Parental health spillover in cost-effectiveness analysis: Evidence from self-harming adolescents in England

Sandy Tubeuf
Associate Professor, Academic Unit of Health Economics, University of Leeds,
s.tubeuf@leeds.ac.uk
Eirini-Christina Saloniki
Research Associate, Centre for Health Services Studies and PSSRU, University of Kent, e.saloniki@kent.ac.uk
David Cottrell
Professor, Academic Unit of Psychiatry and Behavioural Sciences, University of Leeds, d.j.cottrell@leeds.ac.uk

August 2018

Abstract

Objective - This paper presents alternative parental health spillover quantification methods in the context of a randomised controlled trial comparing family therapy with treatment as usual as an intervention for self-harming adolescents, and discusses the practical limitations of those methods.

Methods - The trial followed a sample of 754 participants aged 11 to 17 years. Health utilities are measured using answers to EQ-5D-3L for the adolescent and HUI2 for one parent at baseline, 6 and 12 months. We use regression analyses to evaluate the association between parent's and adolescent's health utilities as part of an explanatory regression model including health-related and demographic characteristics of both the adolescent and the parent. We then measure cost-effectiveness over a 12-month period as mean incremental cost-effectiveness ratios using various spillover quantification methods. We propose an original quantification based on the use of a household welfare function along with an equivalence scale to generate a health gain within the family to be added to the adolescent's QALY gain.

Results - We find that the parent's health utility increased over the duration of the trial and is significantly and positively associated with adolescent's health utility at 6 and 12 months but not at baseline. When considering adolescents' health gain only, the ICER is £40,453 per QALY. When including the health spillover to one parent, the ICERs estimates range from £27,167 per QALY to £40,838 per QALY and can be a dominated option depending on the quantification method used.

Conclusion - According to the health spillover quantification method considered, the incremental cost-effectiveness ratios vary from within the NICE cost-effectiveness threshold range to not being cost-effective.

Keywords: economic evaluation; self-harm; adolescent; EQ-5D-3L; HUI2; health spillover.

Key Points for Decision Makers:

1. Parental health spillover varies with treatment arms, follow-up points and independently from the adolescent's health improvement.
2. Health spillover can be measured using a household welfare function along with an equivalence scale to generate a health gain within the family to be added to the patient's QALY gain.
3. According to the health spillover quantification method considered, the incremental cost-effectiveness ratio varies from being cost-effective to not being cost-effective.

1. Introduction

Self-harm is commonly defined in the UK and Europe as any form of non-fatal self-poisoning or self-injury (such as cutting, taking an overdose, hanging, self-strangulation, jumping from a height, and running into traffic), regardless of the motivation or degree of intention to die. This definition would include US definitions of non-suicidal self-injury and suicidal behaviour. Self-harm in adolescents is a major public health issue with one in ten adolescents self-harming each year [1]. Individuals with mental disorders are heavy users of public health services and require emotional support and care from their family [2, 3]. Their disorders are likely to affect other family members' health and own health care needs, especially because individuals with mental health conditions face elevated rates of all-cause mortality and this places a huge burden of costs and life years lost on the family and the community [4].

It appears that the magnitude of spillovers on the health of other family members is the greatest in parents of ill children [5, 6]. Beyond the effect of caring for an ill child on parents' health [7], treatments that are provided to a self-harming child may have various spillover effects for the family. Indeed, psychotherapeutic treatments such as family-based therapies are often used with self-harming adolescents; they rely on individuals' relational network, involve parents, caregivers, brothers and sisters or other close relatives and friends in the therapies to improve clinical outcomes [8], and typically aim at maximizing cohesion, attachment and support while moderating parental control [9]. Therapy sessions do not necessarily include all family members, but it is expected that they will have an impact beyond the identified patient.

Some prior economic evaluations of psychotherapeutic interventions in young people have looked at the impact of the therapy on the adolescent/child patient and on relatives participating in the therapy. These studies collected parents or carers' outcomes and used them as additional outcomes of interest in a cost-effectiveness analysis [10-12] whilst only two studies combined child and parents' outcomes. Bodden et al. [13] used a compound summary of anxiety specific scores of the child, mother and father, as part of the sensitivity analyses. Their analysis measured the cost-effectiveness per anxiety-free family by including the costs related to the child and other family members' anxiety as self-reported in cost diaries. Cottrell et al. [14] used the same data as this paper over an 18-month follow-up and aggregated QALY of the adolescent and one parent as a sum in a sensitivity analysis. Their application relied on the strong assumption that QALYs can be summed across individuals. This assumption has been used in other studies in child health [15] and is consistent with research showing benefits to other family members involved in mental health family treatment [16, 17]. However, such considerations require a more thorough discussion of the interdependence between the utility functions of the adolescent and the parent, and the most appropriate method to include the overall health benefits.

The NICE reference case underlines that the perspective on outcomes considers "all direct health effects, whether for patients or, when relevant, carers" [18] however, there is no consensus on how these health effects should be measured and valued. Wittenberg and Posser [19] offered a summary of the evidence on the measurement and incorporation of health spillover of illness on family members or caregivers across health conditions, as a disutility. In their review, methods to measure spillovers were of three different types: (i) a direct measure of disutility of family members; (ii) a relative measure of family members' utility with a comparison to a control group; or (iii) an estimation of the utility of family members in a hypothetical scenario in which the patient is healthy or does not require caregiving.

In empirical economic evaluation studies, health spillovers have been included either as accrued health benefits [20-22] or as an estimated multiplier parameter which adjusts the patient's health gain with a spillover for the rest of a wider network (including parents, carers, spouses and other relevant individuals) [23, 24]. Whilst the first method uses a health-related quality of life (HRQoL) questionnaire and directly elicited utilities, the multiplier effect is based on a regression model using observational or primary data collection and consists of two multiplier effects.

In this paper, we use data from a multi-centre, individually randomised controlled trial comparing family therapy (FT) with treatment as usual (TAU) as an intervention for self-harming adolescents aged 11 to 17 [25] as a case study. Both the adolescent and one parent[1] reported their HRQoL as part of the trial across repeated follow-up points. We undertake a within-trial cost-effectiveness analysis (CEA) incorporating parental health spillover effects using alternative quantification methods. We add to the growing literature in three ways. First, we investigate the association between the health utility of the parent and a self-harming adolescent as part of an explanatory regression model using the preference-based HRQoL scores of both the adolescent and one parent. Second, we present a comparative analysis of alternative spillover quantification methods as part of an economic evaluation, bringing together the dyadic and the regression-based perspectives. Finally, we discuss how health spillovers could be adjusted including benefits to the rest of the family using an equivalence scale to adjust parental health gain.

2. The SHIFT trial case study

The SHIFT study was a randomised controlled trial conducted in local child and adolescent mental health services in Yorkshire, Greater Manchester and London for adolescents aged 11 to 17 years who had self-harmed twice. Participants were randomly allocated to receive FT or TAU. The objective of the trial was to assess whether or not FT would reduce the number of times the adolescents attended hospital with further self-harm. The trial results are reported elsewhere [14].

Personal characteristics were collected at baseline including adolescent's gender, age, type and number of self-harm episodes, as well as gender and age of their parent. Additional information were collected on adolescent's mental health using the Hopelessness scale for Children [26], parent's emotion toward

[1] The study collected data on the main caregiver, who was either the mother (86%) or the father (11%), so we will loosely use the term parent in this paper.

the adolescent using the Family Questionnaire (FQ) [27], parent's viewpoint on the family atmosphere through the McMaster Family Assessment Device [28] and parents' General Health Questionnaire (GHQ-12) [29]. All these measurements are defined in Table 1. Adolescent's HRQoL was measured by the EQ-5D-3L [30] whilst the parent's by HUI2 [31, 32]. The original research proposal considered HUI2 as HRQoL measure for both the parent and the adolescent following the NICE guidelines at the time [33, 34]. However, we carried out a pilot study [35] on a sample of 49 adolescents aged 11-18 to test the ability of children to deal with the concepts and language used in the EQ-5D-3L and HUI2. We found that EQ-5D-3L had the least amount of missing data and presented limited problematic wording for that age group therefore EQ-5D-3L was eventually used to measure HRQoL of adolescents in the trial. However, the parents' HRQoL instrument failed to be changed.

Adolescent's responses to the EQ-5D-3L were converted into health state utility scores using national tariff values [36]. Similarly, parent's responses to HUI2 were converted into health state utility values [32, 37]. The area under the curve approach was used to calculate quality-adjusted life-years (QALYs) for the adolescent and the parent.

Resource use of health services was self-reported by the adolescent and/or their parent. Accident and Emergency visits and inpatient stays of the adolescent were available from NHS Digital records. Resource use was combined with national unit costs distinguishing, where possible, by self-harm and not self-harm-related event leading to hospitalisation [38]. Psychotropic medication costs were calculated using trial medication records. The intervention costs were calculated separately for each treatment arm using information on the type and duration of the therapies sessions available from the trial records [14, 39].

Eighty hundred and thirty two adolescents and their parent were recruited in the trial (417 in TAU and 415 in FT). This paper focuses on the first 12-month follow-up so discounting is not required. Missing utility scores and total health and hospital services costs at 6 and 12 months were imputed using multiple imputations via chained equations [40-42]. Imputations were based on a number of demographic and clinical predictors; the process is described elsewhere [39]. Missing utility (4%) and clinical scores (3%) at baseline were not imputed. The sample used in the main analysis is 731 adolescents and their parent (359 in TAU and 372 in FT). As part of the sensitivity analysis, the analysis was also carried out on the complete case sample; the sample reduced to 206 adolescents and their parent (73 in TAU and 133 in FT).

3. Methods

3.1 The association between parent's and self-harming adolescent's health

We firstly modelled the utility of the parent as a function of the adolescent's HRQoL (utility) in the same period controlling for a number of adolescent and parent's characteristics. There is no reason to believe that this association remains consistent over time therefore our approach extends prior research [7, 23] by investigating the relationship empirically at multiple follow-up points in the data as follows:

$$H_i^t = \alpha^t + \beta_1^t H_j^t + \beta_2^t Z_j^0 + \beta_3^t C_i^0 + \varepsilon_i^t \quad (1)$$

where H_i^t denotes the parent's i health-related quality of life measured by the HUI2 index score at time $t = 0{,}1{,}2$ for baseline, 6, and 12 months; H_j^t denotes the adolescent's j HRQoL at time t measured by the overall EQ-5D-3L index score; Z_j^0 is a vector of baseline characteristics of the adolescent such as age, gender, type of self-harm event, total number of self-harm events; C_i^0 is a vector of baseline characteristics of the parent such as gender and mental health measured by the GHQ-12; α^t is the intercept, $\beta_1^t, \dots, \beta_3^t$ are the slope parameters and ε_i^t is the error term with $\varepsilon_i^t \sim N(0,1)$. The distribution of the utility of the parent is skewed to the left so we estimate all regression models using Tobit models. The estimated coefficients were similar to those from the Ordinary Least Squares regressions both in magnitude and sign[2].We initially run Model 1 including demographic controls for both the adolescent (age, gender) and the parent (age). To account for the heterogeneity observed in adolescents' parents, we subsequently run Model 2 controlling for other adolescent characteristics (Hopelessness scale score, type of self-harm) and family characteristics from the parent's perspective (Family questionnaire, McMaster Family Assessment Device) as well as the parent's GHQ-12. We supplemented this simplistic association analysis with a more causal understanding of the impact of a positive change in the adolescent's health over time on parent's HRQoL in line with Bhadhuri et al. [43]. We included a binary variable taking the value 1 if the adolescent's EQ-5D-3L score improved between baseline and follow-up, but this parameter was not significant and did not impact on the results[3].

3.2 Parental health spillover in CEA: five alternative quantifications

The base-case cost-effectiveness analysis (CEA) considers the incremental costs and QALYs associated with FT versus TAU as an intervention for self-harming adolescents. We are interested in quantifying the health spillover effects to the parent in the CEA, which can be used as an extra QALY gain inflating the adolescent's QALY gain. Using the regression model presented in Eq. (1) as a starting point, we suggest four alternative quantification methods to evaluate parental health spillover. We also consider a fifth quantification with a direct measurement of parental QALY gain using answers to the HUI2 index.

3.2.1 Relative health spillover (Quantification 1)

The estimated parameter $\widehat{\beta_1^t}$ in Eq. (1) can be used to extract a spillover coefficient of adolescent's health utility on parents. Assuming policy makers are interested in accounting for broad health benefits

[2] Results available upon request
[3] Results available upon request.

independently of the treatment arm, the parameters $\widehat{\beta_1^0}, \widehat{\beta_1^1}$ and $\widehat{\beta_1^2}$ represent a utility gain for the parent at each time point, which can be transformed into a QALY gain using the area under the curve approach as follows:

$$QALY_i^{Spe\ 1} = \left[\frac{(\widehat{\beta_1^0}+\widehat{\beta_1^1})}{2}\right] \times 0.5 + \left[\frac{(\widehat{\beta_1^1}+\widehat{\beta_1^2})}{2}\right] \times 0.5 \qquad (2)$$

If the relationship between the adolescent and parent's HRQoL remains constant over time, the parameter $\widehat{\beta_1^t}$ represents the full QALY gain, which is similar to what Al-Janabi et al. (2016) called *relative* spillover.

3.2.2 Relative health spillover per treatment arm (Quantification 2)

One might suggest that we should also account for the heterogeneity in the parental health spillover according to the treatment received, especially because parents are directly involved in the FT arm, but not systematically involved in TAU[4]. In this case, the parameter $\widehat{\beta_1^t}$ will also vary by treatment arm. Let us consider the estimated parameter $\widehat{\beta_1^{t,FT}}$ where $FT = 0$ when Eq. (1) is run on the sample of adolescents receiving TAU and $FT = 1$ when it is run on those receiving FT. Three estimated health spillover coefficients (one for each time point) *within* each treatment arm can be used to quantify a utility gain for the parent, and then transformed into a QALY gain as follows:

$$QALY_i^{Spe\ 2} = \left[\frac{(\widehat{\beta_1^{0,FT}}+\widehat{\beta_1^{1,FT}})}{2}\right] \times 0.5 + \left[\frac{(\widehat{\beta_1^{1,FT}}+\widehat{\beta_1^{2,FT}})}{2}\right] \times 0.5 \quad \text{with } FT = \{0,1\} \qquad (3)$$

3.2.3 Absolute health spillover (Quantification 3)

Considering the primary outcome of the study was reducing repetitions of self-harm over 12 months one could argue that measuring spillover coefficients according to the final primary outcome provides an *absolute* health spillover for the parent. Contrary to Quantification 1, Eq. (1) is now run separately on the sub-sample of adolescents who did not have a repeated self-harm at 12 month and on those who did self-harm again. The two sets of estimated health spillover coefficients $\widehat{\beta_1^{t,SH}}$ with $SH = \{0,1\}$ are used to generate an absolute QALY gain for the parent as follows.

$$QALY_i^{Spe\ 3} = \left[\frac{(\widehat{\beta_1^{0,SH}}+\widehat{\beta_1^{1,SH}})}{2}\right] \times 0.5 + \left[\frac{(\widehat{\beta_1^{1,SH}}+\widehat{\beta_1^{2,SH}})}{2}\right] \times 0.5 \quad \text{with } SH = \{0,1\} \qquad (4)$$

[4] TAU included supportive therapy/counselling (25.1%), cognitive–behavioural therapy (17.4%), family work (11.5%), formal systemic FT (10.7%), and various other therapies.

3.2.4 Absolute global health spillover per treatment arm (Quantification 4)

The absolute QALY gain for the parent could additionally account for the heterogeneity in health spillover according to treatment. The health spillover is measured using the estimated coefficient $\widehat{\beta_1^{t,SH,FT}}$ estimating Eq. (1) on four different sub-samples of adolescents.

$$QALY_i^{Spec\ 4} = \left[\frac{(\widehat{\beta_1^{0,SH,FT}} + \widehat{\beta_1^{1,SH,FT}})}{2}\right] \times 0.5 + \left[\frac{(\widehat{\beta_1^{1,SH,FT}} + \widehat{\beta_1^{2,SH,FT}})}{2}\right] \times 0.5 \text{ with } SH = \{0,1\}\ FT = \{0,1\} \quad (5)$$

3.2.5 Additive accrued health benefits (Quantification 5)

Using prior empirical studies [20-22], health spillover could also be measured using an additive approach where the QALY gain of each individual in the dyad adolescent/parent is independently calculated and then the two QALY gains are summed up. Our case study uses two different HRQoL instruments for the adolescent and the parent. Let us consider HUI^t represents the parent's health state utility value at each time point, parent's QALYs are calculated as follows:

$$QALY_i^{Spec\ 5} = \left[\frac{HUI^0 + HUI^1}{2}\right] \times 0.5 + \left[\frac{HUI^1 + HUI^2}{2}\right] \times 0.5 \quad (6)$$

It is worthwhile to note that we assume that the QALY as generated from HUI2 or EQ-5D-3L are of the same nature and meaning, and can be summed even if they are generated for two different individuals and produced from different instruments. This assumption follows from the foundation of resource allocation decisions in health according to which QALY provides an equal valuation between individuals and health care interventions of health improvement, independently of the HRQoL instrument being used to measure quality of life.

3.3 Parental health spillovers in CEA: a new perspective

In addition to all possible quantification methods to account for parental health spillover, outlined in the previous section, we propose an additional method.. While in the context of the economic evaluation of meningitis vaccination, Al-Janabi et al. [23] proposed a unique health spillover estimate that was applied to each family member affected or a health spillover estimate according to their proximity to the patient, we believe that this would not be appropriate in our case study. Three arguments motivate our viewpoint.

First, a single utility value would deny the heterogeneity observed in parents' characteristics at baseline and their potential to benefit over the duration of the study according to their level of engagement in the treatment, whether this is FT or TAU. From a clinical viewpoint, it would be expected that FT has an impact on other members of the family irrespective of whether those members attended the therapy sessions or whether there was any change in the self-harming adolescent. If therapy leads to, those attending, behaving or communicating differently this will inevitably impact others they relate to. The magnitude and even the direction of such impacts will vary from one family member to another, but cannot be ignored. Second, the treatment arm itself might impact on the parent's health independently from the adolescent's health improvement; in the SHIFT trial for a number of secondary outcomes caregivers reported significantly better outcomes than the adolescents [14]. Third, as part of a trial several repeated observations of health utilities are available and it appears important to account for all the available repeated information when quantifying spillover.

These arguments would lead us to consider the additive approach, where the QALY gain of each individual in the dyad adolescent/parent is independently calculated, appealing. At the same time, it is important to ensure that such aggregation does not lead to a decision that deteriorates the health of the adolescent, or more generally, of the patient in the first place. There are clear value judgements about the priority assigned to the identified patient, who is judged the most important individual to benefit from a treatment while the inclusion of health spillover effects for other individuals are of secondary purpose.

For this reason, we propose that health gains are aggregated at the household level if and only if the QALY gain for the patient is positive or equal to zero. When the QALY gain for the patient is positive we need to identify a way of adjusting for the parental health spillover so that the patient's health gain remains a priority for the health care decisions to be made. The concept of equivalence scale (ES) as we will refer to from now on, has been used in economics to measure social welfare and adjust the income of all household members accounting for the size the household and the age of its members [44-46]. In our context, an ES would allow to adjust all health gains for the rest of the household as an additional individual equivalent QALY or utility gain where all the household members (including the patient) are accounted for. The ES transforms a distribution of observed QALY gains across heterogeneous household members into a household health gain. This adjusted health spillover can then simply be summed to the QALY gain of the patient in the cost-effectiveness analysis.

Following Buhmann et al. [44], let us consider that Q measures the adjusted health spillover as follows:

$$Q = \frac{\sum_{r=1}^{R} h_r}{(R+1)^a} \tag{7}$$

where h_r equals the health spillover for each family relative r, R is the number of family relatives with an observed QALY or utility gain, and a is the elasticity of the ES rate, which varies between 0 (when the health spillover is unadjusted and equivalent to a simple sum of the QALY gain available) and 1 (when a per capita QALY is used). The value of a is defined according to the importance given to the QALY gain of the family members beyond the patient (e.g. if $a = 0$ the family members are as important as the patient and this would be equivalent to quantification 5). In our alternative specifications, we consider five examples of quantification of health spillover using an ES where $a = \{0, 0.3, 0.5, 0.8, 1\}$.

4. Results

4.1 Regression models

Descriptive statistics are presented in Table 2. At baseline, more than two thirds of the adolescents were females with about three self-harm episodes over the duration of the trial. Self-harm was caused by self-injury for over 70% of the adolescents with more than 50% reporting some problems with anxiety/depression. For parents, 86% were mothers with average age of 42 years (see Table 3). Parent's average GHQ-12 was 8.52 (SD=5.38) which is within the distressed range (4-12) but lower than the level of psychological distress observed in a sample of caregivers of a dependent relative [47].

Table 4 shows the mean utility scores for adolescents and their parent at baseline, 6 and 12 months, overall and by treatment arm. For the adolescents, utility scores increase monotonically over the 12 months and regardless of the treatment arm. Differences in utility scores between arms were significant at 6 and 12 months favouring FT. The difference from baseline appears to be slightly larger in FT than in TAU (on average 0.145 versus 0.095). Parent's utility also shows an increase in the overall HUI2 score at 6 and 12 months from baseline; this increase however is much smaller than for the adolescent (on average 0.045 versus 0.12) and is not significant when distinguished by treatment arm.

Table 5 presents the Tobit regression results of the parent's health-related quality of life; the association between parent and adolescent's health varies across time points and model specifications. We find a significant and positive association with parent's health at 6 months and 12 months in Model 2 while in Model 1, parent's health is positively associated with adolescent's HRQoL at 6 months only. This is in line with prior studies on the experience of parents' caregiving for an ill child [5, 7, 48], and carers of people with mental health disorders [3].

Parent's HRQoL at every time point also appears to be negatively associated with a higher score of emotion within the family, of poor family functioning and of psychological distress as measured by GHQ-12, all three measured at baseline. The strong association between parent's utility and GHQ-12 has also been shown in other studies [49]. On the other hand, parent's health is positively and significantly associated with adolescent's higher score of hopelessness, however this association substantially reduces in magnitude and significance over time.

4.2 Spillover effects in cost-effectiveness analysis

Table 6 presents the ICERs and their respective probabilities of cost-effectiveness using the base-case analysis when only the adolescent's QALY gain is considered along with the five regression-based alternative spillover quantifications[5] and the equivalence scale based spillover quantification with five

[5] Quantification 1 is based on the Tobit regression results presented in Table 5. Quantification 2 is based on Tobit regression results presented in Table A1. Quantification 3 is based on Tobit regression results presented in Table A2. Quantification 4 is based on Tobit regression results presented in Table A3.

alternative elasticity values. Costs used in the analysis are summarised in Table A4. Since we did not collect health care costs for the parent we note that the costs for each ICER are strictly identical and it is only the level of QALY gain that varies.

Results from the base-case analysis indicate that adolescents in FT incurred £1,207 higher costs on average and gained 0.030 extra QALYs than the adolescents in TAU, which is equivalent to an extra 10.95 days of perfect health annually. The ICER from this analysis (£40,453 per QALY) is above the recommended threshold range specified for NICE decision-making in England and Wales (£20,000-£30,000 per QALY gain), indicating that FT is unlikely to be cost-effective. When considering the relative parental health spillover independently of the treatment arm using quantification 1, the ICER is almost identical to the one obtained from the base-case analysis. However, when accounting for the direct involvement of the parents in the FT arm (quantification 2), parents and adolescents continue to incur higher costs on average but with 24.5 fewer days of perfect health (loss of 0.067 QALYs annually) than those in TAU and therefore indicating that FT is dominated by TAU. The ICER remains above the nationally recommended threshold when we control for the absolute parental health spillover using the number of repeated self-harm events at 12 months (£40,838), implying that FT is not cost-effective. If we further control for any heterogeneity in the absolute parental health spillover, FT is dominated by TAU with adolescents and parents in the FT arm incurring 54.8 fewer days of perfect health (loss of 0.150 QALYs annually) than those in the TAU arm. Any of the regression-based quantification indicates that FT is unlikely to be cost-effective. However, the ICER reduces to £27,167 per QALY when we simply sum adolescent's and parent's QALYs (quantification 5), demonstrating a potential for FT to bring 16.1 extra days at full health annually for both the adolescent and the parent and a value within the NICE threshold range. As expected quantification 5 is equivalent to the quantification with an equivalence scale using an elasticity of $a = 0$. The value of the elasticity a directly impacts on the average QALY gains, and the higher the elasticity, the lower is the cumulated QALY gain and so the higher is the ICER. For smaller values of the elasticity a (less than 5), the quantifications using an ES show an ICER within the NICE cost-effectiveness range. The probability of FT to be cost-effective is higher when using ES to quantify spillover than with regression-based spillover quantifications; at £20K it is comprised between 16% and 28% with ES versus 0%-7% with regressions. At £30K, it respectively reaches 43%-54% versus 0%-28%. It is important to note that with any quantification method, both cost differences between FT and TAU and QALY differences are significant. The same analyses were carried out on the complete case sample to test the sensitivity of the results to missing data imputations (see Table A5). The ICER estimations for each spillover quantification are all larger (between £34,071 and £45,842) with broader standard deviations for both costs and QALYs. It is remarkable that the differences between quantifications present the same pattern as the main analysis.

6. Discussion

We showed that parent's HRQoL is associated with the health of a self-harming adolescent. We investigated how health spillover for the parent could be included in CEA using alternative quantifications based on estimated coefficients and QALY valuations. Sensitivity analyses revealed that the valuation technique had a considerable impact on the magnitude of QALY and could change the inference about the most cost-effective alternative in a trial study. We made two propositions in this paper. Proposition 1 suggests that health gains are only aggregated at the household level when the QALY gain for the patient is positive or equal to zero. Proposition 2 suggests the use of an equivalence scale (ES) to convert a distribution of observed health spillover across other household members into an extra health gain to be added to the patient's QALY gain. We illustrated the use of an ES with a set of alternative elasticity values. There are several advantages with the use of ES. First, ES have been widely used in the literature to measure household social welfare [44-46]. Second, health spillover measured either as a QALY gain from a utility score or a utility parameter generated from a regression model could be summed and transformed in an extra health gain using the ES. Third, the ES adapts to data availability and so every family relative with observed health outcomes can be included. Finally, one could transform easily the ES to account for family members' proximity to the patient including an individual weight in the same way it is done with income equivalence scales[6]. This methodological proposition will require further scrutiny in future research.

Our study presents limitations.

The trial study used two different HRQoL instruments to measure the adolescent and the parent's quality of life. For the purpose of the spillover quantification, we assumed that utilities and QALYs generated from two different generic measures were of the same nature and meaning and could be combined. However these two measures are quite different in descriptive content and in valuation technique. While EQ-5D covers dimensions of physical, mental and general health and is valued with Time Trade-Off, HUI2 additionally considers impairments in vision, hearing, and dexterity and is valued using standard gamble and visual analogue scaling. Research has shown a moderate level of agreement between HRQoL measures in various condition-specific groups [50-53]. The assumption according to which the two preference-based measures can be combined in our spillover quantifications could potentially be biased. For example, if EQ-5D-3L tends to provide lower mean utility estimates than HUI2, this would imply for our study that quantification 5 and the ES quantification with $a = 0$ leads to an aggregation of health gains where the parent's QALY gain from the intervention is relatively higher than for the adolescent's (patient) and so the patient is not the main beneficiary (though respecting proposition 1 ensures that the patient is the priority for the health care decision making). In this context, head-to-head

[6] See Table page 2 in the OECD note on equivalence scales available here URL: http://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf (consulted June 2018)

comparisons between preference-based HRQoL instruments will be useful to develop potential measurement corrections to ensure comparability between utilities and QALYs when measuring health spillover.

Methodologically, the reverse correlation with a focus on the impact of parent's health on adolescent's health could have been of interest to study. Moreover, several authors [13, 17, 54] have argued that potential health care cost savings are transferred to others when treating one family member using family-based psychotherapy; it would be ideal to include the health care resource use of the parent had they been available in the data.

Conceptually, we investigated how social externalities such as the health effects on other individuals could be introduced into the framework of CEA; to some extent, this questions whether cost-utility analysis is appropriate or whether cost-benefit analysis with distributional weights should be considered. We did not enter into this debate and assumed that cost-utility analysis would remain the preferred method for the health spillover quantification [55].

Admittedly, our proposition to rely on an ES is a pragmatic choice. The adoption of a unique scale that would be identical for any CEA would have the advantage to facilitate the generation of evidence that is comparable between individuals and between cost-utility analyses.

Data Availability Statement

Regarding research data, the principal investigator of the SHIFT study has agreed with the NIHR that data would not be made freely available in a repository and interested third parties can apply to the trial team for access Clinical Trials Research Unit – University of Leeds – Leeds – LS2 9JT.

Acknowledgements

We are grateful to Bryony Dawkins, John O'Dwyer, Laetitia Schmitt, Paula Boston, and participants to the LAGV International Conference in Public Economics in Aix en Provence in June and the EuHEA conference in Maastricht in 2018 for comments and feedback on an earlier version of this paper. Two anonymous reviewers and the editors provided us with rich comments that helped improving the paper. We thank Josephine Aikpitanyi for her help with literature references.

Compliance and Ethical Standards

This research work was undertaken as part of the National Institute for Health Research Health Technology Assessment programme (project number: 07/33). The views expressed in this publication are those of the authors and do not necessarily reflect those of the HTA programme, NIHR, NHS, or the Department of Health. Study registration: ISRCTN 59793150. ST designed the trial health economics analysis and ECS lead the analysis. ST and ECS developed the framework for this research paper and equally contributed in writing. DC was PI of the SHIFT trial and provided a clinical insight for the research. ST is the overall guarantor of the research in this paper.

Declared competing interests of authors:

Sandy Tubeuf received funding from NIHR HTA (Grant 07/33/01) including support for travel conference presentation. Eirini-Christina Saloniki received funding from NIHR HTA (Grant 07/33/01) including support for travel and conference presentation. David Cottrell received funding from NIHR HTA (Grant 07/33/01) including travel costs for data collection and to meetings to present results.

References

1. Evans E, Hawton K, Rodham K. In what ways are adolescents who engage in self-harm or experience thoughts of self-harm different in terms of help-seeking, communication and coping strategies? Journal of Adolescence. 2005 August;28(4):573-87.

2. Grad J, Sainsbury P. Mental illness and the family. The Lancet. 1963:544-7.

3. Shah AJ, Wadoo O, Latoo J. Quality of life for parents of children with autism spectrum disorders. British Journal of Medical Practitioners. 2010;3(3).

4. McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the price: the cost of mental health in England to 2026. London: The King's Fund; 2008.

5. Wittenberg E, Saada A, Posser LA. How Illness Affects Family Members: A Qualitative Interview Survey. Patient. 2013;6:257-68.

6. Lavelle TA, Wittenberg E, Lamarand K, Posser LA. Variation in the Spillover Effects of Illness on Parents, Spouses, and Children of the Chronically Ill. Applied Health Economics and Health Policy. 2014;12:117-24.

7. Kuhlthau K, Payakachat N, Delahaye J, Hurson J, Pyne JM, Kovacs E, et al. Quality of life for parents of children with autism spectrum disorders. Research in Autism Spectrum Disorders. 2014;8:1339-50.

8. Stratton P. Report On The Evidence Base Of Systemic Family Therapy. Association for Family Therapy. 2005.

9. Fortune S, Cottrell D, Fife S. Family factors associated with adolescent self-harm: a narrative review. Journal of Family Therapy. 2016;38(2):226-56.

10. Byford S, Harrington R, Torgerson D, Kerfoot M, Dyer E, Harrington V, et al. Cost-effectiveness analysis of a home-based social work intervention for children and adolescents who have deliberately poisoned themselves. Results of a randomised controlled trial. Br J Psychiatry. 1999 Jan;174:56-62.

11. Goldfield G, Epstein L, Kilanowski C, Paluch R, Kogut-Bossler B. Cost-effectiveness of group and mixed family-based treatment for childhood obesity. International Journal of Obesity. 2001 Dec;25(12):1843-9.

12. Flores G, Bridon C, Torres S, Perez R, Walter T, Brotanek J, et al. Improving asthma outcomes in minority children: a randomized, controlled trial of parent mentors. Pediatrics. 2009;124(6):1522-32.

13. Bodden DH, Dirksen CD, Bogels SM, Nauta MH, De Haan E, Ringrose J, et al. Costs and cost-effectiveness of family CBT versus individual CBT in clinically anxious children. Clin. 2008 Oct;13(4):543-64.

14. Cottrell D, Wright-Hughes A, Collinson M, Boston P, Eisler I, Fortune S, et al. Effectiveness and cost-effectiveness of systemic family therapy compared with treatment as usual for young people after self-harm: a pragmatic randomised controlled trial. The Lancet Psychiatry. 2018;5(3):203-16.

15. Griebsch I, Coast J, Brown J. Quality-adjusted life-years lack quality in pediatruc care: a critical review of published cost-utility studies in child health. Pediatrics. 2005;115(5):e600-e14.

16. Fals-Stewart W, O'Farrell TJ, Birchler GR, Cordova J, Kelley ML. Behavioral couples therapy for alcoholism and drug abuse: Where we've been, where we are, and where we're going. Journal of Cognitive Psychotherapy. 2005 Fall;19(3):229-46.

17. Crane D. Does family therapy reduce health care costs for more than the identified patient? Clinical Child Psychology and Psychiatry. 2011;16(1).

18. NICE. Guide to the methods of technology appraisal 2013. Process and methods guides, NICE. 2013:102.

19. Wittenberg E, Posser LA. Disutility of Illness for Caregivers and Families: A Systematic Review of the Literature. Pharmacoeconomics. 2013;31:489-500.

20. Meeuwsen E, Melis R, van der Aa G, Goluke-Willemse G, de Leest B, van Raak F, et al. Cost-Effectiveness of One Year Dementia Follow-Up Care by Memory Clinics or General Practitioners: Economic Evaluation of a Randomised Controlled Trial. PLOS ONE. 2013;8(11):1-7.

21. Søgaard R, Sørensen J, Waldorff FB, Eckermann A, Buss DV, Phung KTT, et al. Early psychosocial intervention in Alzheimer's disease: cost utility evaluation alongside the Danish Alzheimer's Intervention Study (DAISY). BMJ Open. 2014;4:e004105.

22. Sturkenboom IHWM, Hendriks JCM, Graff MJL, Adang EMM, Munneke M, Nijhuis-van der Sanden MWG, et al. Economic Evaluation of Occupational Therapy in Parkinson's Disease: A Randomized Controlled Trial. Movement Disorders. 2015;30(8):1059-67.

23. Al-Janabi H, Exel VJ, Brouwer W, Coast J. A framework for including family health spillovers in economic evaluation. Medical Decision Making. 2016:176-86.

24. Al-Janabi H, Exel VJ, Brouwer W, Trotter C, Glennie L, Hannigan L, et al. Measuring health spillovers for economic evaluation: A case study in meningitis. Health Economics. 2016;25:1529-44.

25. Wright-Hughes A, Graham E, Farrin A, Collinson M, Boston P, Eisler I, et al. Self-Harm Intervention: Family Therapy (SHIFT), a study protocol for a randomised controlled trial of family therapy versus treatment as usual for young people seen after a second or subsequent episode of self-harm. Trials. 2015;16(1):1.

26. Kazdin AE, Rodgers A, Colbus D. The hopelessness scale for children: psychometric characteristics and concurrent validity. Journal of consulting and clinical psychology. 1986;54(2):241.

27. Wiedemann G, Rayki O, Feinstein E, Hahlweg K. The Family Questionnaire: development and validation of a new self-report scale for assessing expressed emotion. Psychiatry research. 2002;109(3):265-79.

28. Miller IW, Epstein NB, Bishop DS, Keitner GI. The McMaster family assessment device: reliability and validity*. Journal of Marital and Family Therapy. 1985;11(4):345-56.

29. Goldberg D. General health questionnaire (GHQ-12). Windsor, UK: Nfer-Nelson. 1992.

30. EuroQoL Group. EuroQoL - A new facility for the measurement of health related quality of life. Health Policy. 1990;16:199-208.

31. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI®): concepts, measurement properties and applications. Health and quality of life outcomes. 2003;1(1):1.

32. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-Attribute Health Status Classification Systems: Health Utilities Index. PharmacoEconomics. 1995;7(6):490-502.

33. Thorrington D, Eames K. Measuring health utilities in children and adolescents: a systematic review of the literature. PLoS ONE. 2015;10(8):e0135672

34. Adlard N, Kinghorn P, Frew E. Is the UK NICE "reference case" influencing the practice of pediatric quality-adjusted life-year measurement within economic evaluations? Value in Health. 2014;17:454-61.

35. Oluboyede Y, Tubeuf S, McCabe C. Measuring health outcomes of adolescents: report from a pilot study. The European Journal of Health Economics. 2013;14(1):11-9.

36. Dolan P. Modeling valuations for EuroQol health states. Medical care. 1997;35(11):1095-108.

37. McCabe C, Stevens K, Roberts J, Brazier J. Health state values for the HUI2 descriptive system: results from a UK Survey. Health Economics. 2005;14(3):231-44.

38. Curtis L. Unit Costs of Health and Social Care 2014, Canterbury, Kent. Personal Social Services Research Unit. 2014:URL: http://www.pssru.ac.uk/project-pages/unit-costs/2014.

39. Cottrell D, Wright-Hughes A, Collinson M, Boston P, Eisler I, Fortune S, et al. A pragmatic randomised controlled trial and economic evaluation of family therapy versus treatment as usual for young people seen after second or subsequent episodes of self-harm: the Self-Harm Intervention – Family Therapy (SHIFT) trial. Health Technology Assessment. 2018; 22(12): 1-222.

40. Little RJ, Rubin DB. Single imputation methods. Statistical Analysis with Missing Data, Second Edition. 2002:59-74.

41. Rubin DB. The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for

creating a few imputations when fractions of missing information are modest: The SIR algorithm. Journal of the American Statistical Association. 1987;82(398):543-6.

42. Schafer JL. Analysis of incomplete multivariate data: CRC press; 1997.

43. Bhadhuri A, Jowett S, Jolly K, Al-Janabi H. A comparison of the validity and responsiveness of the EQ-5D-5L and SF-6D for measuring health spillovers: A study of the family impact of meningitis. Medical Decision Making. 2017.

44. Buhmann B, Rainwater L, Guenther S, Timothy S. Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database. The Review of Income and Wealth. 1988;34(2):115-42.

45. Chanfreau J, Burchardt T. Equivalence Scales: Rationales, Uses and Assumptions. 2008.

46. Anyaegbu G. Using the OECD equivalence scale in taxes and benefits analysis. Economics and Labour Market Review. 2010;4:49-54.

47. Cuellar-Flores I, Sanchez-Lopez MP, Liminana-Gras RM, Colodro-Conde L. The GHQ-12 for the assessment of psychological distress of family caregivers. Behavioral Medicine. 2014;40(2):65-70.

48. Klassen A, Raina P, Reineking S, Dix D, Pritchard S, O'Donnell M. Developing a literature base to understand the caregiving experience of parents of children with cancer: a systematic review of factors related to parental health and well-being. Support Care Cancer. 2007;15:807-18.

49. Lindkvist M, Feldman I. Assessing outcomes for cost-utility analysis in mental health interventions: mapping mental health specific outcome measure GHQ-12 onto EQ-5D-3L. Health and Quality of Life Outcomes. 2016;14(134):1-9.

50. Barton, G. R., Bankart, J., Davis, A. C. & Summerfield, Q. A. (2004). Comparing Utility Scores before and after hearing-aid provision: Results according to the EQ-5D, HUI3 and SF-6D. Applied Health Economics and Health Policy, 3(2), 103–105

51. Marra CA, Woolcott JC, Kopec JA, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med. 2005;60(7):1571–82.

52. McDonough CM, Grove MR, Tosteson TD, Lurie JD, Hilibrand AS, Tosteson AN. Comparison of EQ-5D, HUI, and SF-36-derived societal health state values among spine patient outcomes research trial (SPORT) participants. Qual Life Res. 2005;14(5):1321–32.

53. Polinder S, Haagsma JA, Bonsel G, Essink-Bot ML, Toet H, van Beeck EF. The measurement of long-term health-related quality of life after injury: comparison of EQ-5D and the health utilities index. Injury Prevention 2010 Jun;16(3):147-53.54. Fals-Stewart W, Yates BT, Klostermann K. Assessing the Costs, Benefits, Cost-Benefit Ratio, and Cost-Effectiveness of Marital and Family Treatments: Why We Should and How We Can. Journal of Family Psychology. 2005;19(1):28-39.

55. Labelle RJ, Hurley JE. Implications of basing health-care resources allocations on cost-utility analysis in the presence of externalities. Journal of Health Economics. 1992;11:259-77.

56. Ng ESW, Grieve R, Carpenter JR. Two-stage nonparametric bootstrap sampling with shrinkage correction for clustered data. The Stata Journal. 2013; 13(1): 141-164.

Table 1: Description of clinical scores

Adolescent	**Hopelessness scale**: It measures the degree to which adolescents have negative expectancies about themselves and the future. It consists of 17 items with true or false responses, providing a single overall score with higher scores reflecting greater negative expectations towards the future.
Parent	**Family questionnaire**: It is a 20-item self-report questionnaire relating to the different ways in which families try to cope with everyday problems. It consists of a single overall score with higher scores indicating greater levels of expressed emotion directed at the adolescent by the parent. **McMaster Family Assessment Device**: It measures family functioning across 60-items on six different dimensions: Problem Solving, Communication, Roles, Affective Responsiveness, Affective Involvement, and Behaviour Control. Higher total score is indicative of poorer family functioning. **GHQ-12**: It is a measure of current mental health focusing on two major areas: the inability to carry out normal functions and the appearance of new and distressing experiences. High total scores are indicative of greater psychological distress.

Table 2: Adolescents' characteristics at baseline

		All (N=754)	**TAU (N=371)**	**FT (N=383)**
Gender	Males	n=93 (12%)	n=48 (13%)	n=45 (12%)
	Females	n=661 (88%)	n=323 (87%)	n=338 (88%)
Age	11-14 years old	n=396 (53%)	n=195 (53%)	n=201 (52%)
	15-17 years old	n=358 (47%)	n=176 (47%)	n=182 (48%)
Centre	Yorkshire	n=272 (36%)	n=135 (36%)	n=137 (36%)
	Manchester	n=267 (35%)	n=132 (36%)	n=135 (35%)
	London	n=215 (29%)	n=104 (28%)	n=111 (29%)
Total number of self- harm episodes	Mean (SD)	2.92 (21.51)	3.26 (28.59)	2.60 (10.95)
Type of index episode	Self-poisoning	n=170 (23%)	n=83 (22%)	n=87 (23%)
	Self-injury	n=533 (71%)	n=262 (71%)	n=271 (71%)
	Combined	n=51 (7%)	n=26 (7%)	n=25 (7%)
Source of referral (from hospital)	Yes	n=274 (36%)	n=130 (35%)	n=144 (38%)
	No	n=480 (64%)	n=241 (65%)	n=239 (62%)
EQ-5D-3L score (overall)	Mean (SD)	0.68 (0.27)	0.68 (0.26)	0.68 (0.28)
Hopelessness scale score^	Mean (SD)	7.39 (4.26)	7.21 (4.29)	7.56 (4.22)

^Hopelessness scale score was not available for 11 adolescents.
EQ-5D-3L: EuroQoL 5 Dimensions 3 Levels
SD: Standard Deviation

Table 3: Parents' characteristics at baseline

		All (N=754)	TAU (N=371)	FT (N=383)
Gender	Males	n=89 (12%)	n=47 (13%)	n=42 (11%)
	Females	n=665 (88%)	n=324 (87%)	n=341 (89%)
Relationship to adolescent	Father	n=85 (11%)	n=47 (13%)	n=38 (10%)
	Foster parent	n=2 (0%)	n=1 (0%)	n=1 (0%)
	Guardian	n=11 (1%)	n=3 (1%)	n=8 (2%)
	Mother	n=649 (86%)	n=318 (86%)	n=331 (86%)
	Step-father	n=2 (0%)	n=0 (0%)	n=2 (1%)
	Step-mother	n=5 (1%)	n=2 (1%)	n=3 (1%)
Age^	Mean (SD)	42.38 (6.42)	42.40 (6.18)	42.36 (6.64)
HUI score (overall)	Mean (SD)	0.71 (0.28)	0.70 (0.28)	0.72 (0.27)
McMaster Family Assessment Device^^	Mean (SD)	2.20 (0.37)	2.21 (0.37)	2.20 (0.36)
Family Questionnaire*	Mean (SD)	52.86 (10.75)	52.88 (10.79)	52.84 (10.72)
Parent GHQ**	Mean (SD)	5.70 (4.07)	6.07 (4.07)	5.33 (4.04)

^Age was not available for 81 caregivers.
^^McMaster Family Assessment Device was not available for 9 parents.
*Family Questionnaire was not available for 1 parent.
**Parent GHQ score was not available for 3 parents.
HUI: Health Utility Index; GHQ: General Health Questionnaire
SD: Standard Deviation

Table 4 – Adolescent's and parent's health-related quality of life by time period (N=754)

		Baseline	6 months	12 months
Overall				
Adolescent's EQ-5D-3L score	Mean (SD)	0.68 (0.27)	0.78 (0.17)	0.80 (0.19)
Parent's HUI score	Mean (SD)	0.71 (0.28)	0.76 (0.23)	0.78 (0.23)
Family therapy				
Adolescent's EQ-5D-3L score	Mean (SD)	0.68 (0.28)	0.80 (0.17)	0.81 (0.19)
Parent's HUI score	Mean (SD)	0.72 (0.27)	0.77 (0.23)	0.78 (0.23)
Treatment as usual				
Adolescent's EQ-5D-3L score	Mean (SD)	0.68 (0.26)	0.76 (0.17)	0.78 (0.18)
Parent's HUI score	Mean (SD)	0.70 (0.28)	0.76 (0.22)	0.78 (0.23)
Difference FT vs. TAU				
Adolescent's EQ-5D-3L score	Mean (SD)	-0.003 (0.20)	0.043*** (0.01)	0.036** (0.14)
Parent's HUI score	Mean (SD)	0.019 (0.02)	0.014 (0.03)	0.000 (0.02)

Table 5 – Relative health spillover – Results of Tobit regression model of the parent's health-related quality of life with adolescent's EQ-5D-3L score (full sample with imputations for missing data)

Variables	**Model 1**			**Model 2**		
	Baseline	6 months	12 months	Baseline	6 months	12 months
Adolescent						
EQ-5D-3L	-0.0077	0.2913***	0.0881	0.0558	0.2997***	0.1272**
Female	0.0895*	0.0516*	0.0332	0.0470	0.0314	0.0136
15-17yo vs. 11-14yo	-0.0231	0.0183	0.0162	0.0130	0.0078	0.0398
Type of index episode (ref. self-poisoning)						
Self-injury				0.0336	0.0752***	0.0119
Combined				0.0495	0.0594	0.0078
Repeated SH episodes (ref. >3 events)				-0.0357	-0.0565	0.0263
Hopelessness scale score				0.0108***	0.0047*	0.0050*
Parent						
McMaster Family Assessment				-0.1031**	-0.0702**	-0.0763*
Family Questionnaire				-0.0033**	-0.0027**	-0.0027*
Parent GHQ				-0.0378***	-0.0157***	-0.0150***
Female	-0.1062**	-0.0550*	-0.0511	-0.0218	-0.0213	-0.0101
Centre						
Manchester				-0.0865***	-0.0237	-0.0097
London				-0.0781**	-0.0062	0.0045
Constant	0.8176***	0.5914***	0.7564***	1.3256***	0.9341***	0.9594***
Sigma	0.3382	0.2536	0.2900	0.2702	0.2267	0.2699
Observations	754	754	754	731	731	731
Pseudo R-squared	0.017	0.084	0.010	0.426	0.459	0.151

***p<0.001, **p<0.01, *p<0.05; SH: Self Harm

HUI: Health Utility Index; GHQ: General Health Questionnaire; EQ-5D-3L: EuroQoL 5 Dimensions 3 Levels

yo: years old; ref.: reference; SH: Self-harm

Table 6 – Incremental cost-effectiveness ratios with alternative spillover quantifications

Scenario	Costs (SE)	QALY (SE)	ICER (£/QALY)	CE probabilities of FT (£20,000 – £30,000)[§]
Base-case analysis				
TAU	£3,750.59 (198.34)	0.745 (0.008)		
FT	£4,957.75 (194.13)	0.774 (0.008)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.030*** (0.011)	£40,453.30	(0.080 – 0.263)
Quantification 1 – Relative health spillover (Model 2 – Table 5)				
TAU	£3,750.59 (198.34)	0.940 (0.008)		
FT	£4,957.75 (194.13)	0.970 (0.008)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.030*** (0.011)	£40,453.30	(0.073 – 0.284)
Quantification 2 – Relative health spillover per treatment arm (Model 2 – Table 6)				
TAU	£3,750.59 (198.34)	1.001 (0.008)		
FT	£4,957.75 (194.13)	0.934 (0.008)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	-0.067***(0.011)	Dominated	(0.000 – 0.000)
Quantification 3 – Absolute health spillover (Model 2 – Table 7)				
TAU	£3,750.59 (198.34)	0.943 (0.008)		
FT	£4,957.75 (194.13)	0.972 (0.008)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.030** (0.012)	£40,838.11	(0.083 – 0.267)
Quantification 4 – Absolute health spillover per treatment arm (Model 2 – Table 8)				
TAU	£3,750.59 (198.34)	1.056 (0.019)		
FT	£4,957.75 (194.13)	0.905 (0.010)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	-0.150***(0.021)	Dominated	(0.001 – 0.004)
Quantification 5 – Additive health spillover[#]				
TAU	£3,750.59 (198.34)	1.492 (0.015)		
FT	£4,957.75 (194.13)	1.536 (0.014)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.044** (0.020)	£27,166.45	(0.297 – 0.568)
Equivalence scale health spillover with $a = 0$				
TAU	£3,750.59 (198.34)	1.492 (0.015)		
FT	£4,957.75 (194.13)	1.536 (0.014)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.044** (0.020)	£27,166.45	(0.279 – 0.539)
Equivalence scale health spillover with $a = 0.3$				
TAU	£3,750.59 (198.34)	1.351 (0.013)		
FT	£4,957.75 (194.13)	1.393 (0.013)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.042** (0.018)	£28,951.77	(0.245 – 0.531)
Equivalence scale health spillover with $a = 0.5$				
TAU	£3,750.59 (198.34)	1.273 (0.012)		
FT	£4,957.75 (194.13)	1.313 (0.012)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.040** (0.020)	£30,058.05	(0.202 – 0.487)

Equivalence scale health spillover with $a = 0.8$				
TAU	£3,750.59 (198.34)	1.174 (0.011)		
FT	£4,957.75 (194.13)	1.212 (0.011)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.038** (0.015)	£31,581.72	(0.169 – 0.475)
Equivalence scale spillover with $a = 1$				
TAU	£3,750.59 (198.34)	1.118 (0.010)		
FT	£4,957.75 (194.13)	1.155 (0.010)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,207.16*** (277.53)	0.037** (0.015)	£32,504.48	(0.164 – 0.426)

TAU: Treatment As Usual; FT: Family Therapy
vs.: versus; SE: standard error
QALY: Quality-Adjusted Life Years; ICER: Incremental Cost-Effectiveness Ratio
§ The cost-effectiveness probabilities of FT at £20,000 and £30,000 were estimated using the Stata command *tsbceprob* (Ng et al. 2013 [56])
***p<0.001, **p<0.01, *p<0.05;
#The adolescent's and parent's QALYs are summed, this is equivalent to $a = 0$

Appendix

Table A1– Relative health spillover per arm – Results of Tobit regression model of parent's health-related quality of life with adolescent's EQ-5D-3L score per treatment arm (full sample with imputations for missing data)

Variables	**Treatment as usual**			**Family Therapy**		
	Baseline	6 months	12 months	Baseline	6 months	12 months
Adolescent						
EQ-5D-3L	0.0612	0.3189***	0.3261***	0.0596	0.3100***	-0.0401
Female	0.0563	-0.0289	-0.0192	0.0484	0.1131**	0.0595
15-17yo vs. 11-14yo	-0.0480	-0.0050	-0.0072	0.0245	0.0234	0.0810**
Type of index episode (ref. self-poisoning)						
Self-injury	0.0096	0.0448	0.0263	0.0661	0.1175***	0.0123
Combined	0.0563	0.0679	-0.0049	0.0472	0.0691	0.0069
Repeated SH episodes (ref. >3 events)	-0.0290	-0.0395	0.0479	-0.0873	-0.0876*	0.0013
Hopelessness scale score	0.0115**	0.0059*	0.0102**	0.0085**	0.0033	0.0008
Parent						
McMaster Family	-0.0244	-0.0378	-0.0112	-0.1892***	-0.1163**	-0.1340**
Family Questionnaire	-0.0054**	-0.0034**	-0.0018	-0.0010**	-0.0021	-0.0026
Parent GHQ	-0.0390***	-0.0176***	-0.0178***	-0.0354***	-0.0137***	-0.0102**
Female	-0.0261	-0.0282	0.0038	-0.0310	-0.0116	-0.0331
Centre						
Manchester	-0.0640	-0.0081	-0.01812	-0.1212***	-0.0465	-0.0178
London	-0.0726	-0.0160	0.0049	-0.0836**	-0.0017	-0.0276
Constant	1.2351***	1.0202***	0.6121**	1.4349***	0.8424***	1.1679***
Sigma	0.2795	0.2348	0.2534	0.2560	0.2278	0.2576
Observations	359	359	359	372	372	372
Pseudo- R-squared	0.407	0.358	0.246	0.481	0.439	0.208

***p<0.001, **p<0.01, *p<0.05

HUI: Health Utility Index; GHQ: General Health Questionnaire; EQ-5D-3L: EuroQoL 5 Dimensions 3 Levels

yo: years old; ref.: reference; SH: Self-harm

Table A2 – Absolute health spillover – Results of Tobit regression model of the parent's health-related quality of life with adolescent's EQ-5D-3L score per repeated SH event at 12 months (full sample with imputations for missing data)

Variables	**No repeated self-harm at 12 months**			**Repeated self-harm at 12 months**		
	Baseline	6 months	12 months	Baseline	6 months	12 months
Adolescent						
EQ-5D-3L	0.0389	0.3023***	0.1567**	0.0923	0.3763***	0.0708
Female	0.0350	0.0240	-0.0141	0.1071	0.0788	0.1515
15-17yo vs. 11-14yo	-0.0026	0.0147	0.0311	0.0553	-0.0049	0.0661
Type of index episode (ref. self-poisoning)						
Self-injury	-0.0022	0.0571*	-0.0289	-0.0492*	0.1263**	0.1093*
Combined	0.0186	0.0628	-0.0586	-0.0856*	0.0735	0.1389
Repeated SH episodes (ref. >3 events)	-0.0094	-0.0344	0.0486	-0.2188**	-0.1219**	-0.0310
Hopelessness scale score	0.0084**	0.0027	0.0030	0.0179**	0.0131**	0.0160**
Parent						
McMaster Family Assessment	-0.0977**	-0.0706*	-0.0878*	-0.1269*	-0.1029	-0.0518
Family Questionnaire	-0.0032*	-0.0027*	-0.0017	-0.0038	-0.0036	-0.0057*
Parent GHQ	-0.0346***	-0.0143***	-0.0123***	-0.0434***	-0.0169**	-0.0148*
Female	-0.0309	0.0001	-0.0037	0.0023	-0.0826	-0.0685
Centre						
Manchester	-0.0892**	-0.0033	0.0178	-0.0492	-0.0838	-0.0567
London	-0.0726**	0.0044	0.0227	-0.0856	-0.0482	-0.1284*
Constant	1.3282***	0.2900***	0.5891***	1.3456***	0.7436***	0.7164**
Sigma	0.2639	0.2215	0.2433	0.2730	0.2511	0.2814
Observations	536	536	536	195	195	195
Pseudo- R-squared	0.407	0.376	0.189	0.526	0.417	0.273

***p<0.001, **p<0.01, *p<0.05
HUI: Health Utility Index; GHQ: General Health Questionnaire; EQ-5D-3L: EuroQoL 5 Dimensions 3 Levels
yo: years old; ref.: reference; SH: Self-harm

Table A3 – Absolute health spillover per arm – Results of Tobit regression model of the parent's health-related quality of life with adolescent's EQ-5D-3L score per repeated SH event at 12 months and per arm treatment (full sample with imputations for missing data)

Variables	**Treatment as usual**			**Family Therapy**		
	Baseline	6 months	12 months	Baseline	6 months	12 months
No repeated self-harm						
Adolescent						
EQ-5D-3L	0.0674	0.0878	0.3196***	0.0352	0.4587***	0.0039
Female	0.0610	0.1585	-0.0489	0.0180	0.0804	0.0309
15-17yo vs. 11-14yo	-0.0393	0.0195	-0.0087	0.0228	0.0207	0.0661**
Type of index episode (ref. self-poisoning)						
Self-injury	-0.0041	0.2310***	0.0276	0.0118	0.0620	-0.0631
Combined	0.1001	0.1639*	-0.0320	-0.0489	0.0011	-0.0862
Repeated SH episodes (ref. >3 events)	0.0435	-0.2209**	0.0666	-0.0537	-0.0446	0.0378
YP hopelessness scale score	0.0119**	0.0103	0.0103***	0.0053	0.0015	-0.0027
Parent						
McMaster Family Assessment	-0.0160	-0.1997***	-0.0114	-0.1765***	-0.0762	-0.1493***
Family Questionnaire	-0.0049**	-0.0032	-0.0016	-0.0011	-0.0018	-0.0014
Parent GHQ	-0.0368***	-0.0107	-0.0147***	-0.0321***	-0.0137***	-0.0104**
Female	-0.0609	-0.1393	-0.0031	-0.0193	-0.0040	0.0034
Centre						
Manchester	-0.0643	-0.0569	0.0185	-0.1220***	-0.0428	0.0120
London	-0.0769	0.0175	0.0461	-0.0716	0.0021	0.0113
Constant	1.1704***	1.5139	0.6429***	1.4540***	0.6033***	1.1036***
Sigma	0.2746	0.2146	0.2443	0.2440	0.2098	0.2331
Observations	270	270	270	266	266	266
Pseudo-R-squared	0.409	0.486	0.255	0.476	0.658	0.287
Repeated self-harm						
Adolescent						
EQ-5D-3L	0.0388	0.6925***	0.4393**	0.1371	0.0878	-0.1647
Female	0.0887	-0.0019	0.1264	0.1378	0.1585	0.1815
15-17yo vs. 11-14yo	0.0918	0.0207	-0.0065	0.0190	0.0195	0.1396**
Type of index episode (ref. self-poisoning)						
Self-injury	0.0143	-0.0021	0.0234	0.1809**	0.2310***	0.1810***
Combined	0.0944	-0.0241	0.0527	0.2296*	0.1639*	0.1468
Repeated SH episodes (ref. >3 events)	-0.3173**	-0.0911	0.0686	-0.1710	-0.2209**	-0.1419
Hopelessness scale score	0.0123	0.0128	0.0098	0.0200*	0.0103	0.0171**
Parent						
McMaster Family Assessment	-0.0071	0.0242	0.0267	-0.2147*	-0.1997***	-0.0785
Family Questionnaire	-0.0079*	-0.0042	-0.0024	-0.0001	-0.0032	-0.0078**
Parent GHQ	-0.0452***	-0.0212**	-0.0256**	-0.0466***	-0.0107	-0.0057
Female	0.0489	-0.0759	0.0217	-0.1224	-0.1393	-0.2255
Centre						
Manchester	0.0287	-0.1213	-0.0984	-0.0847	-0.0569	-0.0360
London	-0.0270	-0.0973	-0.1149	-0.1304	0.0175	-0.0909
Constant	1.4848***	0.8156*	0.2465	1.4498***	1.5139***	1.4929***
Sigma	0.2697	0.2444	0.2608	0.2632	0.2142	0.2700
Observations	89	89	89	106	106	106
Pseudo- R-squared	0.537	0.532	0.342	0.600	0.778	0.382

***p<0.001, **p<0.01, *p<0.05
HUI: Health Utility Index; GHQ: General Health Questionnaire; EQ-5D-3L: EuroQoL 5 Dimensions 3 Levels
yo: years old; ref.: reference; SH: Self-harm

Table A4: Average healthcare provider costs by trial arm after imputations

		Baseline to 6 months			6 to 12 months		
		All	**TAU**	**FT**	**All**	**TAU**	**FT**
Health and social services costs	Mean	£569.37	£650.80	£491.56	£385.10	£360.33	£409.08
	(SD)	(£857.95)	(£868.23)	(£841.82)	(£693.05)	(£672.93)	(£712.06)
	Min	£0	£0	£0	£0	£0	£0
	Max	£1,0994.00	£9,648.00	£10,994.00	£5,520.00	£4,894.00	£5,520.00
Hospital services costs (inpatient stays and A&E visits)	Mean	£62.61	£58.66	£66.38	£38.60	£30.61	£46.34
	(SD)	(£288.81)	(£186.43)	(£360.66)	(£144.91)	(£123.37)	(£162.86)
	Min	£0	£0	£0	£0	£0	£0
	Max	£6,415.50	£1,411.41	£6,415.50	£1,283.10	£1,026.48	£1,283.1
Hospital outpatient visits costs	Mean	£241.01	£236.66	£245.22	£246.88	£254.36	£239.64
	(SD)	(£633.57)	(£675.93)	(£590.51)	(£685.87)	(£740.84)	(£628.95)
	Min	£0	£0	£0	£0	£0	£0
	Max	£6,552.19	£6,552.19	£3,658.86	£8,522.76	£8,522.76	£5,035.13
Medication costs	Mean	£0.40	£0.21	£0.57	£0.39	£0.20	£0.57
	(SD)	(£7.62)	(£3.10)	(£10.24)	(£7.58)	(£2.91)	(£10.24)
	Min	£0	£0	£0	£0	£0	£0
	Max	£199.93	£59.06	£199.93	£199.93	£55.20	£199.93

Table A5 – Incremental cost-effectiveness ratios with alternative spillover quantifications (no imputations)

Scenario	Costs (SE)	QALY (SE)	ICER (£/QALY)	CE probabilities of FT (£20,000 – £30,000)[§]
Base-case analysis				
TAU (n=73)	£3,475.22 (380.19)	0.756 (0.021)		
FT (n=133)	£4,568.081 (259.32)	0.779 (0.016)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	0.024 (0.027)	£45,841.96	(0.219; 0.353)
Quantification 1 – Relative health spillover (Model 2 – Table 5)				
TAU (n=73)	£3,475.22 (380.19)	0.955 (0.021)		
FT (n=133)	£4,568.081 (259.32)	0.979 (0.016)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	0.024 (0.027)	£45,841.95	(0.242; 0.393)
Quantification 2 – Relative health spillover per treatment arm (Model 2 – Table 6)				
TAU (n=73)	£3,475.22 (380.19)	1.059 (0.021)		
FT (n=133)	£4,568.081 (259.32)	0.882 (0.016)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	-0.177 (0.027)	Dominated	(0.000; 0.000)
Quantification 3 – Absolute health spillover (Model 2 – Table 7)				
TAU (n=73)	£3,475.22 (380.19)	0.955 (0.021)		
FT (n=133)	£4,568.081 (259.32)	0.980 (0.016)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	0.026 (0.026)	£42,469.92	(0.225; 0.406)
Quantification 4 – Absolute health spillover per treatment arm (Model 2 – Table 8)				
TAU (n=73)	£3,475.22 (380.19)	1.018 (0.028)		
FT (n=133)	£4,568.081 (259.32)	0.920 (0.017)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	-0.097*** (0.031)	Dominated	(0.001; 0.000)
Quantification 5 – Additive health spillover[#]				
TAU (n=73)	£3,475.22 (380.19)	1.537 (0.041)		
FT (n=133)	£4,568.081 (259.32)	1.569 (0.027)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	0.032 (0.047)	£34,070.62	(0.545; 0.455)
Equivalence scale health spillover with $a = 0$				
TAU (n=73)	£3,475.22 (380.19)	1.537 (0.041)		
FT (n=133)	£4,568.081 (259.32)	1.569 (0.027)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	0.032 (0.047)	£34,070.62	(0.362 – 0.472)
Equivalence scale health spillover with $a = 0.3$				
TAU (n=73)	£3,475.22 (380.19)	1.390 (0.037)		
FT (n=133)	£4,568.081 (259.32)	1.421 (0.024)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	0.031 (0.042)	£35,796.37	(0.337 – 0.455)
Equivalence scale health spillover with $a = 0.5$				
TAU (n=73)	£3,475.22 (380.19)	1.308 (0.034)		
FT (n=133)	£4,568.081 (259.32)	1.338 (0.023)		
	Incremental Costs	*Incremental QALY*		

FT vs. TAU	£1,092.85** (449.22)	0.030 (0.040)	£36,841.44	(0.326 – 0.443)
Equivalence scale health spillover with $a = 0.8$				
TAU (n=73)	£3,475.22 (380.19)	1.204 (0.031)		
FT (n=133)	£4,568.081 (259.32)	1.233 (0.021)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	0.029 (0.036)	£38,251.46	(0.311– 0.433)
Equivalence scale spillover with $a = 1$				
TAU (n=73)	£3,475.22 (380.19)	1.146 (0.030)		
FT (n=133)	£4,568.081 (259.32)	1.174 (0.020)		
	Incremental Costs	*Incremental QALY*		
FT vs. TAU	£1,092.85** (449.22)	0.028 (0.035)	£39,089.31	(0.286 – 0.394)

TAU: Treatment As Usual; FT: Family Therapy
vs.: versus; SE: standard error
QALY: Quality-Adjusted Life Years; ICER: Incremental Cost-Effectiveness Ratio
§ The cost-effectiveness probabilities of FT at £20,000 and £30,000 were estimated using the Stata command *tsbceprob* (Ng et al. 2013 [56])
***p<0.001, **p<0.01, *p<0.05;
#The adolescent's and parent's QALYs are summed, this is equivalent to $a = 0$