient flowchart.

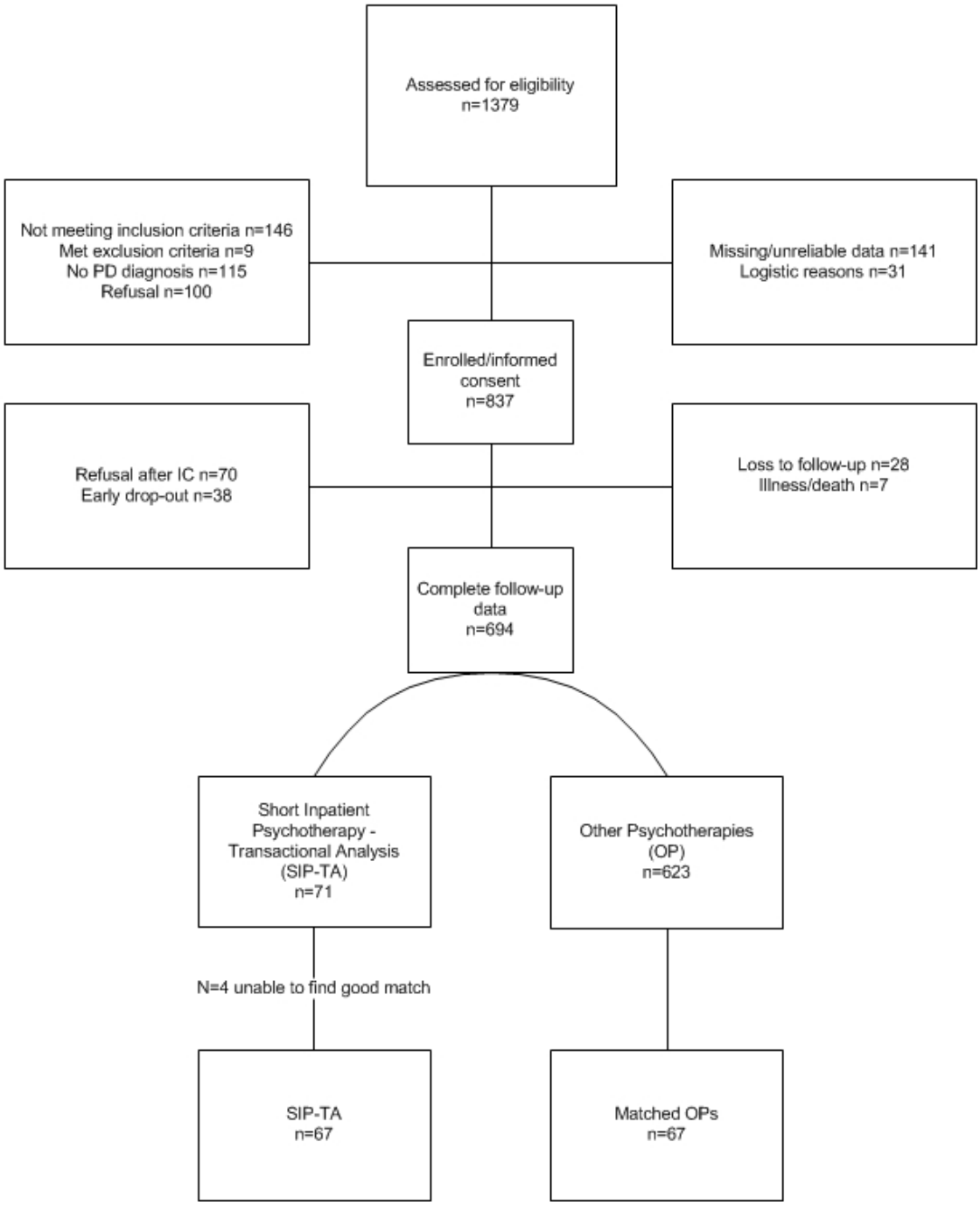

Figure 2. Distribution of costs.

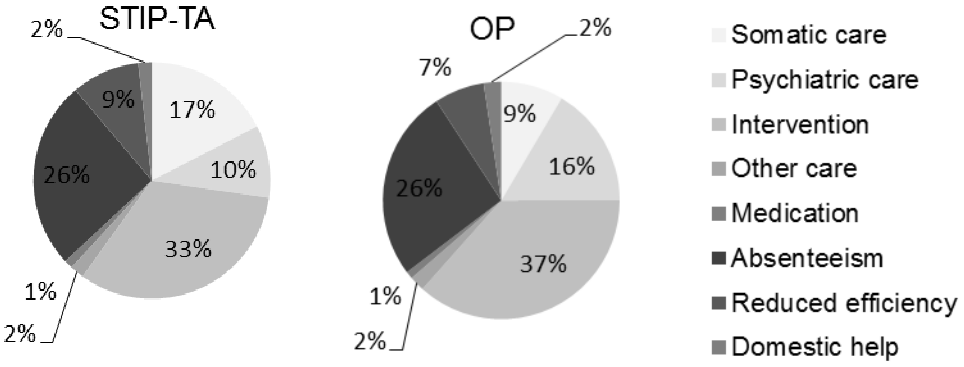

Figure 3. QALY course over 36 months.

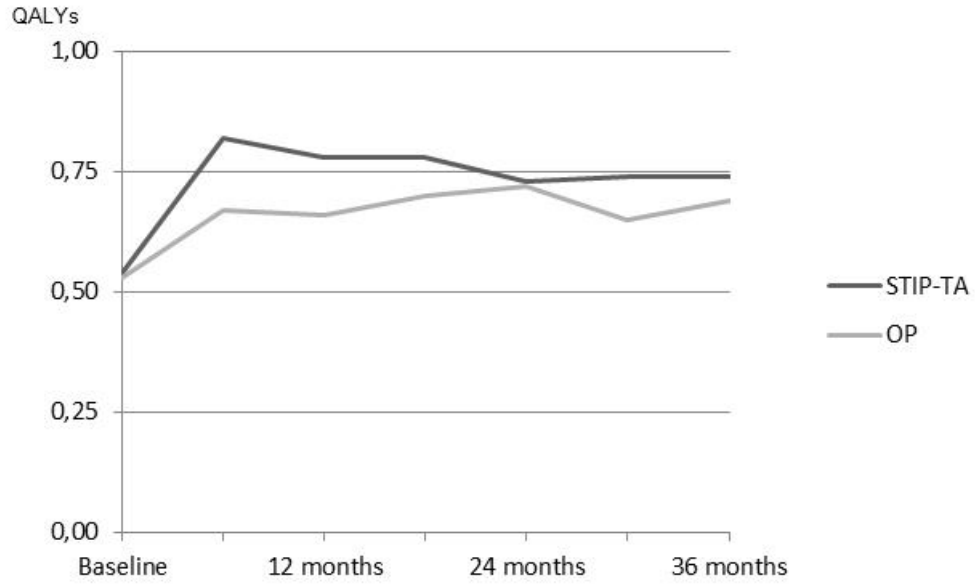

Figure 4. Cost-effectiveness plane.

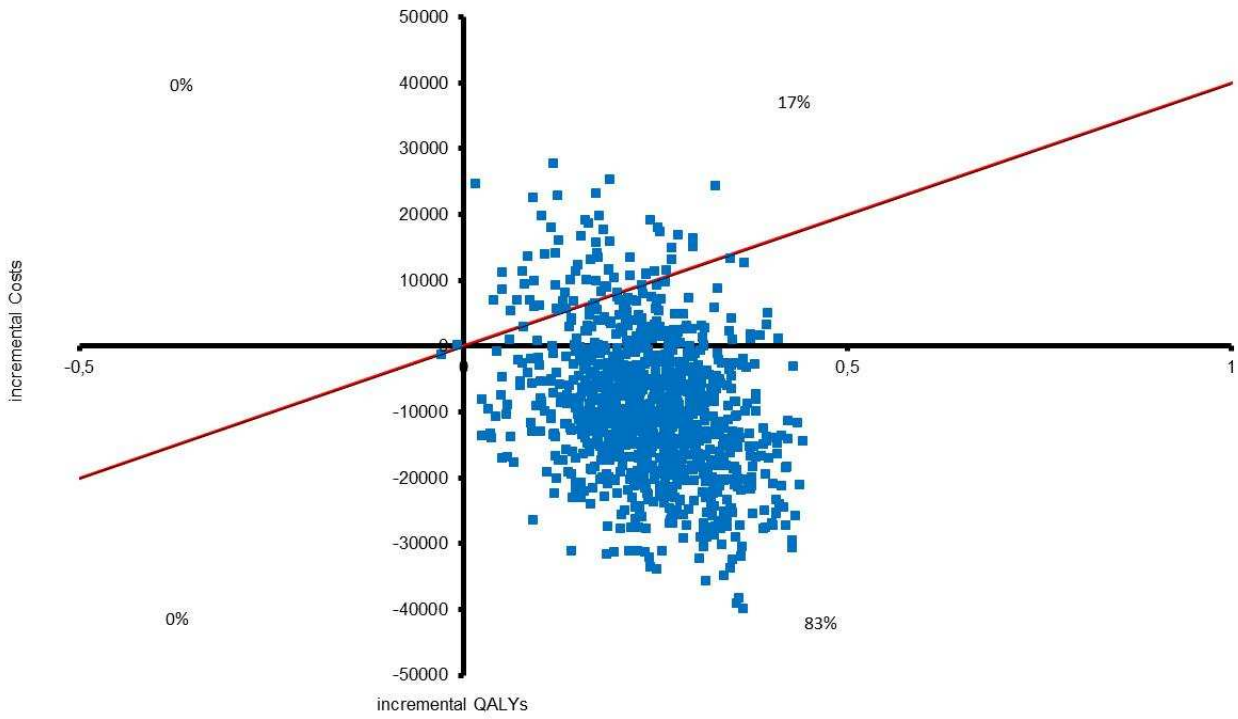

Figure 5. Acceptability curve