HEALTH TECHNOLOGY ASSESSMENT

VOLUME 22 ISSUE 12 MARCH 2018
ISSN 1366-5278

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

David J Cottrell, Alex Wright-Hughes, Michelle Collinson, Paula Boston, Ivan Eisler, Sarah Fortune, Elizabeth H Graham, Jonathan Green, Allan O House, Michael Kerfoot, David W Owens, Eirini-Christina Saloniki, Mima Simic, Sandy Tubeuf and Amanda J Farrin

National Institute for Health Research

DOI 10.3310/hta22120

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

David J Cottrell,[1]* Alex Wright-Hughes,[2] Michelle Collinson,[2] Paula Boston,[1] Ivan Eisler,[3] Sarah Fortune,[1] Elizabeth H Graham,[2] Jonathan Green,[4] Allan O House,[1] Michael Kerfoot,[4†] David W Owens,[1] Eirini-Christina Saloniki,[1] Mima Simic,[5] Sandy Tubeuf[1] and Amanda J Farrin[2]

[1]Leeds Institute of Health Sciences, University of Leeds, Leeds, UK
[2]Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK
[3]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[4]Division of Neuroscience and Experimental Psychology, University of Manchester, Manchester, UK
[5]South London and Maudsley NHS Foundation Trust, London, UK

*Corresponding author
†In memoriam

Declared competing interests of authors: Amanda J Farrin is a member of the Health Technology Assessment (HTA) Clinical Evaluation and Trials Board and the HTA Commissioning Strategy Group. Allan O House is a member of the HTA Efficient Study Designs Board. Sarah Fortune worked as a Consultant Clinical Psychologist in the NHS prior to this project. Sandy Tubeuf is a member of the National Institute for Health Research Programme Grants for Applied Research Committee.

Published March 2018
DOI: 10.3310/hta22120

This report should be referenced as follows:

Cottrell DJ, 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 Technol Assess* 2018;**22**(12).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.236

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 15/141/03. The contractual start date was in November 2015. The draft report began editorial review in June 2016 and was accepted for publication in January 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

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

David J Cottrell,[1]* Alex Wright-Hughes,[2] Michelle Collinson,[2] Paula Boston,[1] Ivan Eisler,[3] Sarah Fortune,[1] Elizabeth H Graham,[2] Jonathan Green,[4] Allan O House,[1] Michael Kerfoot,[4†] David W Owens,[1] Eirini-Christina Saloniki,[1] Mima Simic,[5] Sandy Tubeuf[1] and Amanda J Farrin[2]

[1]Leeds Institute of Health Sciences, University of Leeds, Leeds, UK
[2]Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK
[3]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[4]Division of Neuroscience and Experimental Psychology, University of Manchester, Manchester, UK
[5]South London and Maudsley NHS Foundation Trust, London, UK

*Corresponding author d.j.cottrell@leeds.ac.uk
†In memoriam

Background: Self-harm in adolescents is common and repetition rates high. There is limited evidence of the effectiveness of interventions to reduce self-harm.

Objectives: To assess the clinical effectiveness and cost-effectiveness of family therapy (FT) compared with treatment as usual (TAU).

Design: A pragmatic, multicentre, individually randomised controlled trial of FT compared with TAU. Participants and therapists were aware of treatment allocation; researchers were blind to allocation.

Setting: Child and Adolescent Mental Health Services (CAMHS) across three English regions.

Participants: Young people aged 11–17 years who had self-harmed at least twice presenting to CAMHS following self-harm.

Interventions: Eight hundred and thirty-two participants were randomised to manualised FT delivered by trained and supervised family therapists ($n = 415$) or to usual care offered by local CAMHS following self-harm ($n = 417$).

Main outcome measures: Rates of repetition of self-harm leading to hospital attendance 18 months after randomisation.

Results: Out of 832 young people, 212 (26.6%) experienced a primary outcome event: 118 out of 415 (28.4%) randomised to FT and 103 out of 417 (24.7%) randomised to TAU. There was no evidence of a statistically significant difference in repetition rates between groups (the hazard ratio for FT compared with

TAU was 1.14, 95% confidence interval 0.87 to 1.49; $p = 0.3349$). FT was not found to be cost-effective when compared with TAU in the base case and most sensitivity analyses. FT was dominated (less effective and more expensive) in the complete case. However, when young people's and caregivers' quality-adjusted life-year gains were combined, FT incurred higher costs and resulted in better health outcomes than TAU within the National Institute for Health and Care Excellence cost-effectiveness range. Significant interactions with treatment, indicating moderation, were detected for the unemotional subscale on the young person-reported Inventory of Callous–Unemotional Traits ($p = 0.0104$) and the affective involvement subscale on the caregiver-reported McMaster Family Assessment Device ($p = 0.0338$). Caregivers and young people in the FT arm reported a range of significantly better outcomes on the Strengths and Difficulties Questionnaire. Self-reported suicidal ideation was significantly lower in the FT arm at 12 months but the same in both groups at 18 months. No significant unexpected adverse events or side effects were reported, with similar rates of expected adverse events across trial arms.

Conclusions: For adolescents referred to CAMHS after self-harm, who have self-harmed at least once before, FT confers no benefits over TAU in reducing self-harm repetition rates. There is some evidence to support the effectiveness of FT in reducing self-harm when caregivers reported poor family functioning. When the young person themselves reported difficulty expressing emotion, FT did not seem as effective as TAU. There was no evidence that FT is cost-effective when only the health benefits to participants were considered but there was a suggestion that FT may be cost-effective if health benefits to caregivers are taken into account. FT had a significant, positive impact on general emotional and behavioural problems at 12 and 18 months.

Limitations: There was significant loss to follow-up for secondary outcomes and health economic analyses; the primary outcome misses those who do not attend hospital following self-harm; and the numbers receiving formal FT in the TAU arm were higher than expected.

Future work: Evaluation of interventions targeted at subgroups of those who self-harm, longer-term follow-up and methods for evaluating health benefits for family groups rather than for individuals.

Trial registration: Current Controlled Trials ISRCTN59793150.

Funding: This project was funded by the NIHR Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 22, No. 12. See the NIHR Journals Library website for further project information.

Contents

List of tables

List of figures

List of abbreviations

A&E	accident and emergency
AE	adverse event
BNF	*British National Formulary*
BSS	Beck Scale for Suicide Ideation
CACE	complier average causal effect
CAMHS	Child and Adolescent Mental Health Services
CBT	cognitive–behavioural therapy
CDRS-R	Children's Depression Rating Scale – Revised
CI	confidence interval
CSO	Clinical Studies Officer
CTRU	Clinical Trials Research Unit
DBT	dialectical behaviour therapy
DMEC	Data Monitoring and Ethics Committee
DVD	digital versatile disc
EQ-5D	EuroQoL-5 Dimensions
FAD	(McMaster) Family Assessment Device
FT	family therapy
GHQ-12	General Health Questionnaire, 12 questions
GP	general practitioner
HES	Hospital Episode Statistics
HR	hazard ratio
HTA	Health Technology Assessment
HUI-3	Health Utilities Index 3
ICC	intracluster correlation coefficient
ICER	incremental cost-effectiveness ratio
ICU	Inventory of Callous–Unemotional Traits
ITT	intention to treat
MIU	minor injury unit
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NSSI	non-suicidal self-injury
OR	odds ratio
PPI	patient and public involvement
PQ-LES-Q	Paediatric Quality of Life Enjoyment and Satisfaction Questionnaire
PSSRU	Personal Social Services Research Unit
QALY	quality-adjusted life-year
REC	Research Ethics Committee
SAE	serious adverse event
SASII	Suicide Attempt Self-Injury Interview
SD	standard deviation
SDQ	Strengths and Difficulties Questionnaire
SE	standard error
SHIFT	Self-Harm Intervention: Family Therapy
SOFTA	System for Observing Family Therapy Alliances
TAU	treatment as usual
TMG	Trial Management Group
TSC	Trial Steering Committee
UKCP	UK Council for Psychotherapy
WIC	walk-in centre

Plain English summary

Young people (aged 11–17 years) who had self-harmed at least twice, and their families, were randomly allocated to receive either family therapy (FT) or treatment as usual (TAU) in their local Child and Adolescent Mental Health Services. Eight hundred and thirty-two young people from Yorkshire, Greater Manchester and London agreed to take part. The participants were recruited from hospital and the community.

Information was collected from the young people and their families at the beginning of the study and then again 3, 6, 12 and 18 months later in order to compare the effects of the two treatments. Information was also collected from health records.

The main focus was whether or not FT would reduce the number of times young people attended hospital with further self-harm. No significant differences were found in further self-harm between the two groups, nor was FT more cost-effective than the type of treatment young people usually get in the NHS (TAU).

Looking more closely at the characteristics of those taking part, there was a suggestion that young people who said that they found it harder to talk about feelings did better (self-harmed less) with TAU. On the other hand, when caregivers reported that the family did not talk about feelings easily, young people did better with FT.

Fewer emotional and behavioural problems were found in the group that had FT. FT was not found to be cost-effective.

Scientific summary

Background

Self-harm in adolescents is a major public health issue and, globally, suicide is the second most common cause of death in the 10–24 years age group after road traffic accidents. As many as 10% of adolescents self-harm in the community each year, with the most common methods being cutting and overdose. Only one in eight episodes of self-harm leads to a hospital presentation.

The estimates of the risk of 1-year repetition of self-harm vary between 5% and 25% per year. Actual rates may be much higher when repetition that does not come to clinical or medical attention is considered.

There is limited evidence for the effectiveness of clinical interventions for young people who engage in self-harm. Two recent studies have suggested that dialectical behaviour therapy and mentalisation-based treatment may be effective in reducing self-harm. Both had small numbers of participants and shorter follow-up periods than this study and relied on self-report as the primary outcome measure. A systematic review of interventions to reduce self-harm in adolescents calculated pooled risk differences comparing the proportion of young people who self-harmed at least once in the follow-up period of each study versus those who did not self-harm at all. Overall, the proportion of participants who self-harmed was slightly (but statistically significantly) lower in those allocated to treatment interventions. However, the authors acknowledged that the quality of studies examined was poor and that 'more research and replication of the positive findings by independent groups are urgently required' (Ougrin D, Tranah T, Stahl D, Moran P, Asarnow JR. Therapeutic interventions for suicide attempts and self-harm in adolescents: systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry* 2015;**54**:97–107).

Methods

Design

The Self-Harm Intervention: Family Therapy (SHIFT) trial was a pragmatic, Phase III, multicentre, individually randomised controlled trial of family therapy (FT) compared with treatment as usual (TAU) in 832 adolescents aged 11–17 years who had engaged in self-harm on at least two occasions and for whom a recent self-harm episode was a key reason for contact with Child and Adolescent Mental Health Services (CAMHS).

Objectives

The primary objective assessed the effectiveness of FT compared with TAU as measured by young people's rates of repetition of self-harm leading to hospital attendance 18 months after randomisation.

The secondary objectives assessed were:

- repetition rates of self-harm leading to hospital attendance 12 months after randomisation
- the cost per self-harm event avoided as a result of FT, measured using a structured, trial-specific health economics questionnaire
- the characteristics of all further episodes of self-harm (both those resulting in hospital attendance and self-report of all episodes)
- changes in a range of measures of participant and family functioning (see *Outcome measures*)
- moderator and mediators influencing benefit from treatment
- therapeutic engagement and adherence.

Setting and participants

The participants were young people aged 11–17 years who had self-harmed at least twice presenting to CAMHS following an episode of self-harm, recruited from NHS CAMHS across three 'hubs' in England: Greater Manchester, London and Yorkshire. Young people were screened for trial suitability and approached, if eligible, at their first visit to CAMHS following self-harm.

Interventions

The FT intervention was based on a modified version of the Leeds Family Therapy & Research Centre Systemic Family Therapy Manual. Qualified family therapists were appointed specifically to work on the trial, received standardised training and worked in teams of three or four, providing trial FT as a team for a cluster of CAMHS.

Treatment as usual was the care offered by local CAMHS teams to young people referred following self-harm. It was expected that TAU would be diverse and involve individual and/or family-orientated work, delivered by a range of practitioners with various theoretical orientations.

Outcome measures

The duration of treatment was designed to be approximately 6 months.

Measures were as follows: Inventory of Callous–Unemotional Traits [(ICU) young person and caregiver self-report at baseline], Family Questionnaire (caregiver self-report at 3 and 6 months), System for Observing Family Therapy Alliances (SOFTA; completed by family therapist and participants at FT session 3) and, at 12 and 18 months, the Suicide Attempt Self-Injury Interview (SASII) with the young person, Children's Depression Rating Scale – Revised (CDRS-R), health economics questionnaire, young person and caregiver self-report for the McMaster Family Assessment Device (FAD), Strengths and Difficulties Questionnaire (SDQ), young person self-report for the Beck Scale for Suicide Ideation (BSS), Hopelessness Scale for Children (young person), Paediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q), EuroQol-5 Dimensions [(EQ-5D) also at 6 months] and caregiver self-report for the General Health Questionnaire, 12 questions (GHQ-12), Health Utilities Index 3 [(HUI-3) also at 6 months] and health economics questionnaire.

Results

Characteristics of the sample

A total of 3554 young people were screened within participating CAMHS, of whom the clinician deemed 1603 (45.1%) to be eligible for the trial. The most common reason for a young person to be ineligible was that they had not engaged in self-harm prior to the current CAMHS referral. A total of 993 (61.9%) eligible young people consented to researcher contact, and 832 (83.8%) consented to trial participation and were randomised (51.9% of those eligible): 415 to FT and 417 to TAU.

The mean age at randomisation was 14.3 [standard deviation (SD) 1.38] years in the FT arm and 14.4 (SD 1.35) years in the TAU arm. In both arms there were more females than males: 368 (88.7%) in the FT arm and 369 (88.5%) in the TAU arm. More young people had self-harmed on at least three previous occasions than on two occasions: 369 (88.9%) in the FT arm and 370 (88.7%) the TAU arm. The type of most recent episode was most commonly self-injury, for 297 (71.6%) in the FT arm and 297 (71.2%) in the TAU arm, with self-poisoning attributed to a further 93 (22.4%) and 91 (21.8%), respectively. Those remaining used combined methods. All but two participants were living with their parents/guardians as opposed to in foster care. The majority were in full-time education: 398 (95.9%) in the FT arm and 386 (92.6%) in the TAU arm. Ethnicity was also well balanced between the arms.

Baseline characteristics suggest that participants had experienced significant difficulties and were not dissimilar to UK CAMHS referrals as a whole: 26.2% reported a health or disability problem, 29.3% had been involved with CAMHS in the past, 21.4% reported marked physical abuse and 16.6% reported sexual

abuse. On the total difficulties score of the SDQ, 66.2% of participants scored in the high/very high range, with 69.6% of caregivers reporting scoring participants in this range. On the general functioning subscale of the FAD, 84.7% of participants scored their families as 'unhealthy', with the equivalent figure from caregivers being 75.8%. On the CDRS-R, 65.7% of participants scored themselves as being in the moderate, severely or very severely depressed category. Nearly two-thirds of the participants (63.5%) were referred directly to CAMHS from the community: some had been discharged from hospital without a CAMHS referral and had then been referred via community services; others had never presented to hospital in the first place. The self-harm method used by the young people in the sample was much more slanted towards self-injury than in samples of hospital cases.

Clinical effectiveness

Primary outcome data were available for 795 out of 832 (95.6%) participants. A total of 221 (26.6%) young people experienced the primary outcome event, that is, a repeat self-harm event leading to hospital attendance within 18 months post randomisation: 118 (28.4%) in the FT arm and 103 (24.7%) in the TAU arm. There was no evidence to suggest a statistically significant difference in self-harm repetition rates between the treatment groups. The hazard ratio for FT compared with TAU was 1.14 [95% confidence interval (CI) 0.87 to 1.49] with a *p*-value of 0.3349.

Cost-effectiveness

Both trial arms showed an increase in the mean EQ-5D over 18 months' follow-up. The largest differences in EQ-5D scores between the two arms were at 6 and 12 months, with the FT group exhibiting higher scores at the 5% significance level than the TAU group, but there were no significant differences in quality of life between the two study arms at 18 months.

Family therapy participants incurred higher costs (mean £1266.23, 95% CI £736.04 to £1796.43) and gained more quality-adjusted life-years (QALYs) (mean 0.034, 95% CI –0.004 to 0.065) than TAU patients, equivalent to an extra 12.4 days of perfect health. The incremental cost-effectiveness ratio (ICER) equalled £36,811.80 per QALY, which is above the recommended threshold range currently specified for National Institute for Health and Care Excellence (NICE) decision-making in England and Wales (£20,000–30,000 per QALY gain). FT was unlikely to be cost-effective in most sensitivity analyses and was dominated by TAU in the complete-case analysis (less effective and more costly).

However, when combining young people's and caregivers' QALY gains, the FT arm incurred higher costs and exhibited better health outcomes than those in the TAU arm, resulting in an ICER of £20,808.21 per QALY gain; this ICER is within the NICE cost-effectiveness range, with a probability of being cost-effective of 41% at £20,000 (and 64% at £30,000) per QALY.

Secondary clinical outcomes

There were no significant treatment differences in young person questionnaire outcomes on the CDRS-R, PQ-LES-Q, Hopelessness Scale or FAD. However, adolescents treated with FT reported significantly better outcomes on the prosocial scale of the SDQ, with a mean improvement of 0.4 points (95% CI 0.1 to 0.7 points; $p = 0.0064$) at 12 months and of 0.3 points (95% CI 0.0 to 0.7 points; $p = 0.0337$) at 18 months, and on the impact of their problems scale at 12 months (mean improvement of –0.7 points, 95% CI –1.1 to –0.2 points; $p = 0.0033$), but not at 18 months (mean improvement –0.3 points, 95% CI –0.8 to 0.2 points; $p = 0.2153$). There was good evidence of reduced odds of suicide ideation in FT at 12 months, with an odds ratio of 0.64 (95% CI 0.44 to 0.94; $p = 0.0242$), but not at 18 months.

No significant treatment differences were found for the caregiver questionnaire outcomes on the GHQ-12 or Family Questionnaire. However, caregivers reported a range of significantly better outcomes on the SDQ for FT, with the following improvements in scores:

- total difficulties: mean –1.3 points (95% CI –2.4 to –0.2 points; $p = 0.0260$) at 12 months and mean –1.6 points (95% CI –2.9 to –0.4 points; $p = 0.0131$) at 18 months

- emotional problems: mean –0.5 points (95% CI –1.0 to –0.1 points; $p = 0.0166$) at 12 months and mean –0.6 points (95% CI –1.1 to –0.1 points; $p = 0.0218$) at 18 months
- peer problems: mean –0.3 points (95% CI –0.7 to –0.0 points; $p = 0.0366$) at 12 months and mean –0.5 points (95% CI –0.9 to –0.1 points; $p = 0.0092$) at 18 months
- internalising subscale: mean –0.9 points (95% CI –1.5 to –0.2 points; $p = 0.0111$) at 12 months and mean –1.1 points (95% CI –1.9 to –0.3 points; $p = 0.0074$) at 18 months
- at 18 months only, conduct problems: mean –0.3 points (95% CI –0.6 to –0.0 points; $p = 0.0499$) and externalising –0.7 points (95% CI –1.3 to –0.0 points; $p = 0.0446$)
- impact subscale: mean –0.7 points (95% CI –1.3 to –0.1 points; $p = 0.0309$) at 12 months only.

Caregivers in the FT arm also reported significantly better outcomes on the roles subscale of the FAD at 12 months, with a mean improvement of –0.1 points (95% CI –0.2 to –0.0 points; $p = 0.0020$), but not at 18 months.

The numbers of participants with other 'administrative' outcomes, such as referrals to other services, including to inpatient units, and safety outcomes, including re-referrals to CAMHS, accident and emergency (A&E) attendances and hospital admissions for any reason, were similar in both treatment arms.

Moderator analyses

Significant interactions with treatment, indicating moderation, were detected for the unemotional subscale on the young person-reported ICU ($p = 0.0104$) and for the affective involvement subscale on the caregiver-reported FAD, for both the score ($p = 0.0338$) and the categorisation of healthy versus unhealthy families ($p = 0.0444$).

Young people in the FT arm whose scores on the unemotional subscale suggested that they had difficulty talking about their feelings at baseline had higher risk of self-harm than those in the TAU arm, while those in the FT arm whose scores indicated that they found talking about their feelings to be easier had a lower risk of self-harm than those in the TAU arm.

Among young people whose caregivers reported higher affective involvement scores (the degree to which family members are involved and interested in one another) on the FAD, risk of self-harm was higher in the FT arm than in the TAU arm, while among those with lower affective involvement scores risk of self-harm was lower in the FT arm than in the TAU arm.

Conclusions

This study did not demonstrate that SHIFT manualised FT following repeated self-harm reduced subsequent hospital attendances for self-harm when compared with TAU.

The high proportion of young people whose index episode of self-harm involved self-injury and who were referred into CAMHS through the community rather than recruited directly following admission to hospital means that the sample is representative of self-harm referrals to CAMHS. However, the findings may not be generalisable to the smaller subset of adolescents who present to hospital following a first episode of self-harm.

There was some evidence to support the effectiveness of FT over TAU in reducing self-harm when caregivers reported poor family functioning, particularly in relation to talking about feelings, or young people reported ease in discussing emotions. Conversely, when the young people themselves reported difficulty in expressing emotion, or families reported healthy functioning on the affective involvement scale, FT was not as effective as TAU.

Although there was no evidence of cost-effectiveness of FT in the base-case analysis and most sensitivity analyses focused on health benefits to young people, there is a suggestion that FT may be cost-effective if health benefits to the caregiver are additionally taken into account.

There was clear evidence that FT had a statistically significant, positive impact on young people's prosocial behaviour at 12 and 18 months and on suicidal ideation at 12 but not 18 months and on caregivers' views on young people's total difficulties and emotional and peer problems at 12 and 18 months and conduct problems at 18 months.

Recommendations for future research

There remains a need for research exploring effective interventions to reduce self-harm. Self-harm is likely to be the final common pathway for a wide range of interpersonal and mental health predicaments. Future research needs to evaluate interventions targeted at the characteristics of specific subgroups who self-harm. Obvious candidate groups arising from this research would be families who self-report poorer family functioning and young people who are more unemotional.

Further research into the characteristics of these two groups is also indicated. What is the exact nature of the family dysfunction that some groups report and how might psychological interventions be targeted at this group? Are unemotional traits shared by other family members and is it possible that these two findings are aspects of the same underlying issue?

The accumulation of health benefits for the young person and the carer requires further exploration as to how health economic benefits might be aggregated for family members.

Studies with longer follow-ups are needed to explore any longer-term impact of interventions. The National Institute for Health Research Health Technology Assessment programme has already provided funding to allow follow-up of the SHIFT participants for a further 18 months, looking at the primary outcome only.

The significant differences observed in self-reported episodes of self-harm and episodes requiring hospital attendance and the very different patterns of self-harm recorded suggest that further work is needed to clarify the most appropriate outcome measures in self-harm research and how these might best be measured.

Trial registration

This trial is registered as ISRCTN59793150.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Introduction

What is self-harm?

In this study, in accordance with the commissioning brief and in line with current UK clinical practice, self-harm is defined 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 the degree of intention to die. This definition includes what in the USA would be described as non-suicidal self-injury (NSSI) and suicidal behaviour.[1]

How common is self-harm?

Self-harm in adolescents is a major public health issue and, globally, suicide is the most common cause of death in the 10–24 years age group after road traffic accidents.[2] A relatively recent systematic review reported that 9.7% of adolescents had self-harmed in the previous year in the community,[3] and a recent pan-European study reported a lifetime prevalence of 27.6% for NSSI,[4] with rates of self-harm higher among females than males, although in later adolescence there is nearly parity between male and female hospital attendances.[2]

At the community level, the most common methods of self-harm in young people are cutting and overdose.[5] Only one in eight episodes of self-harm leads to a hospital presentation,[5] but, even so, it is likely that > 30,000 adolescents present to hospital in England each year having harmed themselves.[6] In studies based on presentations to general hospitals in the UK, most adolescents have harmed themselves by taking an overdose, with self-poisoning with analgesics being particularly common.[7]

Self-harm is associated with an elevated risk of overall mortality[2,8–11] and suicide. In one follow-up study of 15- to 24-year-olds who had presented to hospital following an episode of self-harm between 2000 and 2005, the overall number of deaths from all causes was 3% of the cohort at follow-up in 2007, four times higher than expected. This was mainly because of an excess of suicides (2% of the cohort), which were 10 times more frequent than expected. The main risk factors for suicide were male gender, previous multiple episodes of self-harm, prior psychiatric history and high suicide intent.[12,13] Because of the young ages at which these deaths occur, the number of life-years lost to the community as a result of suicide and the impact on family members is significant.

The estimates of the risk of 1-year repetition of self-harm vary between 5% and 25% per year:[14,15] 18% in a recent UK multicentre monitoring study of > 5000 adolescents[16] and as high as 27% over an average of 5 years' follow-up of around 4000 adolescents.[13] In addition, actual rates may be much higher when repetition that does not come to clinical or medical attention is considered.[17] The risk of repetition persists for many years after an episode[9,18] and is associated with self-cutting (rather than overdose), depression, reports of childhood sexual abuse and exposure to models of self-harm.[2]

What is the aetiology?

Self-harm behaviour results from the convergence of a range of biopsychosocial risk factors across a number of domains, including genetic, biological, social, environmental and demographic factors, personality and cognitive styles and psychiatric morbidity.[19] Several reviews synthesise the full range of risk factors in adolescents.[2,14,20,21]

Family factors are also important.[11,22–24] Difficulties in parent–child relationships, including those related to early attachment problems, and perceived low levels of parental caring and communication are related to increased risk of suicide and self-harm among children and adolescents.[25] Exposure to sexual and physical abuse is an important risk factor.[25] Families are not always aware of, or do not have the capacity to respond to, the suicidal behaviour of their offspring,[26,27] particularly families in which there are low levels of social support. Parent reactions to self-harm include a range of strong and often negative emotions.[28]

Depression is the most common psychiatric diagnosis associated with suicide in adolescents[29–31] and with non-fatal self-harm,[20] and hopelessness is an important mediating variable between depression and self-harm.[32] Depression is also a key factor associated with repetition of self-harm in adolescents[33] and may also moderate responsiveness to treatment.[34]

Parental mental illness and substance abuse are significant risk factors,[14] and a family history of self-harm is associated with increased risk of suicide deaths[30,35–37] and non-fatal self-harm by adolescents.[5,30] However, although suicidal behaviour runs in families, this risk is over and above the heritability of mental illness between generations.[38–40]

Young people who self-harm experience higher rates of exposure to recent stressful life events such as rejection, conflict or loss following the break-up of a relationship, conflicts and disciplinary or legal crises,[5] and a strong association exists between adolescent self-harm and both childhood sexual abuse and physical abuse.[41]

What might we do to prevent self-harm?

An overall model of suicide behaviour provides a framework to support research studies in addition to formulations and treatment planning by clinicians. There is a long history of psychological and psychosocial models of suicidal behaviour such as Durkheim, Shneidman and, more recently, Beck, although these focus mainly on adults. Two relevant exceptions are Beautrais[19] and Bridge *et al.*,[14] who take a more developmental approach and are pertinent to the Self-Harm Intervention: Family Therapy (SHIFT) trial. Beautrais[19] posits that suicidal behaviours are the end point of adverse life events in which multiple risk factors combine to encourage their development. This approach represents a stress–diathesis model (for a review, see van Heeringen[42]), which suggests that temperamental and genetic factors and early experiences may make some young people particularly vulnerable to subsequent internal or external stressors. Bridge *et al.*[14] also proposed a developmental transactional model of behaviour, which recognises that most factors associated with suicidal behaviour among children and adolescents are familial. Levers for prevention include positive parent–child connections, active parental supervision and high behavioural expectations.

Building on this and the evidence described above, we suggest that the focus of family-orientated treatment with young people who have self-harmed should be on maximising cohesion, attachment, adaptability, family support and parental warmth, while reducing maltreatment and scapegoating and moderating parental control.[28]

However, there is limited evidence for the effectiveness of any clinical interventions for young people who engage in self-harm, with a failure to demonstrate any effect on reducing repetition of self-harm among adolescents receiving a range of treatment approaches, including therapeutic assessments and compliance enhancement in hospitals, involvement of youth-nominated support teams, tokens for hospital admission and home- and family-focused interventions.[21,43–45] Two recent studies have suggested that dialectical behaviour therapy[46] and mentalisation-based treatment[47] may be effective in reducing self-harm. Both had small numbers of participants, shorter follow-up periods than this study and relied on self-report as the primary outcome measure.

Brent *et al.*[48] concluded that interventions that activated family support or addressed motivations for change were quickly mobilised (to reflect the elevated risk of repetition immediately post episode) and promoted positive affect are the most likely to be able to demonstrate effectiveness. The reviewers also commented that the heterogeneity of treatment as usual (TAU) and the lack of its characterisation, in addition to the many underpowered studies, have held back advances in this field.[48]

Ougrin *et al.*[49] reviewed 19 studies describing interventions to reduce self-harm in adolescents. They calculated pooled risk differences using the outcome of the proportion of young people who self-harmed at least once in the follow-up period of each study versus those who did not self-harm at all. Overall, the proportion of participants who self-harmed was slightly (but statistically significantly) lower in those allocated to treatment interventions. However, the authors acknowledge that the quality of studies examined was poor and that 'more research and replication of the positive findings by independent groups are urgently required'.

The limited evidence that we have suggests that properly powered studies involving the family in treatment, and focusing treatment on both child and parent skills, characterises successful treatment approaches in this area.[50,51]

Chapter 2 Trial design and methods

Trial design

The trial was a pragmatic, Phase III, multicentre, individually randomised controlled trial of family therapy (FT) compared with TAU in 832 adolescents aged 11–17 years who had engaged in previous episodes of self-harm on at least two occasions and where a recent self-harm episode was a key characteristic of the reason for Child and Adolescent Mental Health Services (CAMHS) contact.

The primary outcome was attendance at hospital following self-harm at 18 months post randomisation.

The design and outcomes were influenced by the fact that this project was funded following a commissioned call for research (see *Appendix 1*).

Objectives and outcome measures

The primary objective was to assess the effectiveness of FT compared with TAU as measured by young people's rates of repetition of self-harm leading to hospital attendance 18 months after randomisation.

This outcome was selected following discussion with clinical colleagues and our review of the literature. The following issues were considered:

- high loss to follow-up in other self-harm intervention studies
- clinical reports that young people were not always regular attenders following self-harm
- wide fluctuations in rates of self-harm, with some young people engaging in self-harm many times per day, complicating outcome measurement in other studies
- high rates of self-harm in young people who do not present to hospital but who are difficult to identify and follow up.

We prioritised an outcome that could be operationally defined and measured even if participants were difficult to follow up. This gave us the potential to assist in answering the clinical question 'what is the most effective course of action when a young person presents to CAMHS following self-harm'. The implications of this decision are discussed further in the results (see *Chapters 3* and *4*) and discussion (see *Chapter 5*).

Secondary objectives were to assess:

- the effectiveness of FT compared with TAU as measured by repetition rates of self-harm leading to hospital attendance at 12 months after randomisation
- the characteristics of all further episodes of self-harm leading to hospital attendance
- the cost per self-harm event avoided as a result of FT, measured using a structured, trial-specific health economics questionnaire
- the characteristics of all further episodes of self-harm (not just those resulting in hospital attendance). This included the number of subsequent self-harm events, time to next event, severity of event (fatal, near fatal or not) and dangerousness of method used, as measured by the Suicide Attempt Self-Injury Interview (SASII)[52]
- suicidal ideation, measured by the Beck Scale for Suicide Ideation (BSS)[53]
- quality of life, measured by the Paediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q)[54] and parental completion of the General Health Questionnaire, 12 questions (GHQ-12)[55]
- depression, measured by the Children's Depression Rating Scale – Revised (CDRS-R)[56]

- overall mental health and emotional and behavioural difficulties via young person and parental completion of the Strengths and Difficulties Questionnaire (SDQ)[57]
- hopelessness, via completion of the Hopelessness Scale for Children[58]
- family functioning, measured by the McMaster Family Assessment Device (FAD)[59] and the Family Questionnaire[60]
- mediator and moderator variables that influence engagement with and benefit from treatment (e.g. number of sessions, medication use, referrals)
- therapeutic alliance to FT, via family therapist, young person and parental completion of the System for Observing Family Therapy Alliances (SOFTA)[61]
- therapist adherence to the FT manual.

Participants and setting

Participants were young people who met the eligibility criteria detailed below. In addition, for each young person a primary caregiver was identified who was willing to be involved in the research to provide outcome measures.

Young people and their caregivers were identified from NHS CAMHS across three 'hubs' in England: Greater Manchester, London and Yorkshire. Young people were screened for trial suitability and approached, if eligible, at their first visit to CAMHS following self-harm. To be included in the trial, young people were required to meet the following eligibility criteria and subsequently provide written informed consent, together with their primary caregiver.

Inclusion criteria

1. Aged 11–17 years.
2. Self-harmed prior to assessment by the CAMHS team.
3. Engaged in at least one previous episode of self-harm prior to the index presentation.
4. Assessed in hospital following current episode, or referred directly to CAMHS from primary care with recent self-harm as a key feature of presentation.
5. When the presenting episode was because of alcohol or recreational drugs, the young person had explicitly stated that he or she was intending self-harm by use of these substances.
6. The clinical intention was to offer CAMHS follow-up for self-harm.
7. Lived with primary caregiver.

Exclusion criteria

1. At serious risk of suicide (clinical judgement) and so requiring more than TAU.
2. An ongoing child protection investigation within the family, which would have made treatment difficult to deliver.
3. Would ordinarily have been treated not in generic CAMHS but rather by a specific service, and so would not normally have received TAU.
4. Pregnant at time of trial entry and thus likely to have to interrupt treatment in either arm.
5. Actively being treated in CAMHS (as the possibility of randomisation might disrupt ongoing therapy).
6. In a children's home or short-term foster placement and so no continuity of treatment possible in either arm.
7. Moderate to severe learning disability or lacked capacity to comply with trial requirements.
8. Involved in another research project – at the time of trial entry or within the last 6 months so as not to overburden participants and to avoid possibility of conflicting requirements across two studies.
9. Sibling had been randomised to the SHIFT trial or was receiving FT within CAMHS. As the intervention studied is a family intervention, if a sibling was already in the trial, randomisation would not have been possible.
10. The young person and one main caregiver had insufficient proficiency in English to contribute to the data collection.

Recruitment

All young people presenting via hospital or from primary care to participating CAMHS following a recent episode of self-harm were screened by CAMHS clinicians. When a young person met the criteria, the assessing CAMHS clinician discussed trial participation at the assessment appointment, passed on the participant information sheets and requested consent (written ideally, but verbal was allowed) for subsequent contact by a researcher. Those providing such consent were contacted by the researcher, who arranged to visit the family at home, discussed the trial in detail, obtained consent for participation and administered the baseline assessment.

Randomisation and blinding

Following consent and baseline assessment, researchers randomised participants sequentially via an automated system at the Clinical Trials Research Unit (CTRU) at the University of Leeds. Participants were randomly allocated on a 1 : 1 basis to receive FT or TAU through the use of a computer-generated minimisation program incorporating a random element. The stratification factors were centre (CAMHS team), gender, age (11–14 years/15–17 years), living arrangements (with parents or guardians/foster care), number of previous self-harm episodes including index event (2/≥ 3) and type of index episode (self-poisoning, self-injury, combined). When therapists were not aligned to a specific service but covered a number of services, additional randomisation of the lead therapist took place within the FT team in order to minimise case selection bias. This occurred across the whole Manchester hub, in two sites in Yorkshire and in two sites in London.

The participants and therapists were, of necessity, aware of treatment allocation, but the researchers were blind to allocation to allow the unbiased collection of follow-up data. The CTRU was responsible for informing CAMHS clinicians and family therapists of randomisation outcome in order to maintain researcher blinding. Clinicians and family therapists informed families of their allocation and arranged subsequent appointments for treatment in CAMHS.

Interventions

Both therapeutic interventions were delivered within CAMHS, and all participants were treated within their local service. Family therapists were formally linked with specific CAMHS teams to ensure clear lines of clinical responsibility, and clinicians in both trial arms had access to local child and adolescent psychiatrists if medication or hospitalisation needed to be considered.

Family therapy

The FT intervention was based on a modified version of the Leeds Family Therapy & Research Centre Systemic FT Manual, the development and validation of which was funded by the Medical Research Council to support trials of FT.[62] This manual was updated by the FT expert members from the Trial Management Group (TMG) to ensure that it was appropriate for work with families following self-harm. The SHIFT FT Manual can be requested from the trial team via http://medhealth.leeds.ac.uk/SHIFTManual.

Our systemic orientation to self-harm assumes that there is an interplay between a number of factors that can lead to the development of individual difficulties and problems such as self-harm. However, rather than focus on the causal explanation of individual problems, systemic FT is concerned with the way in which these problems will have become embedded in the matrix of family and wider social relationships, the felt experiences and the meanings and narratives that have become attached to and that shape these difficulties.[63] Looking for a causal account in the individual case is seen as unhelpful, partly because it inevitably oversimplifies the complexity of the individual aetiological pathway and partly because it tends to reinforce disabling, blaming narratives.[64]

Self-harm in an individual case will, in part, have been a response and have meaning in relation to particular issues (e.g. problems in emotional regulation, sense of hopelessness, low self-esteem, developmental and life-cycle issues, school and peer issues, attachment difficulties, relationship problems, parental situations or problems in family communication), but it also highlights that, for some reason, alternative, more constructive solutions were unavailable. The self-harming response will have been influenced by and in turn will influence subsequent patterns of interaction and will give new meaning to them.[65] It becomes a means of communicating and constrains alternative responses. By considering these behavioural and linguistic patterns in greater detail, facilitating understanding and noticing and supporting alternatives that encourage more positive narratives to emerge, the therapist can help the client and family create a different response pattern to the issues that showed themselves in self-harm.

Family therapists [those eligible for registration with the UK Council for Psychotherapy (UKCP)] were appointed specifically to work on the trial, following the formal advertisement of a clear job description and subsequent interview. All were qualified systemic family therapists who had trained in the UK to the level of a master's degree, which included 2 years of supervised practice. The therapists had a primary professional mental health background and qualification (generally social work or nursing), and most of them had experience in CAMHS and post-qualification experience in FT practice.

The therapists joining SHIFT at the outset attended a 2-day group training event prior to the start of the trial and then 1-day training events annually thereafter. The authors of the manual provided the training, which included review and discussion of the manual, reflections on their own experiences with/as young people, their relationship to clinical risk and team building. Additional material on adolescent self-harm was presented by a specialist on the trial. Role play and experiential exercises were included. Each family therapist worked with a pilot case (CAMHS non-trial case) with supervision. After each therapist had completed a number of sessions, they were deemed ready to commence with trial participants. The orientation of replacement family therapists for departing original trial therapists involved one-to-one training with one of the trial's senior family therapists, a period of observation of the remaining team members' therapy and a one-to-one session with the supervisor to review the material.

Family therapists worked in teams of three or four and provided trial FT as a team for a cluster of CAMHS. Within each team, each family therapist would be the lead for a subset of participants (the other family therapists in the team would act as observers and make only a small face-to-face contribution for those participants).

Monthly group supervision of 2 hours was provided for each team by senior family therapists (two of whom were authors of the treatment manual). Supervision included discussion of cases, adherence to the manual, broader trial issues and team dynamics. Generally, the focus of the supervision was determined by the family therapists at each meeting.

Young people and their families were offered FT sessions of approximately 1.25 hours' duration each, delivered over 6 months at approximately monthly intervals, but with more frequent initial appointments. This equated to approximately eight sessions, but there was the expectation that some participants would receive fewer sessions because of dropout or mutually agreed termination of treatment. Equally, it was anticipated that some participants might receive more sessions if this was deemed clinically appropriate.

Wherever possible, and when consent was provided, the sessions were video-recorded, as this is part of good FT practice and facilitates supervision.

Originally, it had been intended that supervision would include a review of recordings of current cases. This proved difficult for technical reasons and so most supervision was based on therapist presentation. The adherence measure was developed at a later stage in the trial and applied retrospectively.

A selection of video-recorded sessions was centrally reviewed, whereby the first and a subsequent therapy session were reviewed for each therapist to monitor their adherence to the manual and to allow the reporting of this. Adherence was assessed through use of a tool developed specifically for the trial.[66] In addition, administrative processes that were required in the manual were also monitored, for example number of sessions offered and attended and whether or not formulation letters (summarising the therapist's understanding of the family situation) were sent to families following attendance at their second therapy session, as expected in the manual.

Treatment as usual

Treatment as usual was the care offered by local CAMHS teams to young people referred following self-harm. It was expected that this treatment would be diverse and involve individual and/or family-orientated work, delivered by a range of practitioners with various theoretical orientations. As SHIFT was a pragmatic trial involving a number of collaborating CAMHS teams, it was not deemed possible or appropriate to specify TAU, although it was expected that CAMHS practitioners would be working in line with best practice as set out in several National Institute for Health and Care Excellence (NICE) guidelines (e.g. guidance on self-harm and depression in childhood).[67,68] Generic TAU was not delivered to the FT group as part of their clinical intervention, unless specific additional assessment or treatment was indicated during or after FT, for example referral for formal mental state assessment or for medication.

Contamination

The possibility of cross-arm contamination was considered during the design stage of the trial, with the following points mandated whenever possible and monitoring processes implemented:

- Different teams of therapists delivered the two interventions in each CAMHS site and SHIFT family therapists were prohibited from treating participants in the TAU arm for the duration of the trial. There were processes in place to record contamination if this occurred.
- Owing to the nature of appointment scheduling and the fact that this was family-specific therapy (i.e. not a group intervention), there was little opportunity for participants to meet and discuss treatment, so contamination at the participant level was very unlikely.
- Any family-orientated clinical interventions in the TAU group were likely to be different from the trial FT intervention, which required adherence to the Leeds Family Therapy & Research Centre manual, fully-trained family therapists eligible for UKCP registration, therapy delivered in a team context and regular supervision. Thus, use of family interventions in the TAU arm was not prohibited, but details of all treatment received by participants in both arms were recorded.

Assessments and data collection

All required data, assessment tools, collection time points and processes are summarised in *Table 1*.

Screening

Data were recorded anonymously for young people who did not meet the eligibility criteria, and for those who were eligible but not willing to provide consent, in order to monitor trial uptake and representativeness of the trial population. Screening data for randomised participants were linked to the main trial data, and also provided details of the outline 'index' event that had prompted their referral to CAMHS and eligibility for the trial. Screening data, eligibility and consent for researcher contact were assessed and recorded by the screening clinician. During the baseline visit, the researcher confirmed the participant's eligibility, obtained and recorded consent for trial participation, and obtained further details, including current physical and mental health comorbidities, psychotropic medications and history of abuse.

Participant assessments and follow-up

At the baseline assessment the researcher administered questionnaires to the young person, and self-reported questionnaires were also completed by the young person and caregiver.

TABLE 1 Data collection schedule

Assessment (including who is involved)	Timeline (months post randomisation)				
	Baseline	3	6	12	18
Eligibility and consent					
Eligibility (assessed by clinician)	✗				
Consent (young person, P, R)	✗				
Background and demographics (young person, P, R – interview and case notes)					
Personal details	✗				
Outline 'index' event details	✗				
Current comorbid physical/mental health	✗				
Current psychotropic medications	✗				
History of abuse	✗				
Follow-up data (collected from case notes)					
Therapy details (provided by therapist)			✗	✗	
Therapist supervision details (provided by therapist/supervisor)			✗	✗	
Details of further self-harm episodes since consent (R)				✗	✗
Psychotropic medication details (R)				✗	✗
Referrals to other MH services (R)				✗	✗
Re-referral to CAMHS (R)				✗	✗
Admissions to hospital relating to mental health (R)				✗	✗
All-cause mortality (NHS Digital)					✗
SAE reporting and hospital attendance (R and NHS Digital)	Ongoing collection				
Questionnaires (completed at researcher visit unless otherwise stated)					
Family Questionnaire (P self-report, CTRU postal administration at 3 and 6 months)	✗	✗	✗		
SOFTA (completed by the family therapist and participants at FT session 3)		✗			
SASII (interview with young person)	✗			✗	✗
BSS (young person self-report)	✗			✗	✗
Hopelessness Scale for Children (young person self-report)	✗			✗	✗
McMaster FAD (young person and P self-report)	✗			✗	✗
GHQ-12 (P self-report)	✗			✗	✗
SDQ (young person and P self-report)	✗			✗	✗
CDRS-R (Interview with young person)	✗			✗	✗
PQ-LES-Q (young person self-report)	✗			✗	✗
ICU (young person self-report)	✗				

TABLE 1 Data collection schedule (*continued*)

Assessment (including who is involved)	Timeline (months post randomisation)				
	Baseline	3	6	12	18
EQ-5D (young person self-report, postal administration at 6 months)	✗		✗	✗	✗
HUI-3 (P self-report, postal administration at 6 months)	✗		✗	✗	✗
Health Economics questionnaire (young person and P self-report, postal administration at 3 and 6 months)	✗	✗	✗	✗	✗

EQ-5D, EuroQol-5 Dimensions; HUI-3, Health Utilities Index 3; ICU, Inventory of Callous-Unemotional Traits; P, parent/caregiver; R, researcher; SAE, serious adverse event.

Postal questionnaires were administered by CTRU at 3 and 6 months post randomisation. These were preceded by a telephone call from researchers, when it was possible to make contact. If questionnaires were not returned, postal reminders were sent 2 weeks after the initial mailing and then again a further 2 weeks later.

Further follow-up to administer questionnaires took place at 12 and 18 months following randomisation via face-to-face researcher interviews. When it was not possible to arrange face-to-face interviews and when participants agreed, self-reported questionnaires were posted to participants for completion. The offer of incentives to young people for participating in the 12- and 18-month follow-up interviews was introduced in August 2013 (see amendments in *Appendix 2*) in an effort to improve compliance.

Clinical follow-up

The primary outcome measure was obtained from accident and emergency (A&E) and inpatient Hospital Episode Statistics (HES) data downloads from NHS Digital. This method of primary outcome data collection was augmented by directed hospital record searches, which were undertaken by researchers at frequent intervals throughout the trial. Researchers searched acute trust records for further details of all attendances that were unclear from the central HES data (i.e. where self-harm relatedness or method of self-harm could not be determined from HES data alone) and for any hospital attendances for those participants who had not consented to data sharing with NHS Digital or could not be linked to HES. Given that HES data sets are England-wide, this maximised the collection of hospital attendance data, allowing for participants moving out of the catchment area (within England). The collection of these routine data also minimised bias by eliminating the possibility of preferential researcher data collection at certain acute trusts. Further details of the comparison of HES records and researcher-collected hospital records and the success of using HES data for the trial will be published separately.

Data relating to treatment received in CAMHS following trial entry were provided by the treating CAMHS clinicians or family therapists, by clinical studies officers (CSOs) employed by the Research Networks in each locality and by the researchers when blinding was no longer required (i.e. after participant-reported follow-up data had been collected). Data included initial session attendance (FT and TAU), therapeutic approach (for TAU), referrals within CAMHS for additional or alternative sessions, requirements for psychotropic medications, liaison and referrals to other agencies outside CAMHS and re-referral to CAMHS following discharge. In addition, treating clinicians and participants completed the SOFTA[44] at the end of the third treatment session. Supervision sessions for family therapists and routinely provided clinical supervision for CAMHS clinicians were also recorded. Details of the scoring of questionnaires are in *Appendix 3*. These appendices provide further detail concerning the collection of clinical follow-up data.

Safety

Non-serious adverse events (AEs) were operationally defined in SHIFT as treatment on an emergency outpatient basis and re-referral to CAMHS, as these events were expected to occur within the adolescent study population. Deaths and hospital admissions were defined as serious adverse events (SAEs).

Adverse event data were obtained through researcher collection of data from CAMHS and acute trusts and via data transfer from NHS Digital.

Assessment instruments

Assessment instruments were incorporated into participant assessment packs for both the young person and their caregiver. The standardised questionnaires used were selected following our literature review because they related to factors thought to be important secondary outcomes in their own right or possible mediators or moderators of treatment effects. In selecting individual measures, we also tried to use those measures that would enable comparisons with other studies in the field. In one case [the Inventory of Callous–Unemotional Traits (ICU)] we were influenced by our knowledge of a large randomised controlled trial about to commence at the same time as SHIFT with a similar age group that we wanted to be able to compare with our sample. All measures were appropriate to the age range of young people and have been widely used in a range of ethnic communities. *Appendix 3* contains details of scoring.

Suicide Attempt Self-Injury Interview

The researcher completed the SASII[52] during the interview with the young person. The SASII is designed to assess the factors involved in non-fatal suicide attempts and intentional self-injury. It contains six screening items, nine open-ended questions to provide information for interviewer coding, and 22 items and associated subitems measuring timing and frequency of self-injurious acts, methods used and lethality of the method, suicidal as well as non-suicidal intent associated with the episode, communication of suicide intent before the episode, impulsivity and rescue likelihood, physical condition and level of medical treatment. The SASII also collects a timeline of all self-reported self-harm with details of the timing, methods and treatment received. At baseline the SASII focused on the index episode of self-harm and provided a timeline over the preceding 12 months; at 12-month completion the focus was on the first post-randomisation episode and the timeline since baseline; and, at 18 months, the focus was on the first episode following the 12-month researcher visit and a timeline since the 12-month researcher visit.

Inventory of Callous–Unemotional Traits

The ICU[71] was completed by the young person and by the caregiver in reference to the young person. It is a 24-item questionnaire in which participants are asked to indicate how well each statement describes them on a 4-point scale and is designed to provide a comprehensive assessment of callous and unemotional traits, with three subscales: callousness, uncaring and unemotional. Higher scores represent higher callous and unemotional traits. ICU measures are helpful in predicting different outcomes and developmental pathways in children with conduct disorder.

Beck Scale for Suicide Ideation

The BSS[53] was completed by the young person and is used to examine suicidal intent. The first 19 of 21 items measure the severity of actual suicidal wishes and plans, each on a 3-point scale, with a higher score indicating a higher level of suicide ideation. The first five items also serve as a screen for suicide ideation.

Paediatric Quality of Life Enjoyment and Satisfaction Questionnaire

The PQ-LES-Q[54] was completed by the young person and is used to examine quality of life. It consists of 15 items, each on a 5-point scale, providing a single overall score with higher total scores indicative of greater enjoyment and satisfaction.

Children's Depression Rating Scale – Revised

The CDRS-R[56] was completed by the researcher during the young person interview and is used to assess the severity of depressive syndrome. It is a brief rating scale based on a semistructured interview rating 17 symptom areas on a 5- to 7-point scale, including impaired schoolwork, difficulty having fun, appetite disturbance, excessive fatigue and low self-esteem. Higher scores are indicative of greater depression and are categorised to indicate mild, moderate, severe and very severe depression.

Strengths and Difficulties Questionnaire

The SDQ[57] was completed by the young person and the caregiver, giving their views of the young person's behaviours, and is a brief behavioural screening questionnaire assessing levels of emotional and behavioural problems. Both young person and caregiver versions ask about 25 attributes, each with three responses, across five scales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and prosocial behaviour. A 'Total difficulties' score combines all scales except the prosocial scale. 'Externalising' and 'Internalising' scores are also generated and an impact supplement is included in the caregiver version, in which the respondent is asked whether or not they think the young person has a problem, enquiring further about chronicity, distress, social impairment and burden to others. Higher scores in all but the prosocial score represent greater issues in that category. For the prosocial score, lower scores represent greater issues. Responses are also categorised according to a four-band categorisation.

Hopelessness scale

The Hopelessness Scale[58] was completed by the young person and is used to measure the degree to which young people 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 hopelessness or negative expectations towards the future.

McMaster Family Assessment Device

The McMaster FAD[59] was completed by the young person and caregiver. It measures family functioning across 60 items, each on a 4-point scale measuring agreement, based on the McMaster Model on six different dimensions: problem-solving (ability to resolve problems), communication (exchange of clear and direct verbal information), roles (division of responsibility for completing family tasks), affective responsiveness (ability to respond with appropriate emotion), affective involvement (degree to which family members are involved and interested in one another) and behaviour control (manner used to express and maintain standards of behaviour); it also includes an independent dimension of general functioning (overall functioning of family). Higher scores are indicative of poorer family functioning. Miller *et al.*[72] documented clinical cut-off scores differentiating 'healthy' versus 'unhealthy' family functioning for each dimension (see *Appendix 3*).

Family Questionnaire

The Family Questionnaire[60] was completed by the consenting caregiver in reference to the young person and is a brief 20-item self-report questionnaire relating to the different ways in which families try to cope with everyday problems. It assesses expressed emotion directed at the young person by family members with responses on a 4-point scale. In addition to the total score, it has two subscales: critical comments (unfavourable statements about the personality or behaviour of the young person) and emotional overinvolvement (overintrusiveness, protectiveness and identification). Higher scores indicate greater levels of expressed emotion.

Parent General Health Questionnaire, 12 questions

The GHQ-12[73] was completed by the caregiver and 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. It includes 12 items, each on a 4-point scale to indicate whether or not symptoms of mental ill health are present, with high total scores indicative of greater psychological distress.

EuroQol-5 Dimensions

The EuroQoL-5 Dimensions (EQ-5D)[74] was completed by the young person. The respondents are asked to describe their levels of health problems using a descriptive system comprising five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each with three levels: no problems, some problems, extreme problems. The combination of answers leads to a health profile of five digits (243 unique health profiles) that can then be converted into a utility that represents the overall quality of life of the young person. The EQ-5D utility score varies from 0 to 1, with 0.0 meaning death and 1.0 complete health. While the original protocol suggested the use of the Health Utilities Index 3 (HUI-3) to measure the young person's quality of life, we undertook a pilot study in a school setting and decided that the EQ-5D was easier for young people to complete.[75]

Health Economics Questionnaire

The Health Economics Questionnaire was completed by the young person and the caregiver. The caregiver reported health care use for the young person as well as for their own care. The questionnaire was designed for use in the SHIFT trial to collect information on A&E attendances, inpatient/outpatient attendances, primary/community care provided by the NHS/social services and medications. Data on personal costs were also collected, such as out-of-pocket expenses and productivity costs to both young persons and caregivers and any use of education and justice services. A reduced self-reported questionnaire was used within the trial from the autumn of 2011; it excluded questions on A&E attendances, inpatient attendances as this was obtained via data downloads from NHS Digital records and questions on private expenses and time off work. The recall period was the past 3 months at every data collection point.

Health Utilities Index 3

The HUI-3[76] was completed by the caregiver and is widely used in population health surveys, clinical studies and cost–utility analyses. The HUI3 is a health status classification system including eight attributes (vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain), each with five or six levels of ability/disability. Answers are then scored using a multiattribute scoring function derived from community preference measures for health states. The scoring systems provide a utility score on a generic scale where dead = 0.00 and perfect health = 1.00.

System for Observing Family Therapy Alliances

The SOFTA[61] was completed in clinic by the lead family therapist, young person and caregiver after the third therapy session and is a self-reported questionnaire used to evaluate the strength of the therapeutic alliance in FT. If the consenting primary caregiver was not present, it could be completed by another member of the family, or during another therapy session. It includes 16 items, each on a 5-point scale, measuring agreement for behaviours in four dimensions: engagement in the therapeutic process, emotional connection with the therapist, safety within the therapeutic system and shared sense of purpose within the family. Higher scores represent greater alliance.

Scoring and missing item data

When a questionnaire was received but questionnaire item data were missing (item non-response), documented instructions for dealing with missing data were followed when these were available (EQ-5D, SDQ, McMaster FAD[77]). Otherwise, in the absence of documented instructions, the half rule[78] was used (CDRS-R, BSS, PQ-LES-Q, Hopelessness Scale, Family Questionnaire, GHQ-12, ICU, SOFTA), allowing for substitution of the mean of answered questions of specific subscales/totals provided that at least half the questions on a given scale were answered. Further details of scoring and categorisation can be found in *Appendix 3*.

Data quality and monitoring

Data were monitored for quality and completeness by the CTRU using established verification, validation and checking processes. Missing data, except individual data items collected via the postal questionnaires, were chased for until they were received or confirmed as not available or until the trial was at analysis.

Protocol adherence, trial uptake, loss to follow-up, withdrawal, participant safety, data quality and data/questionnaire return rates were monitored, as appropriate to each group, at each TMG, Trial Steering Committee (TSC) [which included a patient and public involvement (PPI) member] and Data Monitoring and Ethics Committee (DMEC) meeting.

The DMEC reviewed, on an annual basis and by trial arm, the number and frequency of hospitalisations and deaths as a consequence of self-harm. Processes were in place to escalate any concerns raised by the DMEC or to escalate concerns identified by the trial team to the DMEC in the interim periods between meetings. The DMEC reviewed interim primary outcome data after half the events had occurred.

Quality assurance

The trial was conducted in accordance with current Medical Research Council good clinical practice guidelines, the NHS Research Governance Framework and through adherence to CTRU standard operating procedures.

The trial was reviewed and approved by the National Research Ethics Service Committee (NRES) Yorkshire and the Humber – Leeds West [Research Ethics Committee (REC) reference 09/H1307/20]. SHIFT also received research and development approval from all participating CAMHS' host organisations. For the purposes of collecting routine data from acute trusts to inform the primary outcome, the REC agreed that the study was exempt from site-specific assessment, but researchers required letters of access from each trust to access records.

Protocol amendments

There were 11 substantial amendments to the protocol (see *Appendix 2* for full details) throughout the lifetime of the SHIFT trial. The methods detailed here cover all protocol amendments.

Three substantial amendments were made prior to the start of recruitment. These included clarification of the eligibility criteria and of the secondary outcome measures to be used.

After recruitment commenced, eight further substantial amendments were made, some of which covered more than one issue.

In brief, these included an amended process for obtaining consent for FT sessions to be video-recorded (and confirmation of data transfer processes), clarification of eligibility criteria, randomisation of family therapists (prior to the relevant teams commencing recruitment), augmenting follow-up processes (additional researcher contact prior to postal follow-up, thank-you letters at 12 or 18 months where questionnaires were returned by post), inclusion of verbal consent to researcher contact where written consent was not possible, clarification of who is appropriate to take on the consenting caregiver role, production of a FT leaflet for those allocated FT, sending Christmas cards to participants, collection of multiple contacts at baseline to aid follow-up, introduction of participant incentives and participant information amendments to confirm duration of follow-up.

Patient and public involvement contribution

A local young persons' group in Leeds was consulted regarding the content of the participant information sheets.

A PPI expert was a member of the TSC from the first meeting in 2010 to October 2012.

Unsuccessful attempts were made to invite a young person and her primary caregiver to join the TSC.

Over the course of the trial there was further consultation with young people who work with the charity Young Minds and with the National Institute for Health Research (NIHR)'s Young People's Mental Health Advisory Group.

Statistical methods

Sample size

The power calculation was based on a minimally important reduction in 18-month repetition rates of self-harm (leading to hospital attendance) from 29% in participants receiving TAU[6] to 18.8% in participants receiving FT, that is, a reduction of 35%, providing a constant hazard ratio (HR) of 1.64. Using a 5% significance level log-rank test for equality of survival curves, 374 participants per arm were required, with 172 total events, to give 90%. Assuming, at most, 10% loss to follow-up by 18 months, the total sample size required was 416 per arm, 832 in total.

Although SHIFT was individually randomised, the inherent clustering of participants nested within therapists is known to have an impact on power and is related to the level of the intracluster correlation coefficient (ICC) and the cluster size. It was expected that the level of clustering would be low, possibly around 0.01 but no higher than 0.05 (owing to the use of therapy manuals, therapist selection, training, supervision and monitoring). It was also expected that the number of participants per therapist would be small. It was estimated that between 8 and 15 therapists would be available in the TAU team at any one centre, so across the anticipated 15 participating trusts there would be 120–225 therapists available to treat 416 participants; thus, each therapist would treat between two and four participants (maximum cluster size of four). In FT, we estimated that there would be approximately 35 therapists available across all sites to treat 416 participants; thus, each lead therapist would have direct contact with approximately 12 participants (maximum cluster size).

The design effect, describing the extent to which the sample size must be increased to obtain the same power as an individually randomised study without clustering, was therefore assumed likely to be no greater than 1.55 (ICC of 0.05), effectively reducing the sample size from 416 per group to 270 per group and the power from 90% to around 75%. If the ICC were as low as 0.01, then the design effect would be 1.11, reducing the sample size to 374 per group and the power to around 85%. We anticipated that the ICC would be towards the lower end of the possible range and therefore the trial would still be adequately powered with the sample size planned.

Participant populations

The intention-to-treat (ITT) population consisted of all randomised young people who had consented, regardless of non-compliance with the intervention, whether they were found to be ineligible post randomisation and/or remained in the trial and were analysed according to the treatment they were randomised to receive.

The per-protocol population consisted of all randomised, consented young people, excluding those who were found not to satisfy the eligibility criteria.

Statistical analysis

A single formal interim analysis was planned for the primary end point, repetition of self-harm leading to hospital attendance within 18 months of randomisation, when at least half the required number of events were reached (86 events). It was planned that the DMEC, in light of interim data, would if necessary report to the TSC with a recommendation of trial adaptation or early closure if, compared with TAU, the effect of FT was significantly inferior ($p < 0.005$). Final analyses were planned after all participants had reached the end of the 18-month follow-up period, when it was expected that at least 172 events would have occurred.

Analysis methods

All data analyses and summaries were performed using SAS version 9.4 (SAS institute, Cary, NC, USA), and, unless otherwise specified, were conducted on the ITT population. An overall two-sided 5% significance level was used for all end-point comparisons; however, for the primary end point, the O'Brien and Fleming[79] alpha spending function was used to account for the interim analysis, allowing an alpha level of 0.047 for the final analysis and 0.005 for the interim analysis.

General calculations

Time from randomisation was calculated in days with estimates presented in months, where 1 month was defined as time in days/30.44, and 18 months calculated as 528 days.

All clinical data included data up to 18 months post randomisation. Attendances and other clinical data, that is, referrals, occurring post 18 months were excluded from the analysis as these data were not collected reliably for all participants.

Covariates

All multivariable analyses were adjusted by the randomisation stratification factors age, gender, number of previous self-harm episodes, type of self-harm (if there were errors during randomisation the true value was used) and by participants' recruiting trust and source of referral (through hospital or not). The stratification factor 'living arrangements' was not included as a covariate as all but two young people were living with parents/guardians. Adjustment was made for centre grouped at the trust level as, with 40 centres recruiting to SHIFT and between two and 58 participants recruited from each centre, an adjusted analysis by centre could prove problematic because of instability and lack of convergence (i.e. there could be some small centres where no events were observed).

Study summary, baseline characteristics, treatment and safety

Summaries of screening, accrual, protocol violators, withdrawals, participants' follow-up rates and unblinding were produced and an overall CONSORT diagram presented. All baseline characteristics of participants, treatment received and safety data were summarised by treatment group and described descriptively without statistical comparison.

Primary end-point analysis

Derivation

The timing of repeat self-harm was calculated from the date of randomisation to the date of hospital attendance corresponding to the first instance of repeat self-harm within 18 months of randomisation. Young people without any reported instances of repeat self-harm leading to hospital attendance were censored at their final date of hospital follow-up or 18 months post randomisation, whichever was earlier. Young people without follow-up (i.e. withdrawn prior to a hospital follow-up search) were censored at baseline.

Timing of young people's final hospital follow-up was based on the most recent data cut-off point for HES data collected via NHS Digital and researchers' Acute Trust searches. Follow-up dates based on researchers' Acute Trust searches were calculated for each young person based on the hospital(s) most frequented for participants recruited from the same CAMHS service. For the majority of young people, full 18-month follow-up was achieved via successful linkage to HES, with HES data covering participants' full follow-up period; or, where HES linkage was not possible, via full researcher follow-up at the Acute Trusts to which participants would most likely present. However, full 18-month hospital follow-up was not achieved for participants for whom the final HES data download did not cover their full follow-up period (owing to HES data being provided 3 months in arrears); HES linkage was not possible and researcher follow-up did not cover participants' full follow-up period; or where consent was withdrawn for further clinical data collection and hospital follow-up was based on the HES data cut-off or researcher follow-up preceding the withdrawal.

Attendances at minor injury units (MIUs) and walk-in centres (WIC) identified via HES, without subsequent A&E or inpatient attendance, did not fit the definition of the primary outcome and were excluded.

Attendances reported via both HES and the researcher were combined, in order to avoid double counting. Where there was conflicting information, the more serious case was used, that is, admissions to hospital were deemed more serious than attendance at A&E only, self-harm more serious than non-self-harm and, if uncertain, HES data were used to ensure consistency.

Full details of data cleaning, derivation and use of HES data to identify attendances resulting from self-harm will be published separately.

Missing data

Missing presenting (resulting from self-harm or not) details for hospital attendances reported in HES were queried by trial researchers at relevant acute trusts. At the time of analysis, attendances with missing details were reviewed by the chief investigator (blind to treatment allocation). Information was provided on the length of hospital stay, diagnoses where available and treatment received and as a result it was possible to further categorise 73 of all 915 hospital attendances (8%; A&E, MIU or WIC) as not being self-harm related. Remaining unclassified attendances were assumed not to be related to self-harm in the primary analysis.

Analysis

A multivariable Cox proportional hazards model, accounting for covariates, was used to test for differences in 18-month repetition rates between treatment arms. The assumption of proportional hazards was assessed by plotting the hazards over time (i.e. the log-cumulative hazard plot) and the methods of Lin *et al.*[80] were further used to check the adequacy of the model.

Kaplan–Meier curves were constructed for each group. Repetition estimates at each month post randomisation with corresponding 95% confidence intervals (CIs) are also presented for each treatment arm and for the difference between arms.

Although the significance level was reduced to account for an interim analysis, CIs are still presented at the 95% level.

Sensitivity analysis

The extent of clustering caused by therapists and the impact on the precision of the treatment effect estimate was investigated using a multivariable multilevel survival frailty model,[81,82] in which a common frailty for individuals treated by the same therapist allows for heterogeneity between groups of participants treated by different therapists and accounts for within-group correlations (using the RANDOM statement in PROC PHREG).

Clustering by therapist was derived according to the 'main' therapist delivering the greatest number of sessions for FT or TAU. Where an equal number of sessions were delivered by different therapists, the main therapist was defined as the earliest of the therapists involved. Participants who received no treatment, were missing treatment data or missing the name of their main therapist were classed as belonging to their own individual cluster of size 1.

Missing presenting details A sensitivity analysis was also conducted to assess the impact of missing details of hospital attendance (i.e. where it is unclear whether or not an attendance had resulted from self-harm), in which unclassified attendances were classed as self-harm related, thus contributing to the primary outcome.

Associations with trust Given that the proportional hazard assumption was questionable for a number of trusts, further multivariable Cox proportional hazards regression models were fitted to investigate the effect on the treatment estimate without trust in the model and with hub, and associations between trust and other covariates were explored. The primary analysis model was also fitted accounting for the shared frailty by trust, therefore treating trust as a random effect rather than a fixed effect.

Secondary end-point analysis

Repetition rates of self-harm leading to hospital attendance 12 months after randomisation

The analysis of 12-month repetition rates followed that of the 18-month data detailed in the primary end-point analysis, with events and follow-up curtailed at 12 months rather than 18 months.

Characteristics of further episodes of self-harm leading to hospital attendance

Further episodes of self-harm were analysed using a multivariable recurrent event analysis, incorporating the timing and cumulative number of self-harm events. We used a counting process model[83] with robust sandwich variance estimator for standard errors (SEs) of coefficients[84] to take account of the within-subject correlation (also known as the proportional means model or the independent increment model). The counting process model is an extension of the Cox regression model and, when unrestricted (as per the Andersen and Gill[83] model), regards all subjects to be at risk of an event regardless of the number of events experienced thus far. Sensitivity analysis further analysed recurrent events using the conditional, restricted gap time model of Prentice *et al.*,[85] which assumes that the occurrence of the first event increases the likelihood of further recurrence and participants are considered to be at risk of an event only if the previous event occurred.

Characteristics of all self-harm events (SASII)

No formal statistical analysis was undertaken; characteristics are summarised descriptively only.

Derivation Characteristics of the first self-reported self-harm episode (regardless of hospital attendance) were summarised for participants with sufficient SASII completion from baseline to 12 or 18 months, that is, completion at the 12-month researcher visit at least or at both the 12- and 18-month visits. Furthermore, participants with an 18-month researcher visit only (not a 12-month visit) with self-harm reported between 12 and 18 months post randomisation were included in the overall number of participants with self-harm reported, as these instances of self-harm during follow-up were known.

The timing of first self-reported self-harm post baseline was calculated from the date of randomisation to the first reported date of self-harm within 18 months of randomisation. Participants without any self-reported self-harm and with sufficient SASII completion were censored at their date of follow-up, that is, the date of either the 12-month researcher visit or the 18-month visit if both 12- and 18-month visits were conducted. Participants missing the 12-month SASII were censored at baseline, including those where self-harm was known to have occurred between 12 and 18 months as the timing of the first self-harm post randomisation could not be determined.

Owing to the large variability in the level of self-reported self-harm over the SASII timeline, responses were summarised according to the occurrence and frequency of self-harm for each 3-month period post randomisation (0–3 months, 3–6 months, etc.) with frequency categorised as none (0 episodes), less than once per month (fewer than three episodes), less than once per fortnight (3–6 episodes), less than once per week (7–12 episodes), once or more per week (13–25 episodes), twice or more per week (26–45 episodes), most days (46–91 episodes).

Patient-reported outcomes

For the secondary end points (Beck Scale, CDRS-R, PQ-LES-Q, SDQ, GHQ-12, Hopelessness Scale, McMaster FAD and Family Questionnaire) repeated-measures models were used to estimate differences between the treatment groups over time. Linear repeated-measures models for continuous scores were used for all end points with the exception of the Beck Scale, for which the proportion of participants identified as having suicide ideation was modelled using a logistic repeated-measures model owing to a

zero-inflated total score. A covariance pattern model was used, which accounts for repeated measures on the same participant, which may be correlated, with a first-order heterogeneous autoregressive covariance matrix and allowing for randomised treatment, time effects, baseline score, covariates and treatment–time interactions as fixed effects. This model was chosen over a random-effects model, as there were just two follow-up time points at 12 and 18 months (3 and 6 months for the Family Questionnaire), with only a small proportion (< 3%) of questionnaires completed outside the 2-month acceptable time window. Model assumptions were checked using Pearson and Studentised residuals.

Responder characteristics were summarised by trial arm to explore whether or not differential patterns existed that could impact on the conclusions for the secondary outcomes; characteristics explored included self-harm during follow-up, covariates, age, questionnaire total scores and questionnaire subscales, for subscales for which significant ($p < 0.05$) treatment effects were detected at 12 or 18 months. Logistic regression modelling was used to assess participants' response status against participant characteristics and treatment.

Missing data To allow the ITT population to be used, missing questionnaire data were assumed to be missing at random and missing scores were estimated using multiple imputations for all participants. Multiple imputation uses the distribution of the observed data to estimate a set of plausible values for the missing data. The multivariate normal model via the MCMC (Markov Chain Monte Carlo) method was used to impute missing values for each of the questionnaire scores, allowing for the imputation of normally distributed data with a non-monotone missing pattern.[86] Missing values for each score were imputed separately and the following predictor variables were incorporated in the model: covariates, randomised treatment and participants' scores at all available assessments (i.e. baseline, 6 and 12 months). The imputation process was repeated according to the percentage of missing data. If 20% were missing, then 20 imputations were made, with a minimum of 10 imputations. Results reported were calculated using Rubin's rules[87] for combining the results of identical analyses performed on each of the imputed data sets.

A sensitivity analysis based on an alternative modelling strategy, also assuming that data are missing at random, using data observed for at least one follow-up time point, was also conducted. In this case, as baseline score was a covariate in the model, participants' missing baseline scores were imputed by the mean across observed baseline values.[88]

Predictive and process measures

Moderator analysis

Moderator analysis explored whether or not the treatment effect in the primary analysis Cox proportional hazards model depended upon participants' baseline characteristics via inclusion of the proposed moderator alongside the interaction effect of treatment × moderator in the primary analysis model (including covariates). Covariates and responses to all baseline questionnaires were explored for moderation. Questionnaire responses were also categorised for the young person Beck Scale to indicate whether or not suicidal ideation was present; the young person CDRS-R for whether or not depression was present; and the young person and caregiver McMaster FAD subscales to indicate whether family functioning was healthy or unhealthy. A 5% significance level was used to identify moderation through the interaction of the potential moderator with treatment, irrespective of the main effect of the potential moderator, although it is acknowledged that interaction testing is not a powerful test for moderation. Analysis was of the ITT population to availability of data (complete case) for each proposed moderator.

Mediator analysis

Complier average causal effect analysis A complier average causal effect (CACE) analysis was conducted to model the causal effect of FT receipt (as opposed to randomisation) on the primary outcome (*Figure 1*). An instrumental variable model was used to adjust for potential selection bias occurring because of participants who did not receive their allocated treatment, while addressing potential confounds. Randomisation allocation was used as the instrumental variable, creating additional variance unaccounted for by FT receipt alone to obtain an unbiased estimate of the effect of FT receipt, in a two-stage instrumental variable analysis using the qualitative and limited dependent model (QLIM) procedure in SAS,[89] to simultaneously estimate the first equation of treatment selection (stage 1) and the second equation of the unbiased effect of treatment on outcome (stage 2). Covariates were included in each stage of the model. The primary outcome was considered as a binary variable and a probit transformation was used. The primary ITT analysis was repeated on this transformed scale for comparison with the CACE estimates, alongside a further 'as treated' analysis (using the QLIM procedure). TAU participants with missing treatment data were assumed to have not received FT in the analysis. Kaplan–Meier curves were also constructed for each group by receipt of FT and receipt of FT overall.

Mediation Process variables identified as potential mediators included the overall number of sessions attended, the use of psychotropic medications and therapist characteristics.

Responses to the 3- and 6-month Family Questionnaire (subscores) and 12-month questionnaires (total scores and subscales found to have a significant treatment effect at 12 months in the secondary end-point analysis or moderator analysis) were also explored for mediation. For questionnaire responses, mediation was explored in relation to the time to event, self-harm resulting in hospital attendance, after measurement of the potential mediator. Therefore, events after 3, 6 and 12 months post randomisation were included for 3-, 6- and 12-month questionnaires, respectively. Responses at 18 months were not considered in the analysis, as outcomes were collected up to 18 months only and responses could not mediate an outcome observed before this date.

The Baron and Kenny[90] steps were employed to explore mediation (*Figure 2*):

- **Step 1** – establish an effect of randomisation on outcome that may be mediated, that is, Cox regression with Y as the dependent variable and X as predictor, estimate and test path c the total effect:

$$Y = h_0(t)\exp(\beta_{10}X). \quad (1)$$

- **Step 2** – establish that there is an effect of randomisation on the mediator, that is, linear or logistic regression with M as the dependent variable and X as predictor, estimate and test path a:

$$Me = \beta_{20} + \beta_{21}X. \quad (2)$$

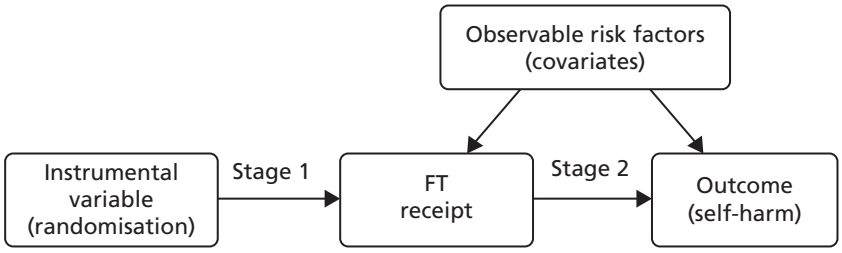


FIGURE 1 Schematic representation of CACE analysis.

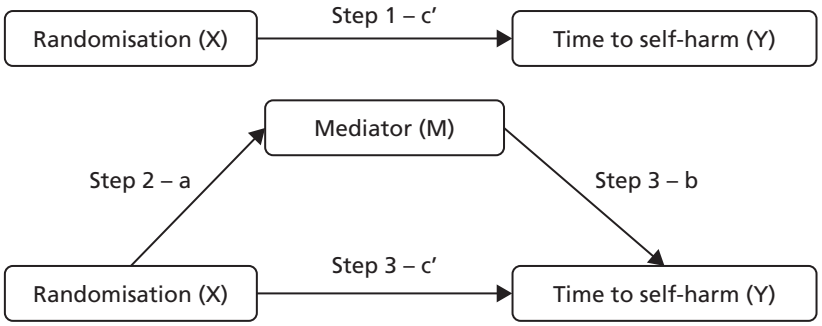


FIGURE 2 Baron and Kenny mediation model.

- **Step 3** – establish that there is an effect of mediator on outcome while controlling for randomisation, that is, Cox regression with *Y* as the dependent variable and *X* and *M* as predictors, estimate and test paths b and c':

$$Y = h_0(t)\exp(\beta_{31}X + \beta_{32}Me). \quad (3)$$

Following these steps, complete mediation is the case in which variable *X* (randomisation) no longer affects *Y* (time to self-harm) after *M* (mediator) has been controlled and so path c' is zero. Partial mediation is the case in which the path from *X* (randomisation) to *Y* (time to self-harm) is reduced in absolute size but is still different from zero when the mediator is introduced.

Steps 1 and 3 were fitted using a Cox proportional hazards model adjusted for all covariates. For continuous mediators, Step 2 was fitted using a linear regression model containing randomised treatment and all covariates; and for binary outcomes (Beck suicide ideation, psychotropic medications, therapist characteristics) logistic regression was used containing randomised treatment and all covariates apart from NHS trust (owing to lack of convergence). For mediators based on questionnaire responses, the baseline score was also included as a covariate in each step.

Adherence and alliance to family therapy

Summaries of therapist baseline characteristics, training, therapy session recordings, initial and specialist adherence review, formulation letters to families and therapist supervision were produced. In addition, mean scores and 95% CIs are presented for the total score and for each subscale of the SOFTA.

Economic methods

The economic evaluation aimed to assess the cost-effectiveness of FT compared with TAU in the management of self-harm in adolescents from a health and social care costs perspective. It consisted of a within-trial analysis, in which cost-effectiveness was assessed within the 18-month trial period and a decision-analytic model in which cost-effectiveness was assessed by extrapolating the trial results to a longer time horizon.

Outcomes

Young people's health-related quality of life was assessed using the EQ-5D,[91,92] which was included at baseline and along with the young person's resource use questionnaire at 6, 12 and 18 months. Changes in EQ-5D scores across time points were evaluated using two-sample *t*-tests to explore any important differences in these end points within the time frame of the trial. These tests are useful in understanding the impact of the length to follow-up on the health outcome measures, as insufficiently long follow-up periods may introduce biases in the subsequent cost-effectiveness analysis. In line with the NICE reference case[93] the primary outcome for the economic evaluation was quality-adjusted life-years (QALYs). Young

people's responses to the EQ-5D were converted into health state utility values and multiplied by the proportion of 1 year the time period represented (baseline to 6 months = 0.5; 6 to 12 months = 0.5; 12 to 18 months = 0.5) to calculate QALYs.[94] Average QALYs between adjacent time points were calculated to generate smoothed estimates between time points. Therefore, QALYs were calculated using the area under the curve approach as shown below:

$$\begin{aligned} QALY_{\text{young person}} = {} & \{[(EQ5D\_Baseline + EQ5D\_6month)/2] \times 0.5\} + \{[(EQ5D\_6month \\ & + EQ5D\_12month/2] \times 0.5\} + \{[(EQ5D\_12month + EQ5D\_18month/2] \times 0.5\}. \end{aligned} \quad (4)$$

Quality-adjusted life-years represent a quality-of-life-weighted survival value in which 1 QALY is equivalent to 1 year of full health.

The number of self-harm events avoided at 18 months was considered as a secondary outcome. Additionally, the sensitivity analysis used the aggregated QALY gain for both the young person and his/her main carer as the health outcome. The caregiver's health-related quality of life was assessed using the HUI,[95,96] which was included along with the carer resource use questionnaire. The carer's responses to the HUI were converted into health state utility values[76] and the carer's QALYs were calculated in a similar way to the young person's QALYs.

$$\begin{aligned} QALY_{\text{caregiver}} = {} & \{[(HUI\_Baseline + HUI\_6month)/2] \times 0.5\} + \{[(HUI\_6month + HUI\_12month)/2] \times 0.5\} \\ & + \{[(HUI\_12month + HUI\_18month)/2] \times 0.5\}. \end{aligned} \quad (5)$$

Resource use

Resource use data were collected at baseline and at 3, 6, 12 and 18 months from the young person and from his/her main carer. Resource usage was converted into costs using unit cost figures from the *British National Formulary* (BNF), Personal Social Services Research Unit (PSSRU) and the Department of Health's National Schedule of Reference Costs.[97–99] The base-case economic evaluation focused on health and social care costs. The currency used was the pound sterling (£) and 2014 was used as the reference financial year.

Costs and benefits were discounted at 3.5% per annum.

Missing data

Missing data were imputed using multiple imputations via chained equations[86,100,101] as recommended for economic analyses alongside clinical trials.[102] For consistency with the statistical analysis and to ensure best fit, imputations were based on predictor variables including treatment allocation, gender, age, centre, total number of self-harm episodes, type of index episode, source of referral from hospital and derived scores for a number of assessment instruments (BSS, CDRS-R, Hopelessness Scale for Children and PQ-LES-Q).

Within-trial cost-effectiveness

The within-trial analysis aimed to determine the intervention that maximised health outcomes at 18 months. It used an ITT perspective.

Base-case analysis 1 was a cost–utility analysis over 18 months examining the cost per QALY gained for all participants. Descriptive statistics of costs and EQ-5D scores and parametric tests to evaluate any important differences between the two trial arms at each time point in the trial were produced. Thereafter, incremental cost-effectiveness ratios (ICERs) were calculated by dividing the average difference in costs between the two arms by the average difference in QALYs between the two trial arms:

$$ICER = \frac{(MeanCost_{TR} - MeanCost_{HR})}{(MeanQALY_{TR} - MeanQALY_{HR})} \quad (6)$$

The ICER represents the additional cost per unit of outcome gained. This indicates the trade-off between total cost and effectiveness when choosing between FT and TAU. When compared against the cost-effectiveness threshold, this gives an indication of whether or not spending additional money on FT appears efficient. As a rule, the NICE implicit cost per QALY threshold of £20,000–30,000 per QALY was used. An intervention is cost-effective so long as its ICER is within or below the £20,000–30,000 per QALY range.

Base-case analysis 2 calculated the incremental cost per self-harm event avoided because of FT at 18 months compared with TAU and was carried out in a similar way; however, there is not a clear decision rule on the cost-effectiveness threshold per self-harm event avoided and so the results were presented for a range of cost-effectiveness thresholds.

Cost-effectiveness over a longer time horizon

The NICE reference case recommends exploring cost-effectiveness over a time horizon that is long enough to reflect any important differences in costs or consequences between treatments being compared. The NHS might be interested to understand the cost and the consequences of FT and TAU beyond the trial follow-up, and decision-analysis modelling is key in this context. As far as we know, there are no prior studies in which a decision-analysis model was built to study the long-term cost-effectiveness of FT for adolescents. Considering that a cohort of young people will face such a large number of events over their adult life, we considered that building a decision-analysis model that reflects a lifetime horizon would require a larger amount of assumptions and probabilities to estimate. Therefore the decision-analysis model evaluated the cost-effectiveness of FT compared with TAU up to 5 years after randomisation. It considered health and social care costs and used young people's QALY gain as the outcome. The results from the within-trial analysis were mainly used to make reliable assumptions for the model as appropriate literature references were not identified from which health states or probabilities beyond an 18-month follow-up could be extracted.

A Markov model with three health states was used (*Figure 3*). The model included three possible health states: self-harm, defined as self-harming at least once in a period of 6 months; no self-harming; and death. Markov models describe participant progression over time through a pathway of health states, with movement between the health states being triggered by events; in this case, self-harm events or death. Resource use and costs were associated with each health state and participants accumulated costs and health benefits in each state over 6-month cycles. Participant cost and utility data were available at 6, 12 and 18 months from the trial data and were directly included into the model to estimate longer-term costs and health benefits.

The model inputs were derived from the trial data. Specifically, the proportion of the participants beginning in each health state in the model was derived directly from the proportion of participants in the trial who remained in the self-harm state or moved to no self-harming. In the first cycle of the model, all participants

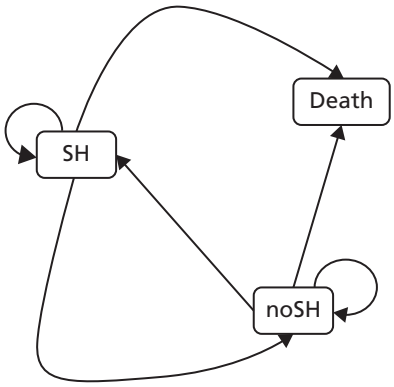


FIGURE 3 Three-state Markov model. SH, self-harm; noSH, no self-harm. Reproduced from Cottrell *et al.*[70]

started from self-harm in both arms and this was informed by the inclusion criteria for the trial. No participants died over the 18-month trial follow-up, but to account for possible death it was assumed that the probability of a participant moving from self-harm to death or from no self-harm to death was minimal and equal to 0.0001, regardless of the trial arm. Derivation of the post-18-month transition rates between the different states required extrapolation beyond the follow-up period of the trial.

Regarding costs and utility related to each model state, death was assumed to be associated with zero utility and zero cost. For self-harm and no self-harming states, the associated costs and utility values at each follow-up time point (6 months, 12 months and 18 months) were derived directly from the trial data.

Any intervention costs were assumed to occur equally over the first 12 months for each arm based on the trial data. Given that it was not possible to distinguish if these costs were incurred only by those in the self-harm state, in the no self-harming state or both, it was assumed to be the same for any of the states.

Presentation of results

The outputs of both the within-trial analysis and the decision-analysis model were presented as expected ICERs of FT compared with TAU, scatterplots on the cost-effectiveness plane and cost-effectiveness acceptability curves using non-parametric bootstrapping to determine the level of sampling uncertainty.

In order to test the robustness of the results of the within-trial analysis, a number of sensitivity analyses were conducted to explore the impact of costs and missing data on the results of the study as described below:

1. a non-parametric bootstrapping with 10,000 replications
2. an analysis in which it was assumed that only one therapist was involved in each treatment session in the FT arm (intervention costs from scenario 1)
3. an analysis in which it was assumed that the average number of therapists were involved in each treatment session in the FT arm (intervention costs from scenario 2)
4. an adjusted analysis to account for EQ-5D differences between arms at baseline, using a set of baseline characteristics such as treatment allocation, gender and age group (baseline EQ-5D was also included as a control to account for any impossible imbalance and to improve the precision of the estimates)
5. a complete-case analysis including only those participants with no missing quality-of-life and cost data at any time point before imputation
6. an analysis using an aggregate QALY, taking into consideration both the young people's and caregivers' QALY gains.

In order to test the robustness of the results of the decision model, deterministic one-way sensitivity analyses were conducted to assess the impact of individual parameter uncertainty and key model assumptions on the results. Probabilistic sensitivity analyses were conducted to assess the impact of parameter uncertainty on the results. Probabilistic analysis accounts for joint parameters uncertainty in non-linear models by assigning probability distributions to each of the input parameters and randomly drawing from these probabilities over a number of Monte Carlo model simulations to produce different cost and QALY estimates in each simulation of the model. In our model, we chose to do 30,000 simulations owing to the results being highly sensitive to a smaller number of simulations.

Chapter 3 Main trial results

Recruitment, retention and participant characteristics

Participant flow

The numbers of young people screened for entry into SHIFT, eligible, randomised, followed up via post at 3 and 6 months, visited by a researcher at 12 and 18 months, withdrawn and analysed are presented in the CONSORT diagram in *Figure 4*.

Screening and recruitment

A total of 3554 young people were screened within CAMHS, of whom the clinician deemed 1603 (45.1%) to be eligible for the trial (see *Figure 4*). Of the remaining young people, 1831 (51.5%) were deemed to be ineligible and no eligibility information was provided for a further 120 (3.4%). The most common reasons for ineligibility were not having engaged in self-harm prior to the current CAMHS referral (44.8% of those ineligible), already being treated within CAMHS (25.2%), not intending to offer CAMHS follow-up for self-harm (21.7%), not living with their primary caregiver (17.6%), recent self-harm was not a key feature of presentation and referral into CAMHS (16.4%), presenting episode resulted from consumption of alcohol or drugs where specific intent to self-harm was not established (13.7%), in a children's home or short-term foster placement (10%), would not ordinarily be treated in CAMHS (9.2%), currently undergoing a child protection investigation (7.8%) and at serious risk of suicide (7%).

A total of 993 (61.9%) eligible young people consented to researcher contact, of whom 481 (48.4%) provided verbal consent and 512 (51.6%) provided written consent; 273 (17%) young people did not consent, and details of researcher consent were missing for 337 (21%) eligible young people. Of the 273 young people not providing consent, 160 (58.6%) refused consent, 85 (31.1%) did not attend their first CAMHS follow-up appointment and 18 (6.6%) were not approached by the CAMHS clinician.

Researcher contact and baseline visits were sought for the 993 young people consenting to researcher contact. A total of 832 (83.8%) consented to trial participation and were randomised into the trial, 51.9% of those eligible. A further 127 (12.8%) young people did not provide consent for trial participation (100 refused; 11 could not be contacted in sufficient time).

Screening and recruitment took place over 4 years: the first participant was randomised on 14 April 2010 and the last on 30 December 2013 once the target of 832 participants had been reached, with 415 participants randomised to receive FT and 417 participants randomised to receive TAU. *Figure 5* shows overall, monthly and cumulative recruitment of participants into the trial, and *Figure 6* presents overall recruitment by CAMHS and hub and treatment arm while *Table 102* in *Appendix 4* provides a summary of overall recruitment activity by CAMHS centre, trust, hub and treatment arm and by CAMHS.

Characteristics of the screened, eligible and randomised participants – generalisability

Tables 2 and *3* compare the characteristics of the screened, eligible and randomised young people and show similar characteristics according to gender, ethnicity and treatment received. Between 82% and 89% were female within each population; and between 79% and 84% were of white ethnicity. The mean age was 14.8 [standard deviation (SD) 1.55], 14.6 (SD 1.38) and 14.3 (1.37) in each respective population with an increasing proportion of younger adolescents (aged 11–14 years) in later stages of the screening process.

There were more marked differences according to referral source; almost 60% of those screened were referred via hospital while 25% had been referred via their general practitioner (GP). However, of those randomised, only 37% had been referred via hospital while 46% had been referred via their GP; this is discussed in relation to the primary outcome in *Primary analysis: Cox proportional hazards regression*.

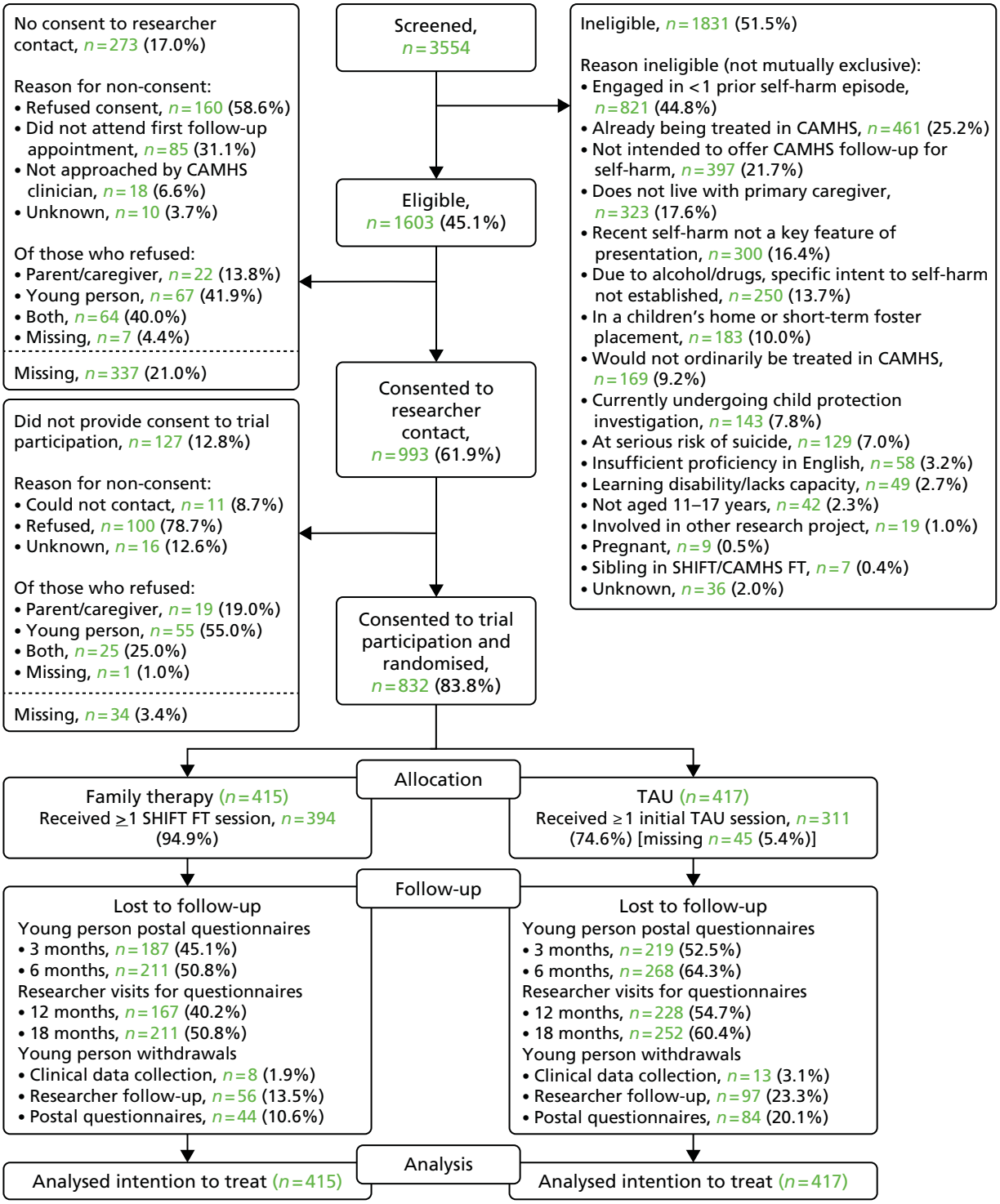


FIGURE 4 Consolidated Standards of Reporting Trials flow diagram. Adapted from Cottrell *et al.*[70] © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Similarly, the proportion of participants whose most recent index self-harm episode was poisoning decreased from 41% of screened to 23% of randomised while the proportion cutting increased from 40% to 63%.

Young people referred to CAMHS through hospital therefore had a far lower eligibility rate than those from the community (31.8% vs. 70.2%) and a lower rate of consent for researcher contact (54.6% vs. 67.8%), suggesting that hospital referrals were a more highly selected group.

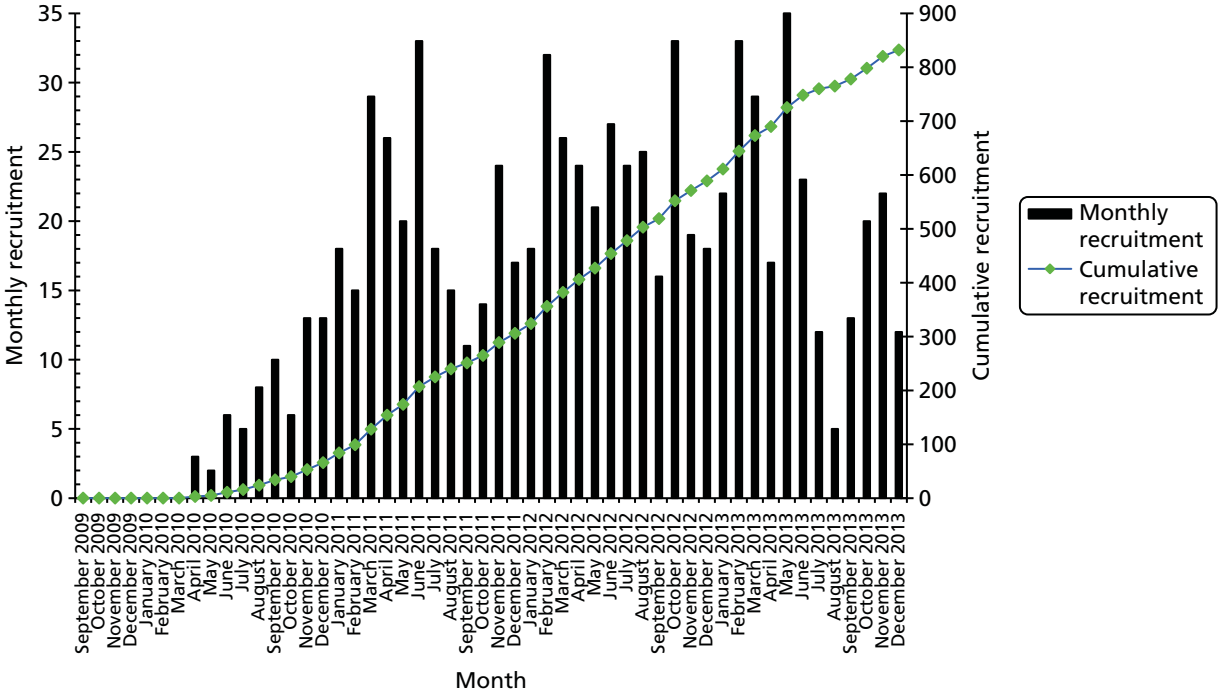


FIGURE 5 Recruitment graph.

Compared with community referrals, young people referred through hospital were more likely to be ineligible because their index episode was their first reported episode of self-harm (48.8% vs. 29.9%) or because they were already being actively treated in CAMH (28.3% vs. 14.5%). The proportion of young people who did not consent to researcher contact because they did not attend their CAMHS follow-up appointment was higher among those referred from hospital than among those referred through the community (37.9% vs. 23.8%), while the proportion of non-consent caused by refusal was far lower (47.9% in hospital referrals vs. 70% in community withdrawals).

The characteristics of young people referred from hospital differed from those of community referrals in that they were slightly older [mean age of 14.9 (SD 1.49) years vs. 14.5 (SD 1.57) years], their index episode of self-harm was more likely to be caused by self-poisoning (63.5% vs. 6.2%), and they were more likely to have received at least some treatment following the episode (31.3% required no treatment, compared with 66.7% of community referrals).

Baseline characteristics

Clinical characteristics

Baseline characteristics and clinical details of the 832 randomised participants are presented by arm in *Tables 4–10*. The two groups are largely well balanced with respect to these characteristics.

The mean age at randomisation was 14.3 (SD 1.38) years in the FT group and 14.4 (SD 1.35) years in the TAU group. There were more females than males in both arms: 368 (88.7%) in the FT arm and 369 (88.5%) in the TAU arm (*Table 4*). Most young people had self-harmed on at least three previous occasions [369 (88.9%) in the FT arm and 370 (88.7%) in the TAU arm] as opposed to on only two occasions, and the type of most recent episode was largely self-injury [297 (71.6%) in the FT arm and 297 (71.2%) in the TAU arm], with a further 93 (22.4%) in the FT arm and 91 (21.8%) in the TAU arm due to self-poisoning; the remainder used combined methods. Almost two-thirds (63.5%) of participants were referred to CAMHS through the community rather than from hospital. Almost all participants were living with their parents/guardians, with only two in foster care, and the majority were in full-time education [398 (95.9%) in the FT arm and 386 in the TAU arm (92.6%)]. Ethnicity was also well balanced across the arms. The overall proportions are presented in *Table 2*.

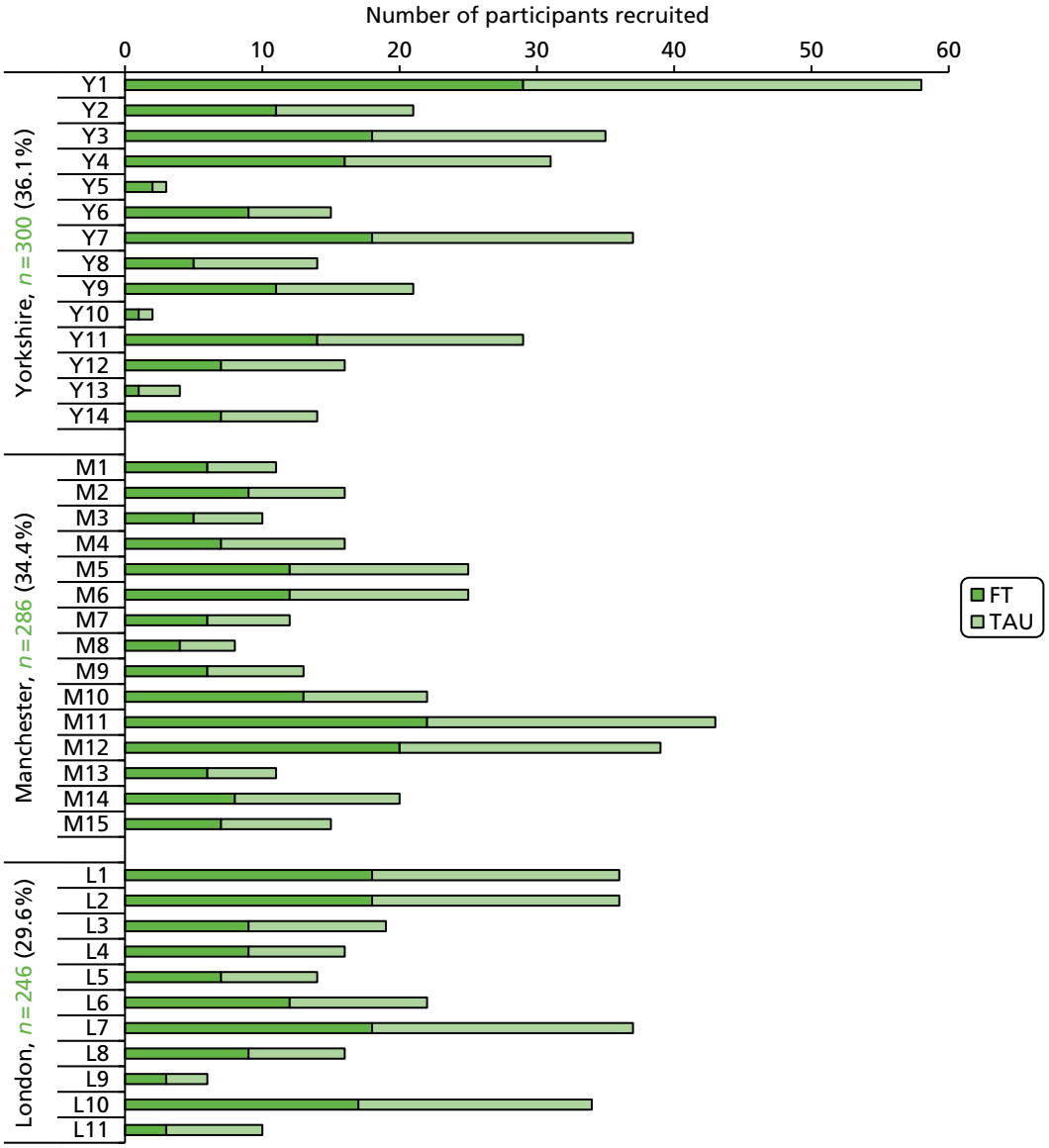


FIGURE 6 Recruitment by CAMHS and hub.

Over one-quarter of participants were reported to have a health or disability problem (218, 26.2%) and a similar proportion had past involvement with CAMHS (244, 29.3%). The most common health or disability problem (*Table 5*) was asthma (reported by 91 participants), followed by musculoskeletal problems (reported by 35 participants). Slightly more participants in the FT arm than in the TAU arm reported past involvement with CAMHS [136 (32.8%) compared with 108 (25.9%), respectively] (*Table 6*). Reasons for prior involvement were mainly self-harm related (128 participants), followed by emotional and behavioural problems, reported by 31 and 24 participants, respectively.

Just under 5% of participants were taking a prescribed psychotropic medication at the time of randomisation, with antidepressants being the most common (taken by 28 participants) (*Table 7*).

Parental abuse was reported by almost one-quarter of participants ($n = 198$, 23.8%). Marked physical abuse was reported by > 20% of participants ($n = 178$, 21.4%) and 138 (16.6%) participants reported sexual abuse (parental or otherwise) (*Table 8*). Furthermore, when asked about any other bad or scary things that had happened to them, 293 (35.2%) participants reported such events, including witnessing or experiencing a traumatic incident, relationship difficulties with family and friends, harassment or bullying, domestic violence, and other sexual, emotional or physical abuse.

TABLE 2 Characteristics of the screened, eligible and randomised populations

Participant characteristics, *n* (%)	Screened (*N* = 3554)	Eligible (*N* = 1603)	Randomised (*N* = 832)
Gender (female)	2904 (81.7)	1404 (87.6)	737 (88.6)
Age category (years)			
< 11	29 (0.8)	1 (0.1)	
11–14	1389 (39.1)	722 (45.0)	441 (53.0)
15–17	2016 (56.7)	861 (53.7)	390 (46.9)
> 17	3 (0.1)	1 (0.1)	1 (0.1)
Missing	117 (3.3)	18 (1.1)	
Ethnicity			
White	2799 (78.8)	1299 (81.0)	701 (84.3)
Black (Caribbean, African, mixed, other)	273 (7.7)	131 (8.2)	60 (7.2)
Asian (Pakistani, Indian, Bangladeshi, mixed, other)	150 (4.2)	56 (3.5)	29 (3.5)
Other ethnic group	121 (3.4)	73 (4.6)	40 (4.8)
Not stated/missing	211 (5.9)	44 (2.7)	2 (0.2)
Young person referred from			
Hospital	2129 (59.9)	678 (42.3)	304 (36.5)
GP	889 (25.0)	651 (40.6)	382 (45.9)
School	186 (5.2)	139 (8.7)	77 (9.3)
Other	213 (6.0)	111 (6.9)	58 (7.0)
Missing	137 (3.9)	24 (1.5)	11 (1.3)
Treatment required			
None	1571 (44.2)	780 (48.7)	428 (51.4)
Minimal	1204 (33.9)	578 (36.1)	301 (36.2)
Significant intervention	444 (12.5)	139 (8.7)	56 (6.7)
Intervention potentially life-saving	27 (0.8)	10 (0.6)	5 (0.6)
Missing	308 (8.7)	96 (6.0)	42 (5.0)
Hospital care			
Discharged from A&E	587 (16.5)	252 (15.7)	124 (14.9)
General admission	1122 (31.6)	368 (23.0)	168 (20.2)
ICU/HDU/other specialist unit	19 (0.5)	5 (0.3)	
N/A, community referral	1246 (35.1)	873 (54.5)	511 (61.4)
Missing	580 (16.3)	105 (6.6)	29 (3.5)

GP, general practitioner; HDU, high-dependency unit; N/A, not applicable.

TABLE 3 Type of most recent self-harm episode (not mutually exclusive)

Type of self-harm	Screened, *n* (%)	Eligible, *n* (%)	Randomised, *n* (%)
Poisoning	1452 (40.9)	444 (27.7)	189 (22.7)
Cutting	1421 (40.0)	919 (57.3)	527 (63.3)
Poisoning and cutting	218 (6.1)	123 (7.7)	54 (6.5)
Other self-injury (biting, burning, electrocution)	176 (5.0)	101 (6.3)	68 (8.2)
Other violent method (drowning, hanging/strangulation)	129 (3.6)	44 (2.7)	24 (2.9)
Threat/thoughts of self-harm or suicide only, or non-self-harm only	112 (3.2)	18 (1.1)	9 (1.1)[a]
Missing	127 (3.6)	12 (0.7)	
Total	3554 (100)	1603 (100)	832 (100)

a Note that the nine randomised participants with threats/thoughts of self-harm or suicide only were confirmed at randomisation to have self-harmed as a result of self-injury (and for one combined self-injury and poisoning) and the SASII confirmed the occurrence of prior self-harm for these nine participants.

TABLE 4 Summary of treatment allocations at randomisation for true values of stratification factors (excluding centre) and education[a]

Stratification factors, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Gender (female)	368 (88.7)	369 (88.5)	737 (88.6)
Age (years)			
11–14	220 (53.0)	221 (53.0)	441 (53.0)
15–17	195 (47.0)	195 (46.8)	390 (46.9)
18	0 (0.0)	1 (0.2)	1 (0.1)
Living arrangements			
With parents/guardians	414 (99.8)	416 (99.8)	830 (99.8)
Foster care	1 (0.2)	1 (0.2)	2 (0.2)
Number of known self-harm episodes (≥ 3 vs. 2)	369 (88.9)	370 (88.7)	739 (88.8)
Type of most recent episode			
Self-poisoning	93 (22.4)	91 (21.8)	184 (22.1)
Self-injury	297 (71.6)	297 (71.2)	594 (71.4)
Combined	25 (6.0)	29 (7.0)	54 (6.5)
Non-stratification factors			
Referred into CAMHS via hospital			
Yes	156 (37.6)	148 (35.5)	304 (36.5)
No	259 (62.4)	269 (64.5)	528 (63.5)
Young person in full-time education	398 (95.9)	386 (92.6)	784 (94.2)

a Discrepancies at randomisation: two females incorrectly randomised as male; two errors in age group (one aged 11–14 years incorrectly randomised as aged 15–17 years and the 18-year-old incorrectly randomised as aged 15–17 years); 12 errors in the number of previous self-harm episodes (six ≥ 3 incorrectly randomised under = 2 and six = 2 randomised incorrectly under ≥ 3); seven errors in the type of most recent episode (two with self-poisoning incorrectly randomised under self-injury, two self-injury incorrectly randomised under combined and three self-injury incorrectly randomised under self-poisoning).

TABLE 5 Young people's physical health

Physical health problems, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person has a health or disability problem	110 (26.5)	108 (25.9)	218 (26.2)
Type of problem (not mutually exclusive)			
Asthma	47 (42.7)	44 (40.7)	91 (41.7)
Blood problem	9 (8.2)	9 (8.3)	18 (8.3)
Diabetes	2 (1.8)	4 (3.7)	6 (2.8)
Eczema	11 (10.0)	9 (8.3)	20 (9.2)
Gastrointestinal disease	1 (0.9)	5 (4.6)	6 (2.8)
Hay fever/allergies	8 (7.3)	6 (5.6)	14 (6.4)
Headaches/migraine	4 (3.6)	7 (6.5)	11 (5.0)
Musculoskeletal problem	16 (14.5)	19 (17.6)	35 (16.1)
Neurological problem	10 (9.1)	2 (1.9)	12 (5.5)
Other	20 (18.2)	18 (16.7)	38 (17.4)
Missing	4 (3.6)	3 (2.8)	7 (3.2)

TABLE 6 Past involvement with CAMHS

Past involvement with CAMHS, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Past involvement with CAMHS	136 (32.8)	108 (25.9)	244 (29.3)
Reason for past referral			
Self-harm related	73 (53.7)	55 (50.9)	128 (52.5)
Emotional problems	18 (13.2)	13 (12.0)	31 (12.7)
Behavioural problems	14 (10.3)	10 (9.3)	24 (9.8)
Emotional and behavioural problems	2 (1.5)	3 (2.8)	5 (2.0)
ADHD/ADD	2 (1.5)	2 (1.9)	4 (1.6)
Eating disorder	3 (2.2)	1 (0.9)	4 (1.6)
Autism assessment	1 (0.7)	0 (0.0)	1 (0.4)
Behavioural problems and autism assessment	1 (0.7)	0 (0.0)	1 (0.4)
Emotional problems and ADHD/ADD	0 (0.0)	1 (0.9)	1 (0.4)
Self-harm related and autism assessment	1 (0.7)	0 (0.0)	1 (0.4)
Unknown	4 (2.9)	8 (7.4)	12 (4.9)
Other	6 (4.4)	4 (3.7)	10 (4.1)
Missing	11 (8.1)	11 (10.2)	22 (9.0)

TABLE 7 Prescribed psychotropic medications

Prescribed psychotropic medication	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person taking prescribed psychotropic medication (*Yes*), *n* (%)	17 (4.1)	24 (5.8)	41 (4.9)
Type of psychotropic (not mutually exclusive), *n* (%)			
Antidepressant drug	14 (82.4)	14 (58.3)	28 (68.3)
ADHD drug	3 (17.6)	3 (12.5)	6 (14.6)
Sedative/sleep medication	0 (0.0)	5 (20.8)	5 (12.2)
Antipsychotic drug	2 (11.8)	0 (0.0)	2 (4.9)
Anti anxiety drug	0 (0.0)	1 (4.2)	1 (2.4)
Missing	0 (0.0)	3 (12.5)	3 (7.3)
Longest length of time (months) on each medication, mean (SD), *n*	4.7 (8.09), 17	3.2 (5.46), 22	3.9 (6.68), 39

TABLE 8 Participant report of bad things that can happen

Types of bad event, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Parental abuse: when your parents get cross with you, have they ever hit you? (*Yes*)	89 (21.4)	109 (26.1)	198 (23.8)
Marked physical abuse: have you ever been hit so hard you had bruises or marks on your body or were hurt? (*Yes*)	80 (19.3)	98 (23.5)	178 (21.4)
Sexual abuse: has someone ever touched you in a way that made you feel uncomfortable (*Yes*)	75 (18.1)	63 (15.1)	138 (16.6)
Other information: anything else bad or scary (*Yes*)	147 (35.4)	146 (35.0)	293 (35.2)
Details of other bad things			
Witnessed/had traumatic incident	46 (31.3)	46 (31.5)	92 (31.4)
Relationship difficulties: family	39 (26.5)	31 (21.2)	70 (23.9)
Harassment/bullying	22 (15.0)	25 (17.1)	47 (16.0)
Domestic violence	19 (12.9)	19 (13.0)	38 (13.0)
Relationship difficulties: friends	7 (4.8)	9 (6.2)	16 (5.5)
Childhood sexual abuse	5 (3.4)	3 (2.1)	8 (2.7)
Emotional abuse	3 (2.0)	1 (0.7)	4 (1.4)
Physical abuse	0 (0.0)	3 (2.1)	3 (1.0)
Unknown	1 (0.7)	3 (2.1)	4 (1.4)
Missing	5 (3.4)	6 (4.1)	11 (3.8)

Note
Marked physical abuse is missing for a large number because of case report form and question wording. If the previous question, parental abuse, had been answered no (as for 37 of the missing here) it could have been assumed that this question was not required.
Furthermore, it is important to note that of the eight reporting other bad things as childhood sexual abuse, one had detailed a different incident for the prior sexual abuse question, one had not provided prior information and six had said no to the prior sexual abuse question, as these incidents may have been deemed to be of a different nature to those reported as sexual abuse (i.e. may not have been touched in a way that made you feel bad).

Self-reported details of the young person's most recent index episode of self-harm are presented in *Table 9*, and further details from the timeline showing all self-harm methods, treatment and suicide attempts reported over the past year are presented in *Table 103, Appendix 4*. The participant's behaviour was rated by the interviewing researcher as a suicide attempt in 313 (37.6%) cases. The remaining episodes were largely classed as NSSI (516, 62%), and three participants reported non-intentional self-injury, which was victim precipitated or planned but not acted on. There was some intent to die for just under half of the participants; for over three-quarters of participants there was a low medical risk of death, and medical treatment was sought or received for 659 (79.2%) participants.

Details of the primary caregiver are provided in *Table 10*. The majority of caregivers consenting to the trial were the mother of the young person (719, 86.4%) and their mean age was 42.5 years (SD 6.39 years). Three hundred and two (36.3%) were in full-time employment and 216 (26%) were in part-time employment.

TABLE 9 Characteristics of the index self-harm episode from the SASII

Characteristics of index self-harm episode, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Interviewer-rated behaviour			
Suicide attempt (SA)	148 (35.7)	165 (39.6)	313 (37.6)
NSSI	265 (63.9)	251 (60.2)	516 (62.0)
Non-intentional self-injury (victim precipitated, did not act)[a]	2 (0.5)	1 (0.2)	3 (0.4)
Interviewer-rated intent to die (some intent to die)	197 (47.5)	215 (51.6)	412 (49.5)
Interviewer-rated medical risk of death			
Low	326 (78.6)	302 (72.4)	628 (75.5)
Moderate	76 (18.3)	105 (25.2)	181 (21.8)
High	12 (2.9)	10 (2.4)	22 (2.6)
Missing	1 (0.2)	0 (0.0)	1 (0.1)
Interviewer-rated impulsivity (planned)	88 (21.2)	78 (18.7)	166 (20.0)
Communication of suicide intent (*Yes*: indirect/direct)	107 (25.8)	118 (28.3)	225 (27.0)
Interpersonal influence mentioned[b]	156 (37.6)	179 (42.9)	335 (40.3)
Emotion relief mentioned[c]	386 (93.0)	391 (93.8)	777 (93.4)
Medical treatment sought/received (*Yes*)	326 (78.6)	333 (79.9)	659 (79.2)
Interviewer-rated physical condition following episode			
No effect	64 (15.4)	56 (13.4)	120 (14.4)
Mild effect	299 (72.0)	310 (74.3)	609 (73.2)
Moderate effect	45 (10.8)	46 (11.0)	91 (10.9)
Severe effect	6 (1.4)	5 (1.2)	11 (1.3)
Missing	1 (0.2)	0 (0.0)	1 (0.1)

a These three events were included as, in the judgement of the research team, the young people were in the process of self-harming as described in the eligibility criteria but the act was interrupted either by the young person or another person. Thus, had medication but called for help and did not take it (participant 100); took a knife with the intention of self-harming but was interrupted (participant 243); hitting fists against head with undue risk-taking/unreasonable expectation of safety (participant 621). All three young people had other prior self-harm confirmed in the SASII timeline.
b With reason(s) for self-harm according to any of the following indicators of interpersonal influence: to communicate to or let others know how desperate I was; to get help; to gain admission into a hospital or treatment program; to shock or impress others; to get other people to act differently or change; to get back at or hurt someone; to demonstrate to others how wrong they are/were; to make others understand how desperate I am.
c With reason(s) for self-harm according to any of the following indicators of emotion relief: to stop bad feelings; to stop feeling angry or frustrated or enraged; to relieve anxiety or terror; to relieve feelings of aloneness, emptiness or isolation; to stop feeling self-hatred, shame; to obtain relief from a terrible state of mind.

Adapted from Cottrell *et al.*[70]

TABLE 10 Primary caregiver details

Primary caregiver details	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Relationship to young person, *n* (%)			
Mother	356 (85.8)	363 (87.1)	719 (86.4)
Father	45 (10.8)	48 (11.5)	93 (11.2)
Guardian	8 (1.9)	3 (0.7)	11 (1.3)
Stepmother	3 (0.7)	2 (0.5)	5 (0.6)
Foster parent	1 (0.2)	1 (0.2)	2 (0.2)
Stepfather	2 (0.5)	0 (0.0)	2 (0.2)
Gender (female), *n* (%)	366 (88.2)	369 (88.5)	735 (88.3)
Age mean (SD)	42.5 (6.68), *n* = 377	42.4 (6.08), *n* = 363	42.5 (6.39), *n* = 740
Employment status, *n* (%)			
Employee/self-employed, full time	152 (36.6)	150 (36.0)	302 (36.3)
Employee/self-employed, part-time	112 (27.0)	104 (24.9)	216 (26.0)
Homemaker	71 (17.1)	87 (20.9)	158 (19.0)
Employee on sick leave	13 (3.1)	12 (2.9)	25 (3.0)
Unemployed	42 (10.1)	51 (12.2)	93 (11.2)
Not in paid employment (e.g. working for charity)	4 (1.0)	4 (1.0)	8 (1.0)
Retired	6 (1.4)	3 (0.7)	9 (1.1)
Learning a trade, government-supported training	1 (0.2)	1 (0.2)	2 (0.2)
Full-time education	6 (1.4)	1 (0.2)	7 (0.8)
Missing	8 (1.9)	4 (1.0)	12 (1.4)

Baseline questionnaires

Baseline questionnaire scores are displayed in *Table 11* for questionnaires completed by both the young person and caregiver (ICU, SDQ, McMaster FAD), in *Table 12* for questionnaires completed by the young person only (CDRS, PQ-LES-Q, Beck, Hopelessness Scale) and in *Table 13* for questionnaires completed by the caregiver only (GHQ-12, Family Questionnaire). *Appendix 3* summarises the description of questionnaire scores and cut-off points applied.

Scores were largely similar across both groups. Over half the young people and caregivers demonstrated a very high total difficulties score on the SDQ, and over three-quarters of young people demonstrated unhealthy family functioning on all scales of the McMaster FAD (while caregivers reported less unhealthy traits). At least mild depression was reported for the majority of participants (762, 91.7%) and suicidal ideation was demonstrated for 543 (66.5%).

Study conduct

Protocol violators

A total of 16 (1.9%) young people were identified as not fulfilling the eligibility criteria (*Table 14*), five in the FT arm and 11 in the TAU arm. The main criterion violated was participants already being treated within CAMHS (defined as three or more sessions of active treatment in the months preceding randomisation) for seven participants: three in the FT arm and four in the TAU arm.

TABLE 11 Young person and caregiver baseline questionnaire outcomes

Outcome, mean (SD), *n*	Young person			Caregiver		
	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
ICU[a]						
ICU total score (0–72)	28.2 (9.10), 410	28.5 (9.09), 403	28.4 (9.09), 813	32.6 (11.59), 413	32.9 (11.43), 415	32.8 (11.50), 828
ICU callousness (0–33)	8.3 (4.61), 410	8.6 (4.99), 403	8.5 (4.80), 813	10.4 (6.62), 413	10.7 (6.68), 415	10.5 (6.65), 828
ICU uncaring score (0–24)	11.0 (4.65), 410	10.8 (4.57), 403	10.9 (4.61), 813	14.5 (4.85), 413	14.5 (4.82), 415	14.5 (4.83), 828
ICU unemotional score (0–15)	8.9 (2.82), 410	9.1 (3.09), 403	9.0 (2.96), 813	7.6 (3.08), 413	7.7 (3.08), 415	7.7 (3.07), 828
SDQ[b]						
Total difficulties score (0–40)	19.6 (5.70), 413	20.1 (5.60), 415	19.8 (5.65), 828	19.4 (6.56), 412	19.8 (6.83), 415	19.6 (6.69), 827
Close to average, *n* (%)	84 (20.3)	70 (16.9)	154 (18.6)	78 (18.9)	86 (20.7)	164 (19.8)
Slightly raised, *n* (%)	58 (14.0)	68 (16.4)	126 (15.2)	51 (12.4)	36 (8.7)	87 (10.5)
High, *n* (%)	60 (14.5)	44 (10.6)	104 (12.6)	67 (16.3)	66 (15.9)	133 (16.1)
Very high, *n* (%)	211 (51.1)	233 (56.1)	444 (53.6)	216 (52.4)	227 (54.7)	443 (53.6)
Prosocial score (0–10)	7.1 (1.84), 414	7.2 (1.87), 415	7.2 (1.85), 829	6.3 (2.33), 414	6.3 (2.30), 416	6.3 (2.31), 830
Emotional problems score (0–10)	6.2 (2.34), 414	6.5 (2.31), 415	6.4 (2.33), 829	6.2 (2.39), 414	6.2 (2.60), 415	6.2 (2.49), 829
Conduct problems score (0–10)	3.9 (2.11), 414	4.0 (1.94), 415	3.9 (2.02), 829	4.2 (2.42), 413	4.3 (2.47), 416	4.2 (2.44), 829
Hyperactivity score (0–10)	6.3 (2.22), 413	6.2 (2.17), 415	6.2 (2.20), 828	5.5 (2.48), 414	5.7 (2.56), 415	5.6 (2.52), 829
Peer problems score (0–10)	3.1 (1.87), 413	3.5 (2.08), 415	3.3 (1.98), 828	3.5 (2.09), 413	3.6 (2.14), 415	3.6 (2.11), 828
Impact score (0–10)	3.3 (2.41), 413	3.5 (2.52), 414	3.4 (2.47), 827	4.4 (2.73), 410	4.3 (2.73), 413	4.4 (2.73), 823
Externalising score (0–20)	10.2 (3.80), 413	10.1 (3.48), 415	10.2 (3.65), 828	9.7 (4.30), 413	10.0 (4.50), 415	9.9 (4.40), 828
Internalising score (0–20)	9.4 (3.41), 413	10.0 (3.63), 415	9.7 (3.53), 828	9.8 (3.69), 413	9.8 (4.02), 415	9.8 (3.86), 828
McMaster FAD (1–4)[c]						
Overall FAD score	2.5 (0.33), 404	2.4 (0.36), 405	2.5 (0.35), 809	2.2 (0.36), 408	2.2 (0.36), 415	2.2 (0.36), 823
General functioning	2.5 (0.53), 410	2.5 (0.56), 408	2.5 (0.54), 818	2.3 (0.48), 415	2.3 (0.47), 416	2.3 (0.47), 831
Unhealthy (≥ 2.0), *n* (%)	354 (86.3)	339 (83.1)	693 (84.7)	314 (75.7)	316 (76.0)	630 (75.8)

continued

TABLE 11 Young person and caregiver baseline questionnaire outcomes (*continued*)

Outcome, mean (SD), *n*	Young person			Caregiver		
	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Behaviour control subscale	2.1 (0.38), 413	2.1 (0.37), 409	2.1 (0.38), 822	1.8 (0.41), 411	1.8 (0.41), 416	1.8 (0.41), 827
Unhealthy (≥ 1.9), *n* (%)	319 (77.2)	311 (76.0)	630 (76.6)	197 (47.9)	226 (54.3)	423 (51.1)
Affective involvement subscale	2.5 (0.48), 412	2.5 (0.49), 409	2.5 (0.48), 821	2.2 (0.45), 412	2.3 (0.51), 415	2.2 (0.48), 827
Unhealthy (≥ 2.1), *n* (%)	345 (83.7)	341 (83.4)	686 (83.6)	270 (65.5)	272 (65.5)	542 (65.5)
Affective responsiveness subscale	2.6 (0.48), 410	2.6 (0.51), 408	2.6 (0.50), 818	2.1 (0.55), 411	2.1 (0.57), 415	2.1 (0.56), 826
Unhealthy (≥ 2.2), *n* (%)	343 (83.7)	335 (82.1)	678 (82.9)	195 (47.4)	197 (47.5)	392 (47.5)
Roles subscale	2.5 (0.35), 412	2.5 (0.37), 409	2.5 (0.36), 821	2.5 (0.42), 415	2.5 (0.42), 415	2.5 (0.42), 830
Unhealthy (≥ 2.3), *n* (%)	310 (75.2)	309 (75.6)	619 (75.4)	304 (73.3)	309 (74.5)	613 (73.9)
Communication subscale	2.6 (0.37), 413	2.6 (0.38), 409	2.6 (0.37), 822	2.3 (0.44), 413	2.3 (0.41), 415	2.3 (0.43), 828
Unhealthy (≥ 2.2), *n* (%)	360 (87.2)	343 (83.9)	703 (85.5)	248 (60.0)	255 (61.4)	503 (60.7)
Problem-solving subscale	2.5 (0.47), 410	2.5 (0.53), 409	2.5 (0.50), 819	2.2 (0.48), 411	2.2 (0.48), 416	2.2 (0.48), 827
Unhealthy (≥ 2.2), *n* (%)	331 (80.7)	306 (74.8)	637 (77.8)	234 (56.9)	249 (59.9)	483 (58.4)

a Higher scores represent higher callous and unemotional traits.
b Higher scores in all but the prosocial score represent greater issues in that category. For the prosocial score lower scores represent greater issues. See *Appendix 3* for categorisation of the total difficulties score (differs for young person and caregiver).
c Higher scores are indicative of poorer family functioning.

Research withdrawals

Participants could withdraw from further clinical data collection, further researcher follow-up visits or postal questionnaire follow-up. In 160 (19.2%) cases, the young person, the caregiver or both withdrew from some or all of the study process: 60 (14.5%) in the FT arm and 100 (24.0%) in the TAU arm.

Table 15 presents the type of withdrawals and who withdrew (young person or caregiver). All 160 participants withdrew from researcher follow-up visits, 136 (16.3%) withdrew from postal questionnaire follow-up (FT arm, 11.6%; TAU arm, 21.1%) and 22 (2.6%) withdrew from further clinical data collection (i.e. the primary end point), nine in the FT arm and 13 in the TAU arm.

Figure 7 presents the timing of withdrawals and shows that these were largely clustered around postal and researcher follow-up time points at 3, 6, 12 and 18 months post randomisation. The main reasons for withdrawal (*Table 16*) included the participant not wanting to be involved any more (38 participants); other issues or events going on in the participants' or their family life, which meant that they were too busy to continue (30 participants); the participant having improved or moved on and not wanting to revisit the past (26 participants); and dissatisfaction with the therapy offered or received (18 participants).

TABLE 12 Young person baseline questionnaire outcomes

Outcome	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
CRDS[a]			
Total CDRS-R score (17–113), mean (SD), *n*	48.0 (14.19), 415	49.4 (13.29), 416	48.7 (13.76), 831
Not depressed (< 30), *n* (%)	45 (10.8)	24 (5.8)	69 (8.3)
Mild depression (30–42), *n* (%)	108 (26.0)	108 (26.0)	216 (26.0)
Moderate depression (43–57), *n* (%)	154 (37.1)	168 (40.4)	322 (38.7)
Severe depression (58–72), *n* (%)	91 (21.9)	102 (24.5)	193 (23.2)
Very severe depression (> 72), *n* (%)	17 (4.1)	14 (3.4)	31 (3.7)
PQ-LES[b]			
Total PQ-LES-Q score (14–70), mean (SD), *n*	41.2 (9.37), 411	41.2 (9.45), 408	41.2 (9.41), 819
Overall self-assessment score (1–5), mean (SD), *n*	2.7 (1.04), 405	2.7 (1.00), 405	2.7 (1.02), 810
Overall, how has your life been?, *n* (%)			
Very poor	57 (14.1)	53 (13.1)	110 (13.6)
Poor	107 (26.4)	102 (25.2)	209 (25.8)
Fair	162 (40.0)	167 (41.2)	329 (40.6)
Good	59 (14.6)	70 (17.3)	129 (15.9)
Very good	20 (4.9)	13 (3.2)	33 (4.1)
Beck[c]			
Beck score (0–38),[d] mean (SD), *n*	10.8 (8.94), 408	10.4 (9.42), 408	10.6 (9.18), 816
Suicide ideation (BSS screening: Q4 and Q5) (*Yes*), *n* (%)	276 (67.6)	267 (65.4)	543 (66.5)
Number of previous suicide attempts (*N*)	395	403	798
Never attempted suicide, *n* (%)	176 (44.6)	170 (42.2)	346 (43.4)
Attempted suicide once, *n* (%)	114 (28.9)	124 (30.8)	238 (29.8)
Attempted suicide two or more times, *n* (%)	105 (26.6)	109 (27.0)	214 (26.8)
Severity of last suicide attempt (*N*)	213	226	439
Wish to die was low, *n* (%)	37 (17.4)	41 (18.1)	78 (17.8)
Wish to die was moderate, *n* (%)	78 (36.6)	79 (35.0)	157 (35.8)
Wish to die was high, *n* (%)	98 (46.0)	106 (46.9)	204 (46.5)
Hopelessness[e]			
Total hopelessness score (0–17),[e] mean (SD), *n*	7.7 (4.30), 409	7.3 (4.19), 406	7.5 (4.25), 815

a Higher scores represent greater levels of depression.
b Higher scores indicative of greater enjoyment and satisfaction.
c Median presented as Beck scores considerably skewed and zero inflated at follow-up.
d Higher scores indicate a higher level of suicide ideation.
e Higher scores reflect greater hopelessness or negative expectations towards the future.
Adapted from Cottrell *et al.*[70] © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license

Questionnaire follow-up

Postal questionnaires were completed and returned by 426 (51.2%) young people and 440 (52.9%) caregivers at 3 months, and by 353 (42.4%) young people and 363 (43.6%) caregivers at 6 months (*Table 17*). Response rates were higher in the FT group than in the TAU group.

TABLE 13 Caregiver baseline questionnaire outcomes

Outcome, mean (SD), *n*	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
GHQ-12[a]			
GHQ-12 score. Likert scale (0–36)	17.7 (7.06), 414	18.6 (7.24), 415	18.2 (7.16), 829
Family Questionnaire[b]			
Total score (20–80)	52.9 (10.67), 415	52.9 (10.85), 416	52.9 (10.75), 831
Emotional overinvolvement (10–40)	27.2 (4.93), 415	27.2 (5.03), 416	27.2 (4.98), 831
Criticism (10–40)	25.7 (6.97), 415	25.7 (7.06), 416	25.7 (7.01), 831

a Higher scores are indicative of greater psychological distress.
b Higher scores indicate greater levels of expressed emotion.

TABLE 14 Number and reasons for eligibility violation

Reasons for eligibility violation, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Has the participant fulfilled the eligibility criteria?			
Yes	410 (98.8)	406 (97.4)	816 (98.1)
No	5 (1.2)	11 (2.6)	16 (1.9)
Reason for protocol violation			
Already being treated in CAMHS	3 (60.0)	4 (36.4)	7 (43.8)
Not intended to offer CAMHS follow-up for self-harm[a]	1 (20.0)	2 (18.2)	3 (18.8)
Currently undergoing child Protection investigation	0 (0.0)	2 (18.2)	2 (12.5)
Not aged 11–17 years	0 (0.0)	1 (9.1)	1 (6.3)
Insufficient proficiency in English	0 (0.0)	1 (9.1)	1 (6.3)
Young person is pregnant	1 (20.0)	0 (0.0)	1 (6.3)
Sibling in SHIFT or already receiving FT in CAMHS	0 (0.0)	1 (9.1)	1 (6.3)
Total	5 (100.0)	11 (100.0)	16 (100.0)

a For young person to whom it was not intended to offer CAMHS follow-up for self-harm, one young person in TAU turned 18 years old 4 days post randomisation and follow-up was offered by adult mental health services; the other young person did go on to receive at least one allocated treatment session (one TAU session and six SHIFT FT sessions each).

Questionnaire completion at the 12- and 18-month researcher follow-up visits (*Table 18*) was similar to postal follow-up, with 465 (55.9%) young person questionnaires completed at 12 months and 395 (47.5%) at 18 months, again with higher follow-up rates in the FT group than in the TAU group.

At the 12-month researcher visit, full follow-up, defined as completion of questionnaires by the young person, caregiver and researcher, was achieved for 411 (49.4%) participants: 237 (57.1%) in the FT arm and 174 (41.7%) in the TAU. In the case of 350 (42.1%) participants, 147 (35.4%) in the FT arm and 203 (48.7%) in the TAU arm, no questionnaires at all were completed. Of the 421 participants for whom at least one questionnaire was missing, the most frequent reasons for loss to follow-up were that the researcher was unable to contact the participant (29%), the researcher made contact but was unable to arrange the visit (24.5%), both the young person and the caregiver withdrew consent (21.1%) and that a visit was arranged but was subsequently cancelled or no one was at home when the researcher visited (8.3%).

TABLE 15 Withdrawals

Withdrawals, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Number of participants with at least one type of withdrawal[a]	60 (14.5)	100 (24.0)	160 (19.2)
Young person with at least one type of withdrawal	56 (13.5)	97 (23.3)	153 (18.4)
Caregiver with at least one type of withdrawal	49 (11.8)	87 (20.9)	136 (16.3)
Withdrawn from			
Clinical data collection	9 (2.2)	13 (3.1)	22 (2.6)
Parent/caregiver only	1 (0.2)	0 (0.0)	1 (0.1)
Both	8 (1.9)	13 (3.1)	21 (2.5)
Researcher follow-up visits	60 (14.5)	100 (24.0)	160 (19.2)
Young person only	11 (2.7)	13 (3.1)	24 (2.9)
Parent/caregiver only	4 (1.0)	3 (0.7)	7 (0.8)
Both	45 (10.8)	84 (20.1)	129 (15.5)
Postal questionnaire follow-up	48 (11.6)	88 (21.1)	136 (16.3)
Young person only	6 (1.4)	10 (2.4)	16 (1.9)
Parent/caregiver only	4 (1.0)	4 (1.0)	8 (1.0)
Both	38 (9.2)	74 (17.7)	112 (13.5)

a Note that withdrawals were made on a total of 167 occasions for 160 participants (five participants withdrew at two different time points and one withdrew at three different time points).

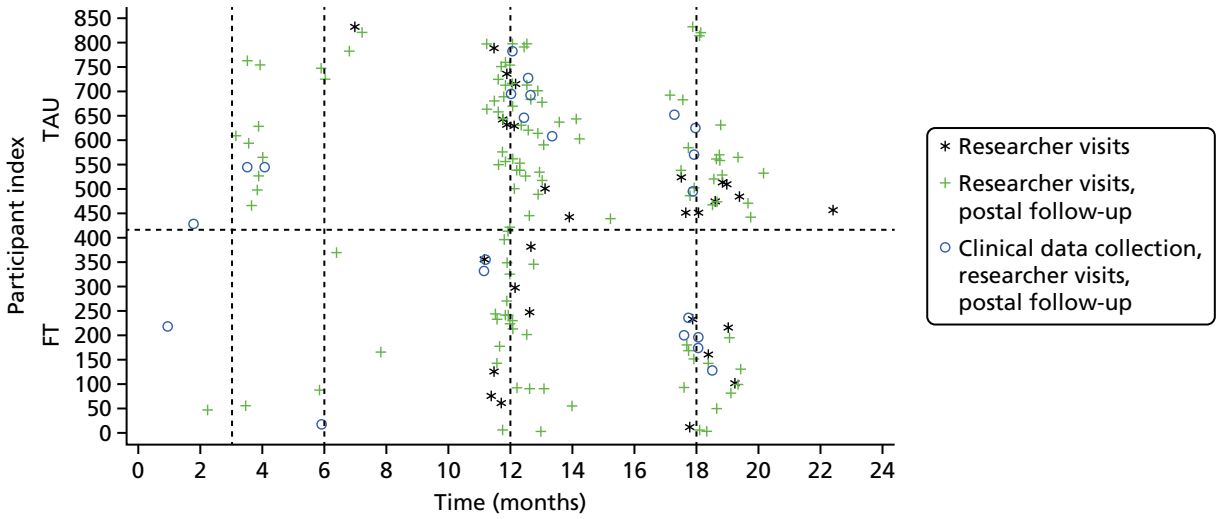


FIGURE 7 Time between randomisation and withdrawal.

At the 18-month researcher visit, full follow-up was achieved for 352 (42.3%) participants: 197 (47.5%) in the FT group and 155 (37.2%) in the TAU group. No questionnaires at all were completed for 412 (49.5%) participants: 186 (44.8%) in the FT group and 226 (54.2%) in the TAU group. Of the 480 participants for whom at least one questionnaire was missing, the most frequent reasons for loss to follow-up largely mirrored the reasons at 12 months; however, inability to contact the participant and withdrawal were more frequently cited as reasons for all questionnaires being missing (37.1% and 29.6% of participants, respectively).

The timing of questionnaire completion for the young person is presented in *Figure 8* with completion largely well defined around required time points and a similar distribution for the caregiver and the researcher questionnaires (at 12 and 18 months).

TABLE 16 Reasons for withdrawal

Reason for withdrawal, *n* (%)	Young person[a]			Caregiver		
	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Did not want to be involved any longer	10 (17.9)	28 (28.9)	38 (24.8)	10 (20.4)	20 (23.0)	30 (22.1)
Other things going on in the young person's or their family life	13 (23.2)	17 (17.5)	30 (19.6)	10 (20.4)	17 (19.5)	27 (19.9)
Improvement/moved on – did not want to revisit past	7 (12.5)	19 (19.6)	26 (17.0)	4 (8.2)	15 (17.2)	19 (14.0)
Dissatisfaction with the therapy offered/ received	10 (17.9)	8 (8.2)	18 (11.8)	9 (18.4)	11 (12.6)	20 (14.7)
Dissatisfaction with research process/ burden of questionnaires	2 (3.6)	4 (4.1)	6 (3.9)	2 (4.1)	4 (4.6)	6 (4.4)
Did not want to think or talk about it	2 (3.6)	4 (4.1)	6 (3.9)	1 (2.0)	3 (3.4)	4 (2.9)
Young person moved away/no longer in contact with caregiver	1 (1.8)	2 (2.1)	3 (2.0)	6 (12.2)	6 (6.9)	12 (8.8)
Missing	11 (19.6)	16 (16.5)	27 (17.6)	7 (14.3)	11 (12.6)	18 (13.2)
Total	56 (100)	97 (100)	153 (100)	49 (100)	87 (100)	136 (100)

a Note that the young person's reasons for withdrawal are not mutually exclusive as one young person withdrew on two occasions and provided different reasons each time.

TABLE 17 Postal questionnaire follow-up

Follow-up completion, *n* (%)	3 months[a]			6 months[b]		
	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person completed	228 (54.9)	198 (47.5)	426 (51.2)	204 (49.2)	149 (35.7)	353 (42.4)
Caregiver completed	241 (58.1)	199 (47.7)	440 (52.9)	213 (51.3)	150 (36.0)	363 (43.6)
Overall questionnaire completion						
Both young person and caregiver completed	213 (51.3)	180 (43.2)	393 (47.2)	187 (45.1)	125 (30.0)	312 (37.5)
Young person completed only	15 (3.6)	18 (4.3)	33 (4.0)	17 (4.1)	24 (5.8)	41 (4.9)
Caregiver completed only	28 (6.7)	19 (4.6)	47 (5.6)	26 (6.3)	25 (6.0)	51 (6.1)
Not completed	159 (38.3)	200 (48.0)	359 (43.1)	185 (44.6)	243 (58.3)	428 (51.4)

a At 3 months, a further three young person questionnaires and one caregiver questionnaire was completed retrospectively at the 12-month researcher visit; and six young person and seven caregiver questionnaires were completed retrospectively at the 18-month visit (health economic data only). Furthermore, six young person and four caregiver questionnaires were completed over the telephone at 3 months.

b At 6 months: a further five young person and eight caregiver questionnaires were completed retrospectively at the 12-month researcher visit and 11 young person and 10 caregiver questionnaires were completed retrospectively at the 18-month visit (health economic data only). Furthermore, seven young person and eight caregiver questionnaires were completed over the telephone at 6 months.

TABLE 18 Researcher-administered questionnaire completion and loss to follow-up

Questionnaire completion, *n* (%)	12 month-follow-up[a]			18-month follow-up[b]		
	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person questionnaire completed	261 (62.9)	204 (48.9)	465 (55.9)	213 (51.3)	182 (43.6)	395 (47.5)
Caregiver questionnaire completed	254 (61.2)	195 (46.8)	449 (54.0)	220 (53.0)	176 (42.2)	396 (47.6)
Researcher questionnaire completed	248 (59.8)	189 (45.3)	437 (52.5)	204 (49.2)	165 (39.6)	369 (44.4)
Overall questionnaire completion						
Young person, caregiver and researcher completed	237 (57.1)	174 (41.7)	411 (49.4)	197 (47.5)	155 (37.2)	352 (42.3)
Young person and researcher completed only	8 (1.9)	12 (2.9)	20 (2.4)	4 (1.0)	10 (2.4)	14 (1.7)
Young person and caregiver completed only	13 (3.1)	14 (3.4)	27 (3.2)	10 (2.4)	12 (2.9)	22 (2.6)
Young person completed only	3 (0.7)	4 (1.0)	7 (0.8)	2 (0.5)	5 (1.2)	7 (0.8)
Caregiver completed only	4 (1.0)	7 (1.7)	11 (1.3)	13 (3.1)	9 (2.2)	22 (2.6)
Researcher completed only	3 (0.7)	3 (0.7)	6 (0.7)	3 (0.7)	0 (0.0)	3 (0.4)
Not completed at all	147 (35.4)	203 (48.7)	350 (42.1)	186 (44.8)	226 (54.2)	412 (49.5)
Reason (least one) not completed						
Unable to contact	61 (34.3)	61 (25.1)	122 (29.0)	84 (38.5)	94 (35.9)	178 (37.1)
Contacted but unable to arrange visit	40 (22.5)	63 (25.9)	103 (24.5)	41 (18.8)	41 (15.6)	82 (17.1)
Withdrawal (young person and/or caregiver)	33 (18.5)	56 (23.0)	89 (21.1)	51 (23.4)	91 (34.7)	142 (29.6)
Visit cancelled/no one in	15 (8.4)	20 (8.2)	35 (8.3)	16 (7.3)	9 (3.4)	25 (5.2)
Young person moved	10 (5.6)	6 (2.5)	16 (3.8)	9 (4.1)	7 (2.7)	16 (3.3)
Visit not arranged in error	1 (0.6)	7 (2.9)	8 (1.9)	5 (2.3)	1 (0.4)	6 (1.3)
Young person in foster/ support care	1 (0.6)	5 (2.1)	6 (1.4)	0 (0.0)	2 (0.8)	2 (0.4)
Young person in hospital/ psychiatric unit	2 (1.1)	1 (0.4)	3 (0.7)	1 (0.5)	1 (0.4)	2 (0.4)
Risk issues	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.5)	0 (0.0)	1 (0.2)
Other	0 (0.0)	1 (0.4)	1 (0.2)	1 (0.5)	1 (0.4)	2 (0.4)
Not known	15 (8.4)	23 (9.5)	38 (9.0)	9 (4.1)	15 (5.7)	24 (5.0)
Total	178 (100.0)	243 (100.0)	421 (100.0)	218 (100.0)	262 (100.0)	480 (100.0)

a At 12 months, a further 24 researcher questionnaires (SASII only) were completed retrospectively at the 18-month visit. The majority of 12-month questionnaires were completed face to face during the researcher visit, however, 30 young person and 34 caregiver questionnaires were completed via post; one young person and five researcher questionnaires were completed over the telephone; and the completion method was unclear for four young person questionnaires, 14 caregiver questionnaires and five researcher questionnaires. The caregiver booklet was known to have been completed by a non-consenting caregiver for three participants.

b At 18 months, the majority of 18-month questionnaires were completed face to face during the researcher visit, however, 23 young person and 33 caregiver questionnaires were completed via post; five young person and 10 researcher questionnaires were completed over the telephone; and the completion method was unclear for four young person questionnaires, 13 caregiver questionnaires and two researcher questionnaires. The caregiver questionnaire was known to have been completed by a non-consenting caregiver for one participant.

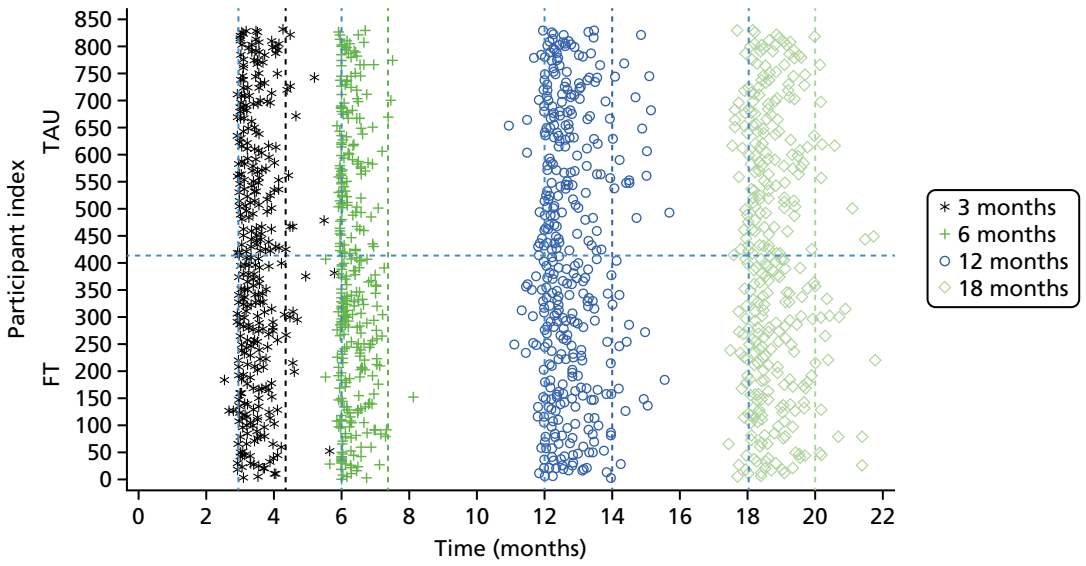


FIGURE 8 Time between randomisation and young person questionnaire follow-up.

Researcher unblinding

Unblinding of one or more researchers occurred on a total of 190 occasions for 161 (19.4%) participants: 109 times in 93 participants (22.4%) in the FT group and 81 times in 68 participants (16.3%) in the TAU group (*Table 19* and *Figure 9*). Unblinding most frequently occurred around the 12- and 18-month researcher follow-up time points (45.3% and 17.4%, respectively); however, 25.8% of unblinding occurred irrespective of a specific follow-up visit, largely within the first few months post randomisation and because a clinician notified the researcher. The rates and reasons for unblinding were similar across trial arms, with the most frequent cause of unblinding arising from the researcher being informed by the caregiver (74 cases, 38.9%), a clinician (44 cases, 23.2%), the young person (34 cases, 17.9%) or both the young person and the caregiver (18 cases, 9.5%), with a smaller number resulting from the young person's response implying a particular intervention (eight cases, 4.2%). Some unblinding occurred during researcher follow-up: 22 prior to the assessment, 56 during the assessment and 30 after the assessment.

To investigate the potential impact of unblinding on the primary outcome, we identified attendances where the only source of reason for presentation (self-harm related or not) was the unblinded researcher record [due to unclassified HES data, or researcher collected only attendances (two cases)]. There were 23 such attendances, involving 16 participants. Ten participants in the FT arm accounted for 16 attendances and six participants in the TAU arm accounted for seven attendances. The researchers reported three of these attendances to be related to self-harm (two in the FT arm and one in the TAU arm) and hospital follow-up took place a mean of 7 (SD 5.3) months after the researchers were unblinded (range 0–16 months).

Analysis populations

Intention-to-treat population

All summaries and analyses have been carried out using the ITT population, which consists of all randomised participants, excluding those who withdrew their full consent to trial participation or for whom written informed consent was not obtained.

Written informed consent was received for all randomised participants and, although some participants withdrew their consent to researcher and postal contact and further clinical data collection, no participants withdrew their full consent and asked for data already collected to be removed. Therefore, the ITT population consists of all 832 participants.

TABLE 19 Researcher unblinding

Research unblinding, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Number of participants for whom unblinding occurred	93 (22.4)	68 (16.3)	161 (19.4)
Number of times researchers were unblinded	109	81	190
Timing of unblinding			
3 months	6 (5.5)	7 (8.6)	13 (6.8)
6 months	4 (3.7)	5 (6.2)	9 (4.7)
12 months	58 (53.2)	28 (34.6)	86 (45.3)
18 months	16 (14.7)	17 (21.0)	33 (17.4)
Other	25 (22.9)	24 (29.6)	49 (25.8)
Total	109 (100.0)	81 (100.0)	190 (100.0)
How did unblinding occur			
Informed by caregiver	44 (40.4)	30 (37.0)	74 (38.9)
Informed by clinician	24 (22.0)	20 (24.7)	44 (23.2)
Informed by young person	19 (17.4)	15 (18.5)	34 (17.9)
Informed by young person and caregiver	14 (12.8)	4 (4.9)	18 (9.5)
Young person's responses implied a particular intervention	2 (1.8)	6 (7.4)	8 (4.2)
Other	4 (3.7)	6 (7.4)	10 (5.3)
Missing	2 (1.8)	0 (0.0)	2 (1.1)
Total	109 (100.0)	81 (100.0)	190 (100.0)
When did unblinding occur during the 12- or 18-month visit			
Before the assessment	14 (18.9)	8 (17.8)	22 (18.5)
During the assessment	38 (51.4)	18 (40.0)	56 (47.1)
After the assessment	17 (23.0)	13 (28.9)	30 (25.2)
Missing	5 (6.8)	6 (13.3)	11 (9.2)
Total	74 (100.0)	45 (100.0)	119 (100.0)

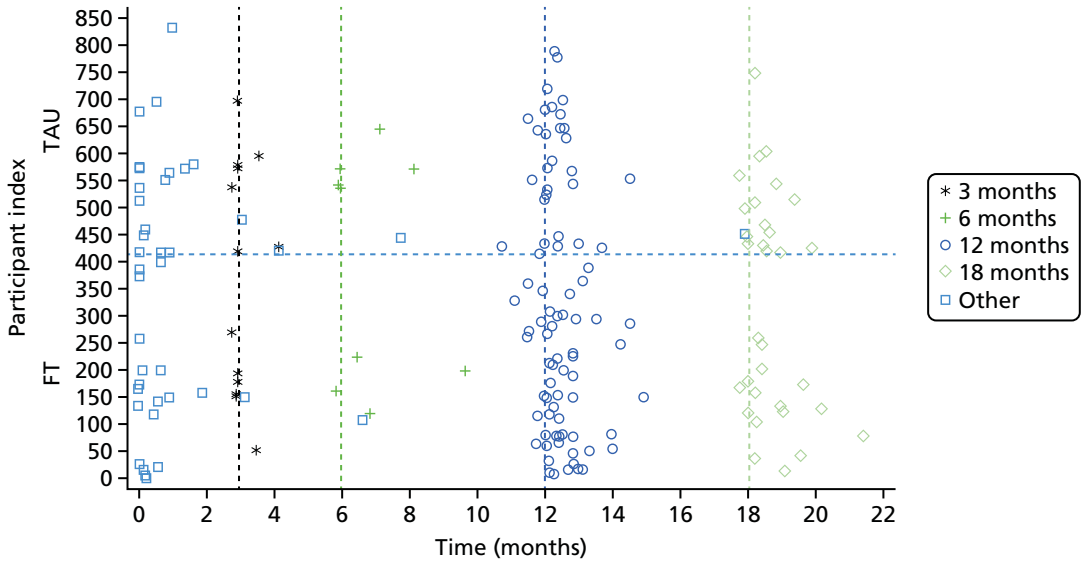


FIGURE 9 Time between randomisation and unblinding.

Two participants consented to all trial procedures excluding the registration of their details with NHS Digital. Primary outcome data for these participants were not obtained from NHS Digital and were collected only during visits by the researcher to acute trusts.

Complete written informed consent was obtained from the primary caregiver for all but two participants: in one case, the caregiver's consent form was not received and in another the form had not been dated. Implied consent was, however, obtained for both caregivers following the completion of the caregiver set of baseline questionnaires.

Per-protocol population

The per-protocol population consisted of the 816 (98.1%) participants without an eligibility violation. The number of patients excluded from the per-protocol population and the reasons are summarised in *Table 14*. No analyses were repeated for the per-protocol population.

Treatment summaries

Process and treatment summaries

Child and Adolescent Mental Health Services treatment pathways

Following randomisation, participants attended sessions within CAMHS following one of three main pathways (*Figure 10* and *Tables 20* and *21*).

- **Initial SHIFT FT or TAU sessions as per randomisation:** at least one session was attended by the young person and/or a family member for 394 (94.9%) participants allocated to FT and 338 (81.1%) participants allocated to TAU.
  Slightly more participants in the TAU arm did not attend initial therapy: 34 (8.2%) participants, compared with 21 (5.1%) participants in the FT arm. At the end of initial treatment, 233 (56.1%) FT participants and 158 (37.9%) TAU participants had completed all of the required sessions.
  In the FT arm, 173 (41.7%) participants attended SHIFT FT for more than the anticipated 6 months or had more than the intended eight sessions. This was primarily as a result of unresolved issues, difficulties with timetabling or cancellations.
- **Referrals within CAMHS for additional or alternative sessions:** referrals were reported for 98 (23.6%) participants allocated to FT during or following their SHIFT FT and for 96 (23.0%) TAU participants referred as part of usual care. In the FT arm, 77 (18.6%) young people and/or a family member attended at least one referred session while 68 (16.3%) participants did so in the TAU arm. Reasons for referral show that the proportion of participants referred for psychiatric mental state/risk assessment including medication was higher in the FT arm than in the TAU arm (43.9% vs. 28.1%), as was the proportion referred for ongoing treatment after trial treatment finished (19.4% vs. 2.1%). In contrast, the proportion of participants referred for alternative/additional therapeutic input was higher in the TAU arm than in the FT arm (47.9% vs. 17.3%).
- **Sessions that took place prior to initial SHIFT FT:** in the FT arm, 84 (20.2%) participants and/or a family member attended at least one session. These sessions occurred because they had been booked prior to randomisation, or were booked as holding appointments because of waiting times, with the majority of sessions (88.1%) consisting of assessment only. Only 13 (11.9%) sessions in 10 (2.4%) participants were for treatment.

Overall, 399 (96.1%) FT participants and 339 (81.3%) TAU participants had at least one session attended within CAMHS by either the participant (young person) or family member.

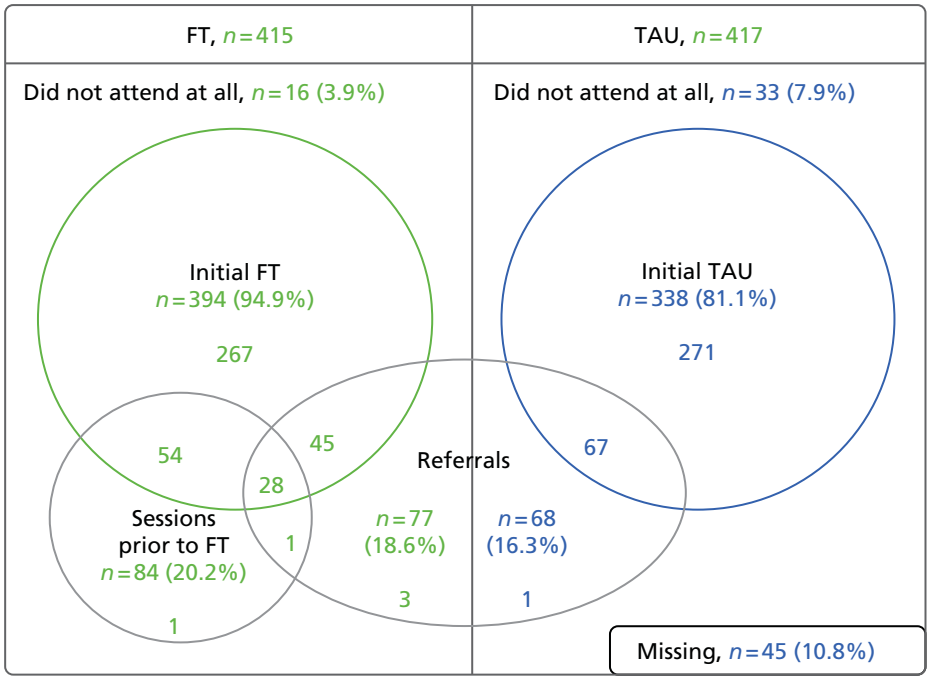


FIGURE 10 Diagram of attendance of at least one session, by anyone, within CAMHS pathways.

TABLE 20 Attendance of sessions within initial randomised treatment in CAMHS

Attendance at sessions, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Any attendance (young person or family) during initial FT/TAU			
Yes	394 (94.9)	338 (81.1)	732 (88.0)
No	21 (5.1)	34 (8.2)	55 (6.6)
Missing	0 (0.0)	45 (10.8)	45 (5.4)
Young person's attendance summary			
Did not attend at all	23 (5.5)	39 (9.4)	62 (7.5)
Dropped out with no negotiation	66 (15.9)	83 (19.9)	149 (17.9)
Negotiated ending in response to threat of dropout	92 (22.2)	57 (13.7)	149 (17.9)
Completed all required sessions	233 (56.1)	158 (37.9)	391 (47.0)
Treatment ongoing	0 (0.0)	10 (2.4)	10 (1.2)
Missing	1 (0.2)	70 (16.8)	71 (8.5)
Did the family attend SHIFT FT for more than 6 months and/or more than eight sessions?			
Yes	173 (41.7)		
No	242 (58.3)		
Rationale for additional time/sessions			
Issues not resolved, more sessions required	104 (60.1)		
Difficulty timetabling/cancellations	42 (24.3)		
Other[a]	5 (2.9)		
Missing	22 (12.7)		
Total – attending for > 6 months or > 8 sessions	173 (100.0)		

a Other reasons for additional time/sessions were that the family was keen to have expected final session; because of concerns arising and to minimise risks between SHIFT ending and referral for assessment and treatment commencing in CAMHS; developmental history had to be taken for autistic spectrum disorder assessment referral; father joined therapy, which extended the work; family felt issues were resolved.

TABLE 21 Referrals within CAMHS and additional sessions attended within CAMHS

Referrals within CAMHS and additional CAMHS sessions, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person referred within CAMHS			
Yes[a]	98 (23.6)	96 (23.0)	194 (23.3)
Attended referred session(s)	*77 (18.6)*	*68 (16.3)*	*145 (17.4)*
No	314 (75.7)	267 (64.0)	581 (69.8)
Missing	3 (0.7)	54 (12.9)	57 (6.9)
Reason for referral (non-mutually exclusive)			
Assessment			
Psychiatric mental state/risk assessment *including medication*	43 (43.9)	27 (28.1)	70 (36.1)
Treatment			
For alternative/additional therapeutic input	17 (17.3)	46 (47.9)	63 (32.5)
Ongoing treatment after trial treatment finished	19 (19.4)	2 (2.1)	21 (10.8)
For more intensive treatment (*e.g. day/inpatient unit*)	3 (3.1)	2 (2.1)	5 (2.6)
Assessment/treatment			
For input to specific comorbid diagnoses (*e.g. ASD, eating disorder*)	12 (12.2)	7 (7.3)	19 (9.8)
Other			
Administrative reasons (e.g. change of address, local protocol, travel problems)	3 (3.1)	4 (4.2)	7 (3.6)
Missing	11 (11.2)	16 (16.7)	27 (13.9)
Total number referred	98 (100)	96 (100)	194 (100)
Young person attended sessions prior to SHIFT FT			
Yes	84 (20.2)	N/A	N/A
Treatment sessions	*8 (1.9)*		
Assessment sessions	*74 (17.8)*		
Assessment and treatment sessions	*2 (0.5)*		
No	302 (72.8)		
Missing	29 (7.0)		
Purpose of attended sessions			
Treatment	13 (11.9)		
Assessment	96 (88.1)		
Total number of sessions prior to SHIFT	109 (100.0)		

N/A, not applicable.

a Twenty-six young persons (11 FT, 15 TAU) were reported to have been referred but were missing details of appointments for the referral and a further 23 young persons (10 FT, 13 TAU) were referred but attended no sessions. There were a total of 202 referrals in 194 young persons with 141 (72 FT, 69 TAU) referred once, 24 (14 FT, 10 TAU) referred twice, one young person (TAU) referred three times and two young persons (one FT, one TAU) referred five times each, and the remaining 26 were missing details of all referrals.

Number and duration of sessions within Child and Adolescent Mental Health Services

During participants' initial randomised treatment, the number of SHIFT FT sessions attended by the family or the young person ranged from 0 to 21, with a mean of 6.1 (SD 4.02) sessions over a mean period of 4.7 (SD 3.03) months. The number of initial sessions attended in the TAU arm was more variable, ranging from 0 to 163, with a mean of 8.2 (SD 12.61) sessions over a mean period of 5.7 (SD 5.41) months (*Tables 22* and *23* and *Figures 11–14*). As a result of highly skewed data, although some TAU participants a large number of sessions, the overall median number of sessions and duration of attendance was lower in the TAU arm, a median of 4 sessions over 4.1 months, compared with 6 sessions over 5.1 months in the FT arm.

TABLE 22 Number of CAMHS sessions attended during initial randomised treatment, additional referrals and sessions prior to SHIFT FT

	FT, *n* (%) (*N* = 415)				TAU, *n* (%) (*N* = 417)			Total, *n* (%) (*N* = 832)
Number of sessions attended	Initial FT	Additional TAU referral	Sessions prior to initial FT	All sessions	Initial TAU	Additional TAU referral	All sessions	All sessions
0	21 (5.1)	324 (78.1)	302 (72.8)	16 (3.9)	34 (8.2)	280 (67.1)	33 (7.9)	49 (5.9)
1	40 (9.6)	16 (3.9)	68 (16.4)	28 (6.7)	52 (12.5)	23 (5.5)	45 (10.8)	73 (8.8)
2	29 (7.0)	19 (4.6)	11 (2.7)	32 (7.7)	37 (8.9)	10 (2.4)	36 (8.6)	68 (8.2)
3	37 (8.9)	2 (0.5)	2 (0.5)	36 (8.7)	45 (10.8)	8 (1.9)	43 (10.3)	79 (9.5)
4	32 (7.7)	5 (1.2)	2 (0.5)	29 (7.0)	23 (5.5)	4 (1.0)	23 (5.5)	52 (6.3)
5	29 (7.0)	6 (1.4)	1 (0.2)	31 (7.5)	28 (6.7)	4 (1.0)	30 (7.2)	61 (7.3)
6	43 (10.4)	3 (0.7)		41 (9.9)	13 (3.1)	1 (0.2)	15 (3.6)	56 (6.7)
7	29 (7.0)	2 (0.5)		32 (7.7)	15 (3.6)	2 (0.5)	10 (2.4)	42 (5.0)
8	50 (12.0)	4 (1.0)		38 (9.2)	11 (2.6)	1 (0.2)	12 (2.9)	50 (6.0)
9	30 (7.2)	3 (0.7)		21 (5.1)	12 (2.9)	2 (0.5)	11 (2.6)	32 (3.8)
10	24 (5.8)	2 (0.5)		23 (5.5)	10 (2.4)	4 (1.0)	9 (2.2)	32 (3.8)
11–15	38 (9.2)	7 (1.7)		48 (11.6)	29 (7.0)	2 (0.5)	32 (7.7)	80 (9.6)
16–20	12 (2.9)	3 (0.7)		25 (6.0)	27 (6.5)	3 (0.7)	26 (6.2)	51 (6.1)
21–25	1 (0.2)	0 (0.0)		6 (1.4)	17 (4.1)	2 (0.5)	19 (4.6)	25 (3.0)
> 26	0 (0.0)	5 (1.2)		9 (2.2)	19 (4.6)	2 (0.5)	28 (6.7)	37 (4.4)
Missing	0 (0.0)	14 (3.4)	29 (7.0)	0 (0.0)	45 (10.8)	69 (16.5)	45 (10.8)	45 (5.4)
Number of sessions attended								
N	415	401	386	415	372	348	372	787
Missing	0	14	29	0	45	69	45	45
Mean (SD)	6.1 (4.02)	1.4 (5.39)	0.3 (0.64)	7.7 (7.23)	8.2 (12.61)	1.1 (4.44)	9.3 (13.85)	8.5 (10.90)
Median (range)	6 (0–21)	0 (0–57)	0 (0–5)	6 (0–70)	4 (0–163)	0 (0–51)	5 (0–163)	6 (0–163)

TABLE 23 Timing and duration of treatment

Timing and duration of treatment	FT			TAU			Total
	Initial FT (*N* = 415)	Additional TAU referral (*N* = 98)	All sessions (*N* = 415)	Initial TAU (*N* = 417)	Additional TAU referral (*N* = 96)	All sessions (*N* = 417)	All sessions (*N* = 832)
Time to first attendance (days/months)	***Days***	***Months***	***Days***	***Days***	***Months***	***Days***	***Days***
n	394	77	399	338	67	339	738
Missing	0	11	0	45	16	45	45
Did not attend sessions	21	10	16	34	13	33	49
Mean (SD)	24.5 (21.49)	4.8 (3.75)	19.9 (18.81)	37.0 (44.87)	6.2 (4.27)	36.8 (44.77)	27.6 (34.36)
Median (range)	18.0 (1–137)	4.1 (0.1–12.9)	15.0 (0–137)	22.0 (0–380)	5.3 (0.2–17.7)	22.0 (0–380)	17.0 (0–380)
Time to last attendance (months)							
n	394	77	399	338	67	339	738
Missing	0	11	0	45	16	45	45
Did not attend sessions	21	10	16	34	13	33	49
Mean (SD)	5.5 (3.06)	9.5 (5.83)	6.2 (4.03)	6.9 (5.40)	9.3 (5.31)	7.4 (5.59)	6.7 (4.84)
Median (range)	5.8 (0.3–17.4)	8.7 (0.7–18.0)	5.8 (0.3–18.0)	5.0 (0.2–18.0)	9.0 (1.1–18.0)	5.3 (0.2–18.0)	5.7 (0.2–18.0)
Duration first to last attendance (months)							
n	394	77	399	338	67	339	738
Missing	21	21	16	79	29	78	94
Did not attend sessions	21	10	16	34	13	33	49
Mean (SD)	4.7 (3.03)	4.7 (4.99)	5.5 (4.06)	5.7 (5.41)	3.1 (3.83)	6.2 (5.63)	5.8 (4.86)
Median (range)	5.1 (0.0–15.7)	3.0 (0.0–17.7)	5.3 (0.0–17.9)	4.1 (0.0–17.7)	1.4 (0.0–15.4)	4.4 (0.0–17.9)	5.1 (0.0–17.9)

The inclusion of additional referred sessions attended by 96 (23.0%) participants in the TAU arm referred within CAMHS to additional sessions or sessions alternative to their initial treatment provides the overall summary of all TAU received. Combined with initial TAU sessions, the overall number of all sessions attended ranged from 0 to 163, with a mean of 9.3 (SD 13.85) sessions over a mean period of 6.2 (SD 5.63) months and a median of 5 sessions over a period of 4.4 months.

In the FT arm, 98 (23.6%) participants received additional TAU sessions as part of a referral within CAMHS. The number of referred sessions attended ranged from 0 to 57, with a mean of 1.4 (SD 5.39) sessions over a mean period of 4.7 (4.99) months and a median of 0 sessions over 3 months. Combined with initial SHIFT FT sessions and sessions prior to starting FT, the overall number of all sessions attended in the FT arm ranged from 0 to 70, with a mean of 7.7 (SD 7.23) sessions over a mean period of 5.5 months (SD 4.06 months) and a median of 6 sessions over a median period of 5.3 months.

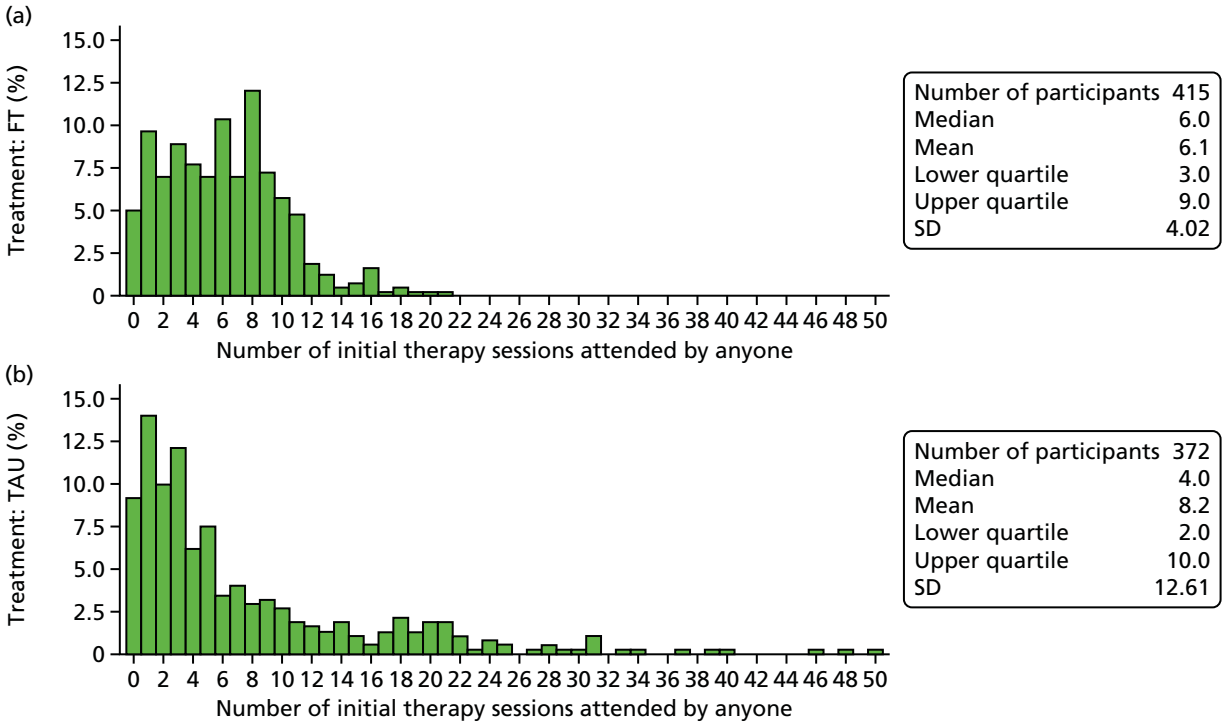


FIGURE 11 Number of initial sessions attended by anyone. Not displayed are three further participants in the TAU arm who had 75, 90 and 163 initial and overall sessions attended by anyone and one participant in the FT arm who had 70 overall sessions attended by anyone. (a) FT and (b) TAU.

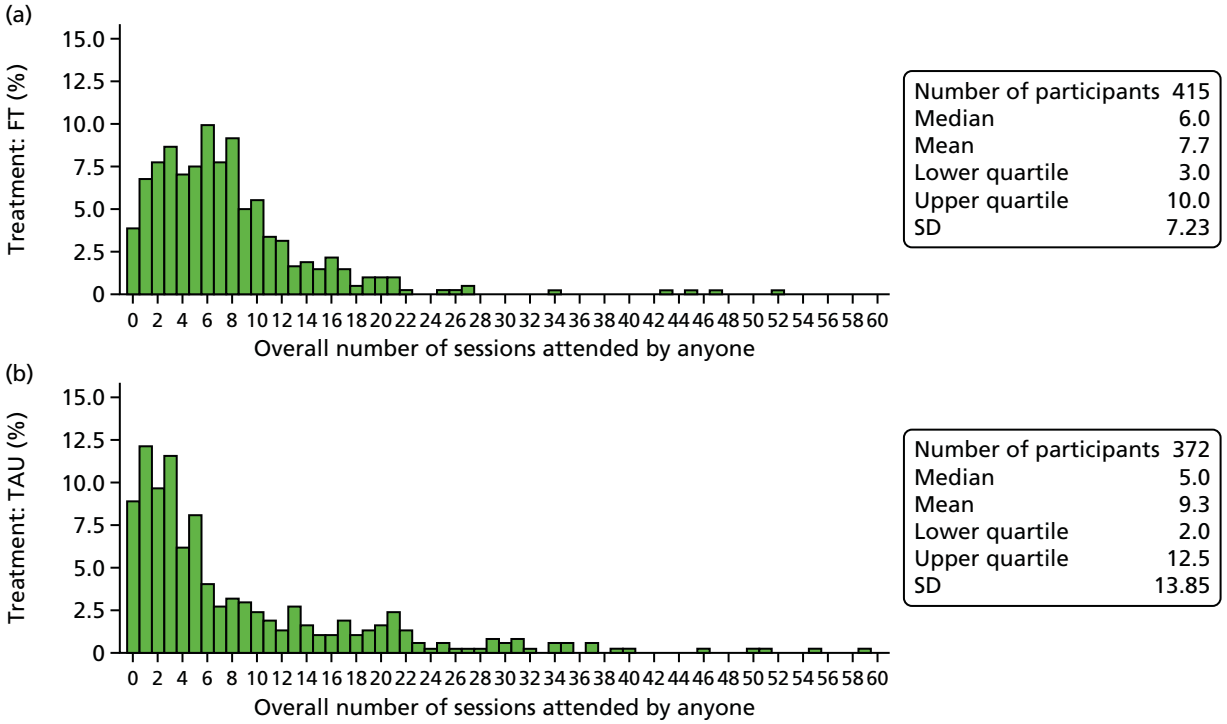


FIGURE 12 Number of overall sessions attended by anyone. Not displayed are three further participants in the TAU arm with attendance of 75, 90 and 163 initial and overall sessions and one participant in the FT arm with attendance of 70 overall sessions. (a) FT and (b) TAU.

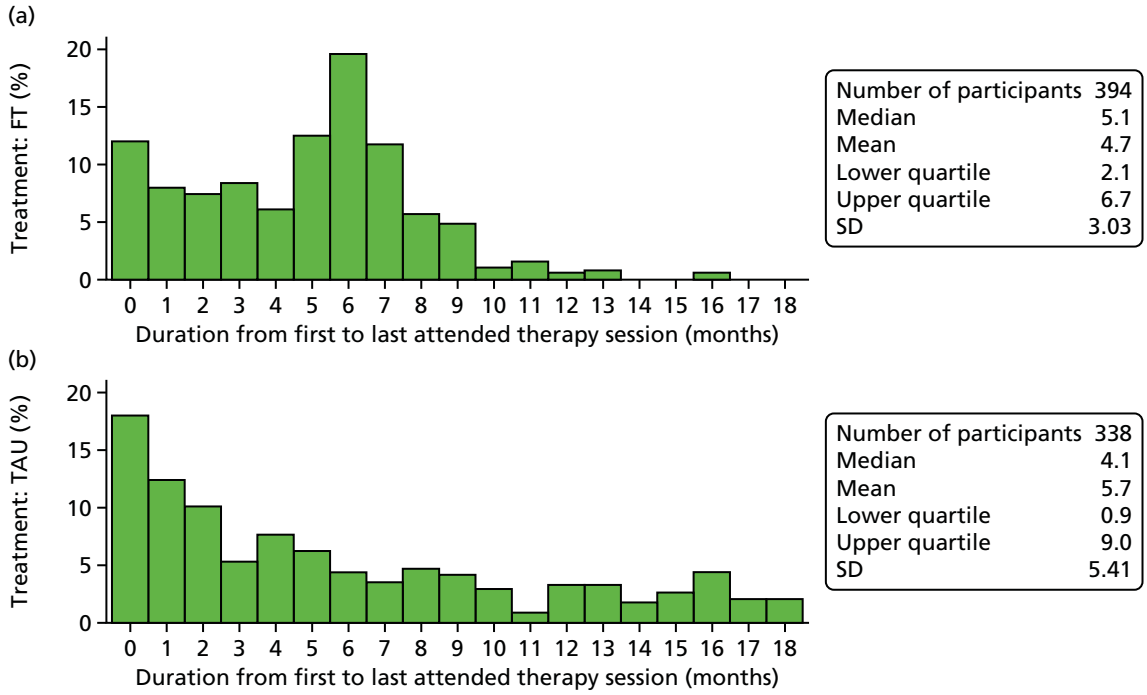


FIGURE 13 Duration of initial randomised treatment (excluding referrals and sessions prior to SHIFT FT). (a) FT and (b) TAU.

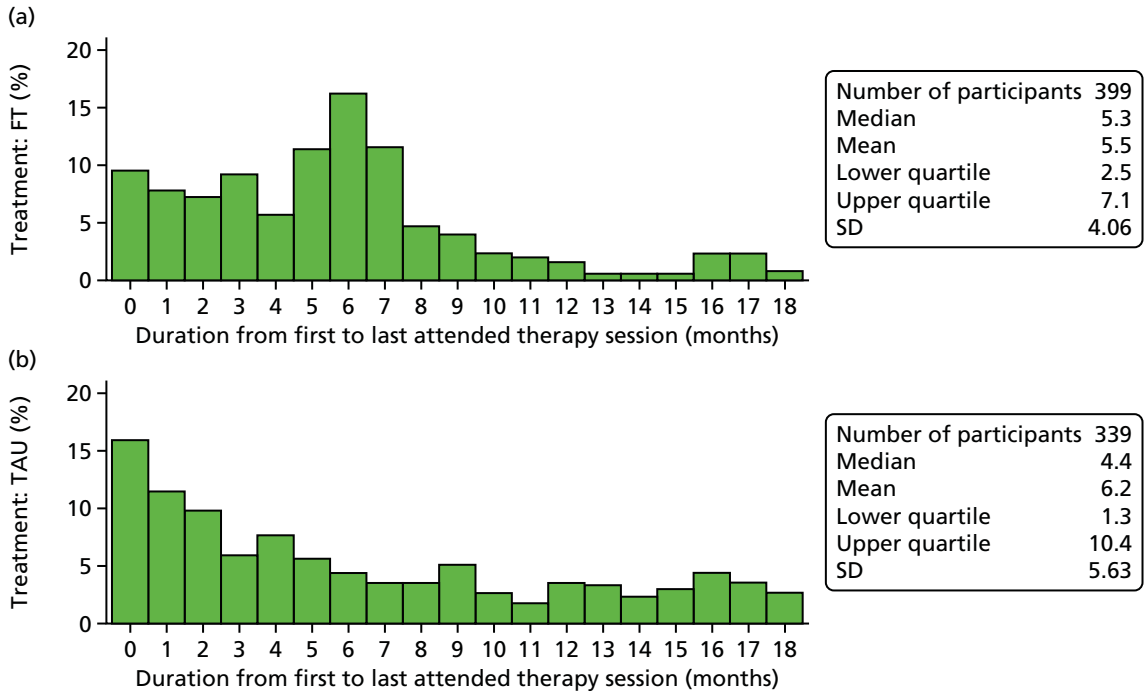


FIGURE 14 Duration of overall treatment (including referrals and sessions prior to SHIFT FT). (a) FT and (b) TAU.

Composition of sessions

The vast majority of SHIFT FT sessions were family sessions (2412, 95.3%) as opposed to individual sessions, with a mean of 1.5 (SD 0.75) family members attending in addition to, or without, the young person. TAU consisted of 1464 (47.7%) individual sessions, 1353 (44.1%) family sessions, 109 (3.6%) individual and family sessions (when a single session was mixed) and 98 (3.2%) group sessions (involving a group of young people led by a therapist). The mean number of therapists per session was lower for TAU sessions (1.2, SD 0.56) than for FT sessions (2.3, SD 0.72) (*Table 24*).

TABLE 24 Length of sessions, number of therapists and family members involved

Length of sessions, number of therapists and family members involved	FT (*N* = 415)				TAU (*N* = 417)			Total (*N* = 832)
	Initial FT	Additional TAU referral	Sessions prior to initial FT	All sessions	Initial TAU	Additional TAU referral	All sessions	All sessions
Number of participants represented at or attending at least one session	394	77	84	399	338	68	339	738
Number of sessions attended	2532	566	109	3207	3066	400	3466	6673
Length of sessions (minutes)								
N	2422	363	Missing	2785	1825	218	2043	4828
Missing	110	203		422	1241	182	1423	1845
Mean (SD)	74.1 (14.36)	60.3 (13.75)		72.3 (15.01)	60.4 (9.88)	60.2 (9.96)	60.3 (9.89)	67.2 (14.36)
Median (range)	75 (15–150)	60.0 (30–120)		70 (15–150)	60 (10–180)	60 (15–120)	60.0 (10–180)	60 (10–180)
Number of therapists involved								
N	2525	549	Missing	3074	2944	394	3338	6412
Missing	7	17		133	122	6	128	261
Mean (SD)	2.3 (0.72)	1.3 (0.71)		2.2 (0.82)	1.2 (0.56)	1.4 (0.74)	1.2 (0.59)	1.7 (0.85)
Median (range)	2.0 (1.0–4.0)	1.0 (1.0–5.0)		2.0 (1.0–5.0)	1.0 (1.0–8.0)	1.0 (1.0–5.0)	1.0 (1.0–8.0)	1.0 (1.0–8.0)
Type of session,* n *(%)								
Individual	115 (4.5)	274 (48.4)	56 (51.4)	445 (13.9)	1464 (47.7)	203 (50.8)	1667 (48.1)	2112 (31.6)
Family	2412 (95.3)	247 (43.6)	47 (43.1)	2716 (84.7)	1353 (44.1)	174 (43.5)	1527 (44.1)	4243 (63.6)
Group	N/A	26 (4.6)	1 (0.9)	27 (0.8)	98 (3.2)	21 (5.3)	119 (3.4)	146 (2.2)
Individual and family	N/A	10 (1.8)	5 (4.6)	5 (0.2)	109 (3.6)	0 (0.0)	109 (3.1)	114 (1.7)
Missing	5 (0.2)	9 (1.6)	0 (0.0)	14 (0.4)	42 (1.4)	2 (0.5)	44 (1.3)	58 (0.9)
Number of family members in family sessions								
N	2412	251	Missing	2663	1392	153	1545	4208
Missing	0	6		58	70	21	91	149
Mean (SD)	1.5 (0.75)	1.3 (0.62)		1.5 (0.75)	1.4 (0.67)	1.4 (0.78)	1.4 (0.68)	1.5 (0.73)
Median (range)	1.0 (1.0–5.0)	1.0 (1.0–4.0)		1.0 (1.0–5.0)	1.0 (1.0–6.0)	1.0 (1.0–7.0)	1.0 (1.0–7.0)	1.0 (1.0–7.0)

N/A, not applicable.

Overall therapeutic orientation and therapy pathways

Overall, 398 (95.9%) FT participants and 301 (72.2%) TAU participants attended at least one therapeutically orientated session (excluding assessment sessions), and one FT participant and 32 TAU participants attended assessment sessions only (*Table 25*). A further six TAU participants attended only sessions for which the content was unknown, while 16 (3.9%) FT participants and 33 (7.9%) TAU participants attended no sessions at all.

Therapeutic orientation of sessions in TAU varied considerably (*Table 26*). One-quarter of all sessions were for supportive therapy/counselling, 17.4% were for cognitive–behavioural therapy (CBT), 11.5% were for family work, 10.7% were for formal systemic FT and < 5% were for each of psychodynamic therapy, communication skills/problem-solving, interpersonal therapy, dialectical behaviour therapy (DBT), psychoeducational therapies and other therapy. The most frequently attended sessions on a per-participant basis were supportive therapy/counselling, family work, CBT and formal systemic FT. A substantial proportion of sessions were for assessment rather than therapy, with 12.5% of all sessions in 161 (38.6%) participants for other assessment/review.

Seventy-nine per cent of all sessions attended in the FT arm were SHIFT sessions, with other types of sessions accounting for < 5% each. Eighty-nine participants (21.4%) attended an assessment/review and 47 participants (11.3%) attended a mental state/risk assessment.

The main therapeutic orientation(s), making up one-third or more of all therapy sessions per participant (excluding assessment and unknown sessions), varied from one to a maximum of three main therapies per participant (*Table 27*).

In the TAU arm, supportive therapy/counselling was again one of the main therapies for 115 (27.6%) participants, CBT was the main therapy for 73 (17.5%), family work was the main therapy for 66 (15.8%) and formal systemic FT was the main therapy for 57 (13.7%). As the main therapy comprised up to three

TABLE 25 Treatment mix of sessions: therapy and assessment

Treatment mix of sessions, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Any therapy sessions attended?			
Yes	398 (95.9)	301 (72.2)	699 (84.0)
No	17 (4.1)	71 (17.0)	88 (10.6)
Missing	0 (0.0)	45 (10.8)	45 (5.4)
Any assessment sessions attended?			
Yes	117 (28.2)	197 (47.2)	314 (37.7)
No	298 (71.8)	175 (42.0)	473 (56.9)
Missing	0 (0.0)	45 (10.8)	45 (5.4)
Overall session mix			
No sessions attended	16 (3.9)	33 (7.9)	49 (5.9)
Assessment only	1 (0.2)	32 (7.7)	33 (4.0)
Therapy only	282 (68.0)	136 (32.6)	418 (50.2)
Assessment and therapy	116 (28.0)	165 (39.6)	281 (33.8)
Unknown only	0 (0.0)	6 (1.4)	6 (0.7)
Missing	0 (0.0)	45 (10.8)	45 (5.4)

TABLE 26 Therapeutic orientation of all sessions and per participant

Therapeutic orientation of sessions, *n* (%)	Therapeutic orientation per session		Therapeutic orientation per participant (non-mutually exclusive)	
	FT (*N* = 3207)	TAU (*N* = 3466)	FT (*N* = 415)	TAU (*N* = 417)
Therapy				
SHIFT FT	2532 (79.0)	0 (0.0)	394 (94.9)	0 (0.0)
Supportive therapy/counselling	142 (4.4)	871 (25.1)	21 (5.1)	158 (37.9)
CBT	79 (2.5)	602 (17.4)	10 (2.4)	88 (21.1)
Family work	35 (1.1)	397 (11.5)	9 (2.2)	116 (27.8)
Formal systemic FT	14 (0.4)	371 (10.7)	5 (1.2)	87 (20.9)
Communication skills/problem-solving	0 (0.0)	62 (1.8)	0 (0.0)	29 (7.0)
Psychoeducational	22 (0.7)	29 (0.8)	6 (1.4)	18 (4.3)
Interpersonal therapy	5 (0.2)	38 (1.1)	1 (0.2)	10 (2.4)
DBT	0 (0.0)	30 (0.9)	0 (0.0)	8 (1.9)
Psychodynamic	1 (0.0)	131 (3.8)	1 (0.2)	5 (1.2)
Other therapy	78 (2.4)	168 (4.8)	17 (4.1)	28 (6.7)
Assessment				
Mental state/risk assessment	110 (3.4)	137 (4.0)	47 (11.3)	58 (13.9)
Other assessment/review	153 (4.8)	434 (12.5)	89 (21.4)	161 (38.6)
Medication review	26 (0.8)	80 (2.3)	13 (3.1)	28 (6.7)
Non-therapy[a]	2 (0.1)	27 (0.8)	1 (0.2)	23 (5.5)
Unknown: per session/all sessions	8 (0.2)	89 (2.6)	5 (1.2)	27 (6.5)
No sessions attended	–	–	16 (3.9)	33 (7.9)
Missing all treatment data	–	–	0 (0.0)	45 (10.8)

DBT, dialectical behaviour therapy.
a For example, introductory meetings/administration.

different therapies, these summaries are not mutually exclusive and full details of combinations of therapies can be found in *Table 28*. In the TAU arm, 256 (61.4%) participants attended only one type of therapy in over one-third of all attended sessions, 46 (11.0%) participants attended two main therapies and five (1.2%) participants attended three main therapies.

In the FT arm, the allocated therapy, SHIFT, was the main therapy attended by the vast majority (94%) of participants; however, the main therapies attended in addition to, or instead of, SHIFT are detailed in *Table 29*.

Eighty-seven (20.9%) participants in the TAU arm were reported to have attended at least one formal systemic FT session, albeit of a type different from SHIFT, which is a specific manualised form of formal systemic FT (*Table 30* and *Figure 15*). Furthermore, five (1.2%) participants in the FT arm, in addition to attending 7–16 SHIFT sessions, also attended 1–5 sessions of formal systemic FT as part of a referral within CAMHS for ongoing treatment after trial treatment finished.

TABLE 27 Main therapeutic orientation(s) of sessions (> 33% of therapy sessions)

FT (*N* = 415)		TAU (*N* = 417)	
Main therapy (not mutually exclusive)	***n* (%)**	**Main therapy (not mutually exclusive)**	***n* (%)**
No sessions attended	16 (3.9)	No sessions attended	33 (7.9)
Assessment only	1 (0.2)	Assessment only	32 (7.7)
SHIFT FT	389 (93.7)	Supportive therapy/counselling	115 (27.6)
Supportive therapy/counselling	14 (3.4)	CBT	73 (17.5)
CBT	6 (1.4)	Family work	66 (15.8)
Other therapy	6 (1.4)	Formal systemic FT	57 (13.7)
Psychoeducational	3 (0.7)	Communication skills/problem-solving	15 (3.6)
Family work	2 (0.5)	Other therapy	15 (3.6)
Interpersonal therapy	1 (0.2)	Psychoeducational	5 (1.2)
		Interpersonal therapy	4 (1.0)
		Psychodynamic	4 (1.0)
		DBT	3 (0.7)
		Unknown session orientation	6 (1.4)
		Missing all treatment data	45 (10.8)

Note
Up to three main therapeutic orientations were derived per young person. Note that four TAU young persons had evenly spread sessions across three orientations, with each slightly below the 33% threshold, and these have been included.

TABLE 28 Main therapeutic orientation of sessions (> 33% of therapy sessions) for participants randomised to receive TAU (mutually exclusive)

Main therapy, *n* (%)	TAU (*N* = 417)
Single main therapy	256 (61.4)
Supportive therapy/counselling	85 (20.4)
CBT	51 (12.2)
Family work	44 (10.6)
Formal systemic FT	43 (10.3)
Other therapy	9 (2.2)
Communication skills/problem-solving	8 (1.9)
Unknown only	6 (1.4)
Interpersonal therapy	4 (1.0)
Psychodynamic	4 (1.0)
Psychoeducational	2 (0.5)
Two main therapies	46 (11.0)
Supportive therapy/counselling, family work	14 (3.4)
CBT, formal systemic FT	7 (1.7)
CBT, supportive therapy/counselling	6 (1.4)
CBT, family work	2 (0.5)
Communication skills/problem-solving, family work	2 (0.5)

TABLE 28 Main therapeutic orientation of sessions (> 33% of therapy sessions) for participants randomised to receive TAU (mutually exclusive) (*continued*)

Main therapy, *n* (%)	TAU (*N* = 417)
DBT, supportive therapy/counselling	2 (0.5)
Formal systemic FT, other therapy	2 (0.5)
Supportive therapy/counselling, communication skills/problem-solving	2 (0.5)
Supportive therapy/counselling, other therapy	2 (0.5)
CBT, DBT	1 (0.2)
CBT, other therapy	1 (0.2)
CBT, psychoeducational	1 (0.2)
Communication skills/problem solving, formal systemic FT	1 (0.2)
Family work, formal systemic FT	1 (0.2)
Psycho-educational, supportive therapy/counselling	1 (0.2)
Supportive therapy/counselling, formal systemic FT	1 (0.2)
Three main therapies	5 (1.2)
CBT, family work, formal systemic FT	1 (0.2)
CBT, formal systemic FT, other therapy	1 (0.2)
CBT, supportive therapy/counselling, communication skills/problem-solving	1 (0.2)
CBT, supportive therapy/counselling, family work	1 (0.2)
Psychoeducational, communication skills/problem-solving, family work	1 (0.2)
Assessment only	32 (7.7)
No sessions attended	33 (7.9)
Missing	45 (10.8)

TABLE 29 Main therapeutic orientation of sessions (> 33% of therapy sessions) attended by participants randomised to receive FT (mutually exclusive)

Main therapy, *n* (%)	FT (*N* = 415)
Single main therapy	375 (90.4)
SHIFT FT	367 (88.4)
Supportive therapy/counselling	4 (1.0)
Other therapy	2 (0.5)
Family work	1 (0.2)
Psychoeducational	1 (0.2)
Two main therapies	23 (5.5)
SHIFT FT, supportive therapy/counselling	10 (2.4)
SHIFT FT, CBT	5 (1.2)
SHIFT FT, other therapy	4 (1.0)
SHIFT FT, psychoeducational	2 (0.5)
SHIFT FT, interpersonal therapy	1 (0.2)
CBT, family work	1 (0.2)
Assessment only	1 (0.2)
No sessions attended	16 (3.9)

TABLE 30 Summary of FT sessions received

FT sessions received, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Any SHIFT FT?			
Yes	394 (94.9)	0 (0.0)	394 (47.4)
No	21 (5.1)	372 (89.2)	393 (47.2)
Missing	0 (0.0)	45 (10.8)	45 (5.4)
Any TAU formal systemic FT?			
Yes	5 (1.2)[a]	87 (20.9)	92 (11.1)
No	410 (98.8)	285 (68.3)	695 (83.5)
Missing	0 (0.0)	45 (10.8)	45 (5.4)
Any FT?			
Yes	394 (94.9)	87 (20.9)	481 (57.8)
SHIFT FT	*389 (93.7)*	*0 (0.0)*	*389 (46.8)*
SHIFT FT and TAU FT	*5 (1.2)*	*N/A*	*5 (0.6)*
TAU FT	*0 (0.0)*	*87 (20.9)*	*87 (10.5)*
No	21 (5.1)	285 (68.3)	306 (36.8)
Missing	0 (0.0)	45 (10.8)	45 (5.4)

N/A, not applicable.

a For three young persons, the TAU Formal systemic FT was delivered by non-SHIFT clinicians. For one young person, a trained SHIFT clinician was involved; however, this clinician had not been part of the assigned team initially treating the young person for SHIFT, whereas for the remaining young person, two additional non-SHIFT formal systemic FT sessions were received immediately following the end of SHIFT FT involving the same clinicians, albeit with the addition of a non-SHIFT clinician who continued to then see the young person for family work and other therapy.

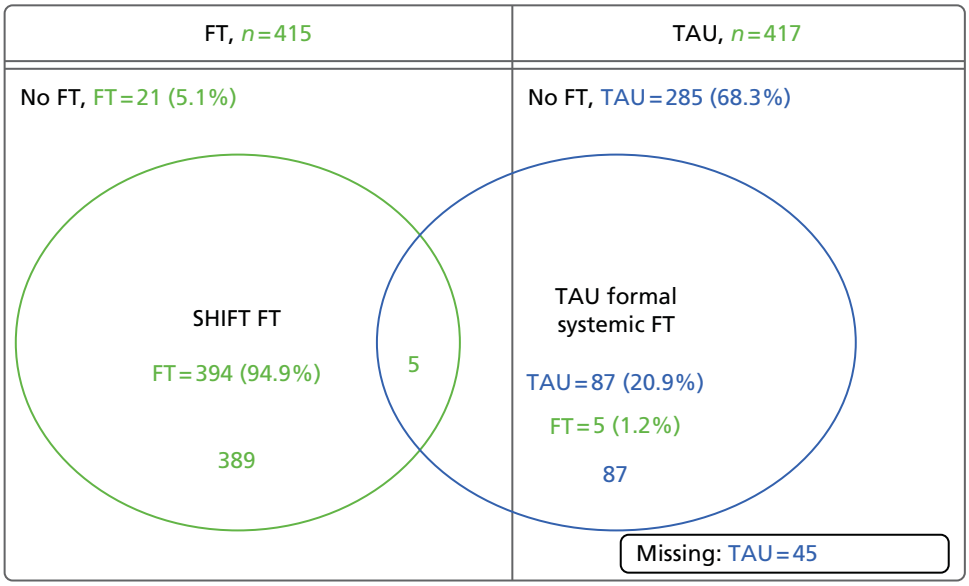


FIGURE 15 Diagram of FT attendance across trial arms.

Telephone contact, liaison and case manager meetings during initial treatment

During initial randomised treatment (i.e. not including further referrals within CAMHS), rates of telephone contact with participants' families were similar in both trial arms with 190 (45.8%) FT participants and 170 (40.8%) TAU participants receiving telephone contact, lasting in total a mean of 54.1 minutes and 49.1 minutes, respectively (*Table 31*).

Rates of liaison with other agencies by CAMHS were also similar in both groups (*Table 32*), this being the case for 151 (36.4%) FT participants and 136 (32.6%) TAU participants. The most frequently contacted host agencies were school/education services, which were contacted about 194 young people (67.6% of

TABLE 31 Details of telephone contact with participants between sessions

Telephone contact	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Any telephone contact with family between sessions, *n* (%)			
Yes	190 (45.8)	170 (40.8)	360 (43.3)
No	212 (51.1)	194 (46.5)	406 (48.8)
Missing	13 (3.1)	53 (12.7)	66 (7.9)
Total length of telephone contact (minutes)			
N	180	125	305
Missing	10	45	55
Mean (SD)	54.1 (47.47)	49.1 (50.56)	52.1 (48.74)
Median (range)	40.0 (7–295)	30.0 (5–442)	40.0 (5–442)

TABLE 32 Liaison with other agencies during initial therapy

Liaison with other agencies, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Liaison with other agencies during treatment			
Yes	151 (36.4)	136 (32.6)	287 (34.5)
No	258 (62.2)	227 (54.4)	485 (58.3)
Missing	6 (1.4)	54 (12.9)	60 (7.2)
Host agency – per participant (non-mutually exclusive)			
School/education	104 (68.9)	90 (66.2)	194 (67.6)
Social services	43 (28.5)	44 (32.4)	87 (30.3)
Child protection agency	5 (3.3)	4 (2.9)	9 (3.1)
Youth offending teams	2 (1.3)	7 (5.1)	9 (3.1)
Counselling services	7 (4.6)	15 (11.0)	22 (7.7)
Other[a]	47 (31.1)	53 (39.0)	100 (34.8)
Missing	2 (1.3)	7 (5.1)	9 (3.1)
Total	151 (100.0)	136 (100.0)	287 (100.0)

a Other agencies included A&E/hospital, adult mental health, adoption services, other CAMHS, dietitian, domestic violence group, psychiatrist, family support service, GP, housing support, insight team, voluntary/charity (Connexions, Befriending Society, Signpost), pharmacy, police, safeguarding, youth worker.

those for whom liaison took place), followed by social services, which were contacted about 87 young people (30.3%).

During the delivery of initial SHIFT FT, meetings were held with the case manager of 149 (35.9%) participants. A total of 299 such meetings were held, with a mean of two meetings per participant, and involved one professional; in the majority of meetings (71.2%) the family were not in attendance and the meeting took place face to face (62.2%).

Contamination

Nine SHIFT-trained family therapists were involved in the direct contamination of 11 (2.6%) young people in the TAU arm, with these participants attending at least one session in CAMHS with a SHIFT-trained family therapist. Six (1.4%) attended at least one formal systemic FT session (between one and nine sessions) with the SHIFT family therapist acting as the lead therapist for three (0.7%) participants and as part of the reflecting team for three (0.7%) participants. Two (0.5%) further participants received family work-based sessions with the SHIFT family therapist and sessions with the SHIFT family therapist for the remaining three (0.7%) participants were for supportive therapy/counselling and review only.

Formal systemic FT or family work sessions involving SHIFT-trained family therapists and TAU participants led to the possible contamination of a further nine TAU clinicians who were also present during these sessions. Four of these TAU clinicians subsequently saw at least one further participant, a further six (1.4%) TAU participants in total, of whom four (1%) received family work (but no systemic FT).

A further source of contamination occurred during 11 SHIFT FT sessions for seven (1.7%) FT participants, during which a TAU clinician was present in addition to the SHIFT family therapists (all different SHIFT FT teams). A total of seven TAU clinicians were contaminated in this way, three of whom subsequently saw at least one further participant, a further seven (1.7%) TAU participants in total, of whom one received two sessions of systemic FT and a further two received family work sessions.

Therapist characteristics and training

Thirty-one family therapists received training in SHIFT FT. Training was conducted in formal group sessions for the majority of family therapists and via individual training and observation for those few therapists who joined the trial late following staff changes. Family therapists recruited early to the trial had the opportunity to deliver therapy to at least one pilot case prior to delivering trial FT; therapists recruited following staff changes always observed existing team members prior to commencing trial FT.

Baseline characteristics were received for 189 (65.6%) of the 288 treating therapists: 29 (93.5%) SHIFT family therapists and 160 (62.3%) TAU clinicians. The mean age of family therapists was 48.9 (SD 6.95) years and of TAU clinicians was 43.6 (SD 10.58) years. There were more female therapists than males: 21 family therapists (72.4%) and 117 (73.1%) TAU clinicians were women. Job titles varied for TAU clinicians, with over half listed as practitioners (31, 19.4%), clinical psychologists (28, 17.5%) or nurses (27, 16.9%). Eighteen (11.3%) TAU clinicians had family therapist as their job title. As the use of job titles changes from service to service, TAU clinicians were classified into ‘senior’ (doctors at consultant level and all other non-medical staff on pay grade 8C or above) and ‘non-senior’ (all other clinicians). All SHIFT family therapists were classified as ‘senior’ whereas only five (3.1%) TAU clinicians were. The majority of family therapists (20, 60.6%) reported that their basic qualification was in social work, while TAU clinicians were trained in nursing (51, 25.2%), social work (30, 16.9%) and clinical psychology (29, 16.4%). Of the additional training received, 45.9% of training received by the SHIFT family therapists was in formal systemic therapy, compared with 21.4% for the TAU clinicians. TAU clinicians received more additional training in CBT and behaviour therapy and communication skills/problem-solving techniques.

Adherence and supervision

One thousand six hundred and fifty-four (65.3%) initial FT sessions were recorded and, of those, the digital versatile disc (DVD) was returned for 1433 (86.6%). Fifty-two (3.6%) DVDs underwent initial review and, of these, the young person was present in all but one session. Twenty-three (44.2%) reviewed recordings related to the first therapy session and 29 (55.8%) related to later sessions. Twenty-six (83.8%) family therapists had therapy sessions reviewed; two sessions were reviewed for each therapist.

Overall adherence for the 52 reviewed sessions was scored between 0 and 5, with 5 indicating greater adherence. Overall competence was scored between 0 and 6, with 6 indicating greater competence. Initial reviews scored overall mean adherence as 4.6 (SD 0.72) and overall competence as 4.4 (1.03).

Specialist reviews were conducted when the overall competence score was 0, 1 or 2 or when requested by the initial reviewer. Specialist reviews were conducted for 7 (13.5%) of the initially reviewed sessions and the young person was present for all of the sessions. Three (42.9%) related to the first session and four (57.1%) related to later therapy sessions. Five family therapists had specialist reviews; two sessions were reviewed for two therapists and one session was reviewed for three therapists.

Tables 33 and *34* present a comparison of overall adherence scores and competence ratings between the initial and specialist reviews.

Two hundred and thirty-eight (67.2%) families who attended at least two sessions of SHIFT FT received a formulation letter from their family therapist. Formulation letters were sent by 25 (80.6%) of the 31 SHIFT family therapists. Compliance by therapist, for those who sent letters, ranged between 22.0% and 100.0%, with a median of 80.0%.

All but one of the family therapists had at least one session of supervision; the family therapist who did not receive supervision was not a main therapist and attended only one session of FT, as part of the reflecting team. Two hundred and twenty-five supervision sessions were conducted in total: 98 in Yorkshire, 38 in Manchester and 89 in London. The number of sessions received by individual therapist is presented by hub in *Table 35*. Family therapists in Yorkshire received the most supervision followed by London and Manchester. Two hundred and sixty-six young people (119 in Yorkshire, 56 in Manchester and 97 in London) were discussed 584 times in supervision; overall, each young person was discussed a mean of 2.2 (SD 1.45) times, and this was similar for both Yorkshire and London; however, young people in Manchester were discussed less often (1.5 times, SD 0.76 times). The mean length of supervision sessions overall was 96.3 (SD 38.97) minutes and this was similar between hubs.

TABLE 33 Comparison of overall adherence scores

	Specialist review					
Initial review	0	1	2	3	4	5
0						
1						
2						1
3				1	1	1
4		1	1			1
5						

TABLE 34 Comparison of overall competence scores

	Specialist review						
Initial review	Highly inappropriate performance	Inappropriate performance with major problems evident	Evidence of competence but numerous problems/lack of consistency	Competent but some problems and/or inconsistencies	Good but minor problems and/or inconsistencies	Very good, minimal problems and/or inconsistencies	Excellent performance even in the face of patient difficulties
Highly inappropriate performance							
Inappropriate performance with major problems evident							
Evidence of competence but numerous problems/ lack of consistency		1	3				
Competent but some problems and/or inconsistencies		1	1				
Good but minor problems and/or inconsistencies				1			
Very good, minimal problems and/or inconsistencies							
Excellent performance even in the face of client difficulties							

TABLE 35 Number of sessions of supervision per therapist received by hub

Hub	Number of supervision sessions received per therapist
Yorkshire	
N	10
Missing	0
Mean (SD)	21.0 (6.68)
Median (range)	23 (9–29)
Manchester	
N	6
Missing	0
Mean (SD)	10.8 (7.41)
Median (range)	11.5 (1–18)
London	
N	15
Missing	1
Mean (SD)	15.5 (9.01)
Median (range)	18 (0–27)
Overall	
N	31
Missing	1
Mean (SD)	16.4 (8.61)
Median (range)	18 (0–29)

Family therapy alliance

The strength of therapeutic alliance as reported by the young person, caregiver and SHIFT family therapist as captured in the SOFTA questionnaire at the participant's third treatment session is summarised in *Table 36* and *Figure 16*, with higher scores representing greater alliance. Caregivers consistently reported the highest levels of alliance across all four dimensions, while the therapist generally reported the lowest, with the exception of 'engagement in the therapeutic process', for which the young person reported lower alliance. Young person and therapist alliance were most closely aligned, with overlapping CIs, for the two dimensions reflecting the uniqueness of conjoint treatment 'shared sense of purpose' and 'safety within the therapeutic system', while there was no overlap for engagement or 'emotional connection', nor for any of the caregiver scores. Alliance is further explored by the participant's primary outcome in the mediation results section.

Psychotropic medications

During treatment, clinicians and CAMHS notes reported that 104 (12.5%) young people had been prescribed a psychotropic medication (*Table 37*): 44 (10.6%) young people in the FT arm and 60 (14.4%) young people in the TAU arm. The most frequently prescribed psychotropic medications were antidepressants, prescribed for 80 (76.9%) of these young people (reported by clinician, CSO or family), which in the majority of cases (all but five prescriptions) were selective serotonin reuptake inhibitors as opposed to tricyclic antidepressants.

TABLE 36 Summary statistics for the SOFTA

SOFTA dimensions	Young person (*N* = 415)	Caregiver (*N* = 415)	Family therapist (*N* = 415)
Shared sense of purpose (4–20), *n*	273	279	294
Mean (SD)	15.2 (3.31)	16.6 (2.75)	15.1 (2.87)
95% CI	(14.9 to 15.6)	(16.2 to 16.9)	(14.8 to 15.4)
Median (range)	16.0 (5.0–20)	17.0 (4.0–20)	15.0 (4.0–20)
Engagement in the process (4–20), *n*	274	279	293
Mean (SD)	13.8 (3.01)	15.8 (2.62)	14.8 (2.65)
95% CI	(13.4 to 14.1)	(15.5 to 16.1)	(14.5 to 15.1)
Median (range)	14.0 (4.0–20)	16.0 (4.0–20)	15.0 (4.0–20)
Emotional connection (4–20), *n*	274	278	293
Mean (SD)	15.2 (2.83)	16.5 (2.33)	14.3 (1.82)
95% CI	(14.9 to 15.5)	(16.2 to 16.7)	(14.1 to 14.5)
Median (range)	16.0 (7.0–20)	17.0 (7.0–20)	14.0 (8.0–20)
Safety (4–20), *n*	274	279	294
Mean (SD)	13.8 (3.45)	16.7 (2.64)	13.3 (2.73)
95% CI	(13.3 to 14.2)	(16.3 to 17.0)	(13.0 to 13.6)
Median (range)	14.0 (4.0–20)	17.0 (4.0–20)	14.0 (4.0–19)
Total score (16–80), *n*	274	279	293
Mean (SD)	57.9 (10.51)	65.4 (8.65)	57.5 (8.30)
95% CI	(56.7 to 59.2)	(64.4 to 66.4)	(56.5 to 58.4)
Median (range)	58.3 (30–80)	66.0 (19–80)	59.0 (28–77)

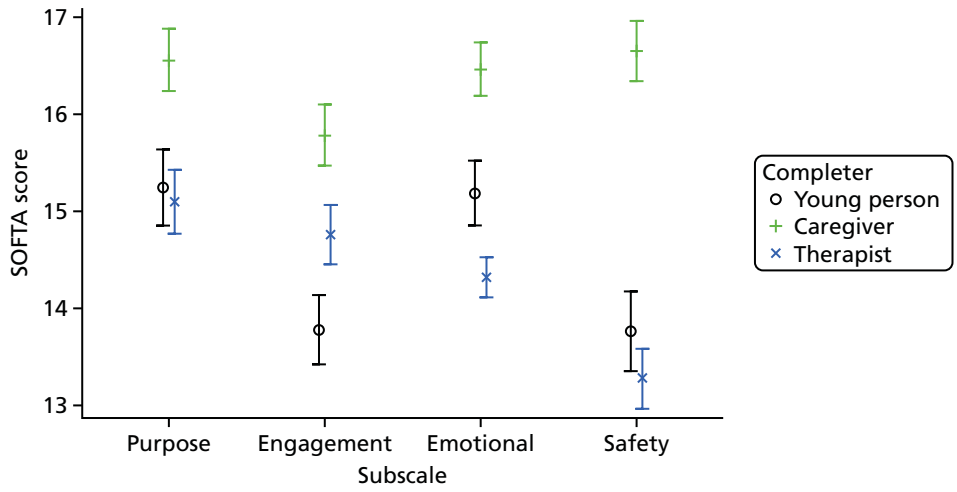


FIGURE 16 The SOFTA subscale scores with 95% CIs.

TABLE 37 Prescribed psychotropic medication use during treatment and overall at baseline and during 18-month follow-up

Psychotropic medication, *n* (%)	During treatment			At baseline and during 18-month follow-up		
	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person taking prescribed psychotropic						
Yes	44 (10.6)	60 (14.4)	104 (12.5)	48 (11.6)	70 (16.8)	118 (14.2)
No[a]	367 (88.4)	303 (72.7)	670 (80.5)	366 (88.2)	347 (83.2)	713 (85.7)
Missing	4 (1.0)	54 (12.9)	58 (7.0)	1 (0.2)	0 (0.0)	1 (0.1)
Total	415 (100.0)	417 (100.0)	832 (100.0)	415 (100.0)	417 (100.0)	832 (100.0)
Type of psychotropic[b]						
ADHD drug	4 (9.1)	5 (8.3)	9 (8.7)	5 (8.2)	7 (9.1)	12 (8.7)
Anti anxiety drug	1 (2.3)	1 (1.7)	2 (1.9)	1 (1.6)	1 (1.3)	2 (1.4)
Antipsychotic drug	2 (4.5)	8 (13.3)	10 (9.6)	3 (4.9)	9 (11.7)	12 (8.7)
Antidepressant drug[c]	33 (75.0)	47 (78.3)	80 (76.9)	48 (78.7)	59 (76.6)	107 (77.5)
Sedative/sleep medication	5 (11.4)	13 (21.7)	18 (17.3)	8 (13.1)	18 (23.4)	26 (18.8)
Missing	5 (11.4)	5 (8.3)	10 (9.6)	3 (4.9)	4 (5.2)	7 (5.1)
Total	44 (100.0)	60 (100.0)	104 (100.0)	61 (100.0)	77 (100.0)	138 (100.0)

a Of those not on medications overall at baseline or during follow-up, six were missing data at baseline but not on medications during initial treatment (five FT, one TAU) and 53 had no baseline medications but were missing data during initial treatment (three FT, 50 TAU).
b Reported by clinician, CSO or additionally the researcher during follow-up (and baseline for overall use).
c Of the young people on antidepressants, only five were taking tricyclics (three FT, two FT), while the rest were taking selective serotonin reuptake inhibitors..

A total of 118 (14.2%) young people were taking or had been prescribed a psychotropic medication at baseline or during the 18-month trial period: 48 (11.6%) in the FT arm and 70 (16.8%) in the TAU arm.

Re-referrals to Child and Adolescent Mental Health Services

Re-referrals to CAMHS following discharge (or inactivity for 3 months) were reported from clinical notes within CAMHS by the participant's CAMHS clinician or a CSO, and also by the family during the 18-month researcher visit (*Table 38*).

Based on clinical data, 108 (13%) participants were re-referred to CAMHS within 18 months of randomisation, with similar rates between the trial arms; however, this information was missing for 19% of participants. Of all re-referred participants, 12 (11.1%) were re-referred on more than one occasion and the referrals of 83 (76.9%) were related to self-harm: 44 (84.5%) in the FT arm and 39 (69.6%) in the TAU arm.

Details of family-reported re-referrals were missing for over three-quarters of participants (because no researcher had visited or family had not been asked); however, re-referrals were reported for 39 (4.7%) participants.

TABLE 38 Re-referrals to CAMHS within 18 months of randomisation

	Method of reporting					
	Clinical data			Family reported		
Re-referrals	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person re-referred to CAMHS?, *n* (%)						
Yes	52 (12.5)	56 (13.4)	108 (13.0)	19 (4.6)	20 (4.8)	39 (4.7)
No	287 (69.2)	281 (67.4)	568 (68.3)	85 (20.5)	70 (16.8)	155 (18.6)
Missing	76 (18.3)	80 (19.2)	156 (18.8)	311 (74.9)	327 (78.4)	638 (76.7)
Number of re-referrals, *n* (%)						
1	46 (88.5)	50 (89.3)	96 (88.9)	18 (94.7)	16 (80.0)	34 (87.2)
2	6 (11.5)	5 (8.9)	11 (10.2)	1 (5.3)	3 (15.0)	4 (10.3)
3	0 (0.0)	1 (1.8)	1 (0.9)	0 (0.0)	1 (5.0)	1 (2.6)
Total	52 (100.0)	56 (100.0)	108 (100.0)	19 (100.0)	20 (100.0)	39 (100.0)
Any referral related to self-harm?, *n* (%)						
Yes	44 (84.6)	39 (69.6)	83 (76.9)	13 (68.4)	16 (80.0)	29 (74.4)
No	8 (15.4)	16 (28.6)	24 (22.2)	5 (26.3)	4 (20.0)	9 (23.1)
Missing	0 (0.0)	1 (1.8)	1 (0.9)	1 (5.3)	0 (0.0)	1 (2.6)
Total – participants	52 (100.0)	56 (100.0)	108 (100.0)	19 (100.0)	20 (100.0)	39 (100.0)
Timing of first re-referral (months)						
N	51	56	107	18	20	38
Missing	1	0	1	1	0	1
Mean (SD)	11.8 (3.79)	11.5 (3.91)	11.6 (3.84)	12.6 (4.13)	13.2 (3.64)	12.9 (3.84)
Median (range)	12.2 (4.2–17.7)	12.2 (0.8–17.9)	12.2 (0.8–17.9)	12.9 (4.2–17.4)	13.5 (4.4–17.6)	13.5 (4.2–17.6)

Referrals to psychiatric day or inpatient units

Referrals to psychiatric inpatient units within 18 months of randomisation were reported from clinical notes within CAMHS by participants' CAMHS clinician or CSO or via HES data, and also by the family during the 18-month researcher visit (*Table 39*).

Based on clinical data, a total of 25 (3%) participants were referred to a psychiatric inpatient unit during the 18-month follow-up period, of whom 21 (84.0%) had a referral related to self-harm, with similar rates between the trial arms. However, information was missing for 149 (17.9%) participants. The total length of inpatient stays over all referrals ranged from 1 to 185 days, with a mean of 72.2 days in the FT arm and 43.3 days in the TAU arm.

A total of seven referrals to psychiatric inpatient units were reported for five (0.6%) participants by the family, of which five had also been reported via the clinical data, with one referral in each arm reported by the family but not from the clinical data.

TABLE 39 Referrals to a psychiatric inpatient unit within 18 months of randomisation

Referrals to a psychiatric inpatient unit	Clinical data			Family reported[a]		
	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person referred to an inpatient unit?, *n* (%)						
Yes	13 (3.1)	12 (2.9)	25 (3.0)	3 (0.7)	2 (0.5)	5 (0.6)
No	331 (79.8)	327 (78.4)	658 (79.1)	104 (25.1)	84 (20.1)	188 (22.6)
Missing	71 (17.1)	78 (18.7)	149 (17.9)	308 (74.2)	331 (79.4)	639 (76.8)
Number of referrals, *n* (%)						
1	10 (76.9)	9 (75.0)	19 (76.0)	2 (66.7)	1 (50.0)	3 (60.0)
2	2 (15.4)	3 (25.0)	5 (20.0)	1 (33.3)	1 (50.0)	2 (40.0)
3	1 (7.7)	0 (0.0)	1 (4.0)	0 (0.0)	0 (0.0)	0 (0.0)
Total number referred	13 (100.0)	12 (100.0)	25 (100.0)	3 (100.0)	2 (100.0)	5 (100.0)
Any referral related to self-harm?, *n* (%)						
Yes	11 (84.6)	10 (83.3)	21 (84.0)	3 (100.0)	1 (50.0)	4 (80.0)
No	2 (15.4)	2 (16.7)	4 (16.0)	0 (0.0)	0 (0.0)	0 (0.0)
Missing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (50.0)	1 (20.0)
Total number referred	13 (100.0)	12 (100.0)	25 (100.0)	3 (100.0)	2 (100.0)	5 (100.0)
Length of stay (over all referrals, days)						
N	13	12	25	3	2	5
Missing	0	0	0	0	0	0
Mean (SD)	72.2 (49.74)	43.3 (46.89)	58.4 (49.62)	68.0 (89.44)	36.0 (28.28)	55.2 (67.13)
Median (range)	77.0 (11–185)	21.0 (1–139)	43.0 (1–185)	31.0 (3–170)	36.0 (16–56)	31.0 (3–170)
Timing of first referral (months)						
N	13	12	25	3	2	5
Missing	0	0	0	0	0	0
Mean (SD)	7.8 (5.76)	5.8 (4.56)	6.9 (5.21)	3.2 (4.37)	6.7 (8.99)	4.6 (5.78)
Median (range)	8.0 (0.8–17.8)	6.0 (0.3–12.8)	6.5 (0.3–17.8)	1.3 (0.1–8.2)	6.7 (0.3–13.0)	1.3 (0.1–13.0)

a Five of the seven referrals reported by the family were also reported in the clinical data and, conversely, one of the TAU referrals and one of the FT referrals were reported by the family at the researcher visit but not from the clinical data.

Referrals to other agencies

Referrals to adult mental health services or other agencies for the participating young person, their caregiver(s) or sibling(s) within 18 months of randomisation were reported from clinical notes within CAMHS by the participant's CAMHS clinician or a CSO, and also by the family during the 18-month researcher visit (*Table 40*).

Based on clinical data, a total of 18 (2.2%) young people or family members were referred to adult mental health services: eight young people only referred, nine family members only referred, and one young person and their family member referred. Information was missing for 155 (18.6%) young people. Referrals were reported within CAMHS notes only if the referral had come from CAMHS or if the young person had been in contact with CAMHS at the time. A higher proportion of adult mental health referrals were reported by the family: 25 (3%) participants (nine young people, 11 caregivers and five siblings).

TABLE 40 Referrals to adult mental health services and other agencies within 18 months of randomisation

Referrals to adult mental health services and other agencies, *n* (%)	Clinical data			Family reported		
	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Young person, caregiver or sibling referred to adult mental health services						
Yes	6 (1.4)	12 (2.9)	18 (2.2)	12 (2.9)	13 (3.1)	25 (3.0)
Young person	1 (0.2)	8 (1.9)	9 (1.1)[a]	4 (1.0)	5 (1.2)	9 (1.1)[b]
Parent/caregiver	5 (1.2)	5 (1.2)	10 (1.2)	5 (1.2)	6 (1.4)	11 (1.3)
Sibling	0 (0.0)	0 (0.0)	0 (0.0)	3 (0.7)	2 (0.5)	5 (0.6)
No	333 (80.2)	326 (78.2)	659 (79.2)	97 (23.4)	76 (18.2)	173 (20.8)
Missing	76 (18.3)	79 (18.9)	155 (18.6)	306 (73.7)	328 (78.7)	634 (76.2)
Young person, caregiver or sibling referred to other agency						
Yes	69 (16.6)	74 (17.7)	143 (17.2)	23 (5.5)	15 (3.6)	38 (4.6)
Young person	63 (15.2)	67 (16.1)	130 (15.6)	21 (5.1)	12 (2.9)	33 (4.0)
Parent/caregiver	20 (4.8)	18 (4.3)	38 (4.6)	6 (1.4)	4 (1.0)	10 (1.2)
Sibling	3 (0.7)	4 (1.0)	7 (0.8)	1 (0.2)	1 (0.2)	2 (0.2)
No	272 (65.5)	262 (62.8)	534 (64.2)	86 (20.7)	70 (16.8)	156 (18.8)
Missing	74 (17.8)	81 (19.4)	155 (18.6)	306 (73.7)	332 (79.6)	638 (76.7)
Agency referred to (non-mutually exclusive)						
Social services	32 (46.4)	36 (48.6)	68 (47.6)	13 (56.5)	5 (33.3)	18 (47.4)
Special Education Services	7 (10.1)	17 (23.0)	24 (16.8)	4 (17.4)	4 (26.7)	8 (21.1)
Voluntary sector	2 (2.9)	8 (10.8)	10 (7.0)	6 (26.1)	1 (6.7)	7 (18.4)
Youth offending team	4 (5.8)	2 (2.7)	6 (4.2)	0 (0.0)	0 (0.0)	0 (0.0)
Other[c]	58 (84.1)	60 (81.1)	118 (82.5)	13 (56.5)	9 (60.0)	22 (57.9)
Total	69 (100)	74 (100)	143 (100)	23 (100)	15 (100)	38 (100)

a All but one young person referral was related to self-harm.
b Referrals for five young persons were related to self-harm.
c Other agencies include, for example: support groups, counselling services, housing support, dermatologist, dietitian, eating disorder team, drugs and alcohol team, family support/relate/welfare, homeless shelter, Improving Access to Psychological Therapies, midwife paediatrician, pain clinic, parenting groups, physiotherapy, psychiatry, refuge, solicitor, youth theatre.

Conversely, referrals to other agencies were reported for 143 (17.2%) young people, caregivers or siblings via clinical data and for 38 (4.6%) young people, caregivers or siblings by the family. Based on clinical report, 68 referrals were to social services, 24 referrals were to special education services, 10 referrals were to the voluntary sector, six referrals were to youth offending team(s) and 118 referrals were to a wide range of other agencies or services.

Overall safety

A total of 1036 AEs were reported, defined as treatment on an emergency outpatient basis through an A&E attendance, attendance at a MIU or WIC and re-referral to CAMHS. A&E attendances constituted the largest portion of AEs (*Table 41*). AEs were reported in 443 (53.2%) participants, with a similar number of events and participants between the trial arms, and a mean of 1.2 AEs, ranging from 0 to 26 per participant.

TABLE 41 Adverse events and serious adverse events

AEs and SAEs	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Number of participants with one or more AE, *n* (%)	226 (54.5)	217 (52.0)	443 (53.2)
Number of participants with each type of AE (not mutually exclusive), *n* (%)			
A&E attendance	189 (45.5)	176 (42.2)	365 (43.9)
MIU/WIC attendance	33 (8.0)	40 (9.6)	73 (8.8)
Re-referral to CAMHS[a]	52 (12.5)	56 (13.4)	108 (13.0)
Total number of AEs reported	512	524	1036
A&E attendance	*409*	*372*	*781*
MIU/WIC attendance	*45*	*89*	*134*
Re-referral to CAMHS	*58*	*63*	*121*
Number of AEs per participant			
n	415	417	832
Mean (SD)	1.2 (2.00)	1.3 (2.25)	1.2 (2.13)
Median (range)	1.0 (0–17)	1.0 (0–26)	1.0 (0–26)
Number of participants with one or more SAE, *n* (%)	156 (37.6)	141 (33.8)	297 (35.7)
Total number of SAEs reported	275	323	598
Number of SAEs per participant			
n	415	417	832
Mean (SD)	0.7 (1.26)	0.8 (1.94)	0.7 (1.64)
Median (range)	0.0 (0–11)	0.0 (0–27)	0.0 (0–27)

a Re-referrals to CAMHS from clinical data only (i.e. excluding family reports).

No deaths were reported. All 598 SAEs reported for 297 (35.7%) participants were hospital admissions and readmissions. Again, similar rates were observed in both trial arms, although slightly more SAEs were reported in fewer participants in the TAU arm than in the FT arm: 323 SAEs in 141 (33.8%) in the TAU arm compared with 275 SAEs in 156 (37.6%) participants in the TAU arm.

Further details of the total number of hospital-related events – MIU/WIC attendances, A&E attendances and hospital admissions – are presented in *Table 42*. Overall, 1513 attendances were reported in 511 (61.4%) participants, with similar rates in both trial arms and a mean of 1.8 attendances per participant.

Over half of all attendances were for non-mental health reasons (59.7%), while 26.7% were for self-harm, 4.2% were for other mental health reasons and 3.4% resulted from a combination of mental health and non-mental health reasons (*Table 43*). A total of 387 (46.5%) participants attended hospital or a MIU or WIC at least once for non-mental health reasons; 223 (26.8%) participants attended for self-harm, 85 (10.2%) participants for other mental health reasons and in eight (1.0%) participants the reasons for all attendances were unknown.

TABLE 42 Number and details of all hospital, MIU and WIC attendances

Hospital, MIU and WIC attendances	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Number of participants with one or more hospital or MIU/WIC attendance, *n* (%)	258 (62.2)	253 (60.7)	511 (61.4)
Total number of attendances	729	784	1513
Number of attendances per participant			
n	415	417	832
Mean (SD)	1.8 (2.63)	1.9 (3.22)	1.8 (2.94)
Median (range)	1.0 (0–23)	1.0 (0–30)	1.0 (0–30)
Outcome of presentation, *n* (%)			
Discharged	454 (62.3)	461 (58.8)	915 (60.5)
Admitted to hospital ward	275 (37.7)	323 (41.2)	598 (39.5)
Total attendances	729 (100.0)	784 (100.0)	1513 (100.0)
Discharge type, *n* (%)			
Self-discharge	19 (4.2)	24 (5.2)	43 (4.7)
Did not require any follow-up treatment	172 (37.9)	195 (42.3)	367 (40.1)
With referral to GP or outpatient follow-up	250 (55.1)	232 (50.3)	482 (52.7)
Transferred to other health-care professional/provider	9 (2.0)	4 (0.9)	13 (1.4)
Other discharge	2 (0.4)	4 (0.9)	6 (0.7)
Missing	2 (0.4)	2 (0.4)	4 (0.4)
Total A&E/MIU/WIC attendances	454 (100.0)	461 (100.0)	915 (100.0)
Admitted to hospital ward, *n* (%)			
Paediatric ward	66 (24.0)	61 (18.9)	127 (21.2)
Assessment unit	7 (2.5)	5 (1.5)	12 (2.0)
Adult ward	5 (1.8)	3 (0.9)	8 (1.3)
ICU/HDU/other specialist unit	6 (2.2)	4 (1.2)	10 (1.7)
Psychiatric inpatient stay	14 (5.1)	11 (3.4)	25 (4.2)
Missing	177 (64.4)	239 (74.0)	416 (69.6)
Total admissions	275 (100.0)	323 (100.0)	598 (100.0)
Length of admission (days)			
n	261	313	574
Missing	14	10	24
Mean (SD)	6.2 (29.45)	2.2 (9.15)	4.0 (21.05)
Median (range)	1.0 (0.0–371.0)	1.0 (0.0–114.0)	1.0 (0.0–371.0)

TABLE 43 Reason for all attendances, overall and per participant

Reason for attendance, *n* (%)	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Per attendance			
Self-harm	207 (28.4)	197 (25.1)	404 (26.7)
Other mental health reason	31 (4.3)	32 (4.1)	63 (4.2)
Non-mental health reason	433 (59.4)	470 (59.9)	903 (59.7)
Other mental health and non-mental health reason	22 (3.0)	29 (3.7)	51 (3.4)
Missing[a]	36 (4.9)	56 (7.1)	92 (6.1)
Total attendances	729 (100.0)	784 (100.0)	1513 (100.0)
Per participant (not mutually exclusive)			
Self-harm[b]	119 (28.7)	104 (24.9)	223 (26.8)
Other mental health reason	44 (10.6)	41 (9.8)	85 (10.2)
Non-mental health reason	191 (46.0)	196 (47.0)	387 (46.5)
Missing reason for all attendances	2 (0.5)	6 (1.4)	8 (1.0)
No attendances reported	157 (37.8)	164 (39.3)	321 (38.6)
Total participants	415 (100.0)	417 (100.0)	832 (100.0)

a Ninety-eight attendances were classed as unknown in the primary outcome sensitivity analysis for missing presenting details. This is because six attendances for mental health reasons were missing sufficient further information to rule out self-harm.
b MIU and WIC attendances are included, hence the number of participants with an attendance caused by self-harm is higher than the primary outcome.

Primary end point

Hospital follow-up

Following the interim review of primary outcome results, it was agreed that the stopping rule had not been reached and the DMEC recommended that the trial continue.

Table 44 details the length of follow-up for hospital attendance for all participants.

A final HES data download was obtained from NHS Digital in May 2015, providing HES data to 31 January 2015. The final download provided full data for the majority of participants: 724 (87.0%) recruited prior to 31 July 2013 for whom HES data covered the entire 18 month follow-up period. A further 75 participants had partial HES coverage during their follow-up period: 71 (8.5%) recruited after 31 July 2013 for whom HES data covered at least the first 13 months' follow-up, and four (0.5%) for whom linkage to HES had been unsuccessful for at least one prior NHS Digital download resulting in incomplete HES data from randomisation until their first successful linkage (1.3, 3.8, 11.1 and 14.3 months).

For the 75 participants with partial HES coverage and a further 12 (1.4%) participants for whom linkage to HES was not made during any data download, hospital attendance data were obtained via researcher follow-up at acute trusts within participants' recruiting hub. A further 21 (2.5%) participants withdrew from further clinical data collection during the trial, thus preventing further linkage to HES data or researcher follow-up.

TABLE 44 Duration of hospital follow-up for all participants

Duration of hospital follow-up	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Follow-up description, *n* (%)			
Full HES linkage and follow-up	362 (87.2)	362 (86.8)	724 (87.0)
Linked to HES but researcher follow-up only from February 2015 onwards	35 (8.4)	36 (8.6)	71 (8.5)
Linked to HES data after randomisation	2 (0.5)	2 (0.5)	4 (0.5)
Not linked to HES data: researcher follow-up only	8 (1.9)	4 (1.0)	12 (1.4)
Withdrawal	8 (1.9)	13 (3.1)	21 (2.5)
Length of hospital follow-up (months), *n* (%)			
18	398 (95.9)	397 (95.2)	795 (95.6)
< 18	15 (3.6)	18 (4.3)	33 (4.0)
< 3	0 (0.0)	1 (0.2)	1 (0.1)
≤ 3 and < 6	1 (0.2)	1 (0.2)	2 (0.2)
≤ 6 and < 9	1 (0.2)	2 (0.5)	3 (0.4)
≤ 9 and < 12	0 (0.0)	2 (0.5)	2 (0.2)
≤ 12 and < 15	5 (1.2)	5 (1.2)	10 (1.2)
≤ 15 and < 18	8 (1.9)	7 (1.7)	15 (1.8)
No follow-up	2 (0.5)	2 (0.5)	4 (0.5)
Length of hospital follow-up (months)			
N	415	417	832
Mean (SD)	17.8 (1.59)	17.7 (1.93)	17.7 (1.77)
Median (range)	18.0 (0.0–18.0)	18.0 (0.0–18.0)	18.0 (0.0–18.0)

Note
Of the 21 young persons who withdrew from clinical data collection, four had no hospital follow-up, three had < 6 months' follow-up, five had < 12 months' follow-up, eight had < 18 months' follow-up and one had the full 18 months' follow-up.

In summary, a total of 795 (95.6%) participants had complete follow-up over their entire 18 months post randomisation. Four (0.5%) participants were completely lost to follow-up and 33 (4%) participants had at least some but not complete 18-month hospital follow-up.

The duration of follow-up for all participants can be seen to be evenly balanced between the trial arms (see *Table 44*).

Summary of primary outcome events

A total of 221 (26.6%) young persons experienced the primary outcome event, that is, having a repeat self-harm event leading to hospital attendance within 18 months post randomisation: 118 (28.4%) young persons randomised to receive FT and 103 (24.7%) young persons randomised to receive TAU.

Table 45 details the type and outcome of all primary outcome events. The overall median time to first post-randomisation self-harm event is 5.2 months. This is earlier for young persons randomised to TAU (4.3 months, compared with 5.7 months in the case of young persons randomised to FT).

The most common reason for attendance was self-poisoning (57.5% of primary outcome events), with a further 19.9% attending as a result of cutting. The most common outcome was admission to a hospital ward, for 161 (72.9%) events, with similar rates of admission between trial arms.

TABLE 45 Timing, type and outcome of first repeat self-harm events

Timing, type and outcome of first repeat self-harm events	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Number of participants with one or more event, *n* (%)	118 (28.4)	103 (24.7)	221 (26.6)
Time to first self-harm event (months)			
n	118	103	221
Mean (SD)	6.9 (5.47)	6.2 (5.25)	6.6 (5.37)
Median (range)	5.7 (0.1–17.8)	4.3 (0.0–16.8)	5.2 (0.0–17.8)
Type of first self-harm event, *n* (%)			
Poisoning	68 (57.6)	59 (57.3)	127 (57.5)
Cutting	28 (23.7)	16 (15.5)	44 (19.9)
Poisoning and cutting	8 (6.8)	8 (7.8)	16 (7.2)
Other self-injury	6 (5.1)	10 (9.7)	16 (7.2)
Other violent method	1 (0.8)	4 (3.9)	5 (2.3)
Other	6 (5.1)	6 (5.8)	12 (5.4)
Missing	1 (0.8)	0 (0.0)	1 (0.5)
Total	118 (100)	103 (100)	221 (100)
Outcome of first self-harm event, *n* (%)			
Discharged from A&E	32 (27.1)	28 (27.2)	60 (27.1)
Admitted to hospital ward	86 (72.9)	75 (72.8)	161 (72.9)
Total	118 (100)	103 (100)	221 (100)
Treatment received, *n* (%)			
Yes	106 (89.8)	97 (94.2)	203 (91.9)
No	10 (8.5)	6 (5.8)	16 (7.2)
Missing	2 (1.7)	0 (0.0)	2 (0.9)
Total	118 (100)	103 (100)	221 (100)
Level of treatment received, *n* (%)			
Minimal	53 (50.0)	60 (61.9)	113 (55.7)
Significant intervention	23 (21.7)	25 (25.8)	48 (23.6)
Missing	30 (28.3)	12 (12.4)	42 (20.7)
Total	106 (100)	97 (100)	203 (100)

For participants with a primary outcome event, *Table 46* presents the type of event by the method used in the index event prompting referral into CAMHS. Among those participants whose index episode was self-poisoning, the first repeat event was also self-poisoning in the majority (38, 73.1%), whereas 11 (21.2%) self-injured and three (5.8%) used combined methods. Of those who used combined methods at the index event, similar proportions were observed, with the majority self-poisoning at first repeat (16, 66.7%). Of those in whom the index event was self-injury, a greater proportion switched method, with only 59 (40.7%) self-injuring at first repeat, 73 (50.3%) self-poisoning and 12 (8.3%) using combined methods.

Table 47 presents the number of participants with one or more event, the number without an event and censored in the analysis, the reasons for censoring and the time to censoring. The majority of censored participants (*n* = 583, 95.4%) were censored because no observed self-harm events resulting in hospital

TABLE 46 Type of primary outcome event by index method

Index event	Primary outcome event, *n* (%)			
	Self-poisoning	Self-injury	Combined	Missing
Self-poisoning (*N* = 52)	38 (73.1)	11 (21.2)	3 (5.8)	0 (0.0)
Self-injury (*N* = 145)	73 (50.3)	59 (40.7)	12 (8.3)	1 (0.7)
Combined (*N* = 24)	16 (66.7)	7 (29.2)	1 (4.2)	0 (0.0)
Total (*N* = 221)	127 (57.5)	77 (34.8)	16 (7.2)	1 (0.5)

TABLE 47 Summary of censoring for participants without an event: reasons and time to censoring

Censored participants	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Number of censored participants, *n* (%)			
Not censored (one or more event)	118 (28.4)	103 (24.7)	221 (26.6)
Censored	297 (71.6)	314 (75.3)	611 (73.4)
Total	415 (100)	417 (100)	832 (100)
Reason for censoring, *n* (%)			
No event in 18-month follow-up period	283 (95.3)	300 (95.5)	583 (95.4)
Not linked to HES data (and without complete 18-month researcher follow-up)	1 (0.3)	0 (0.0)	1 (0.2)
Participants still in follow-up after the final HES data set (without further researcher follow-up)	6 (2.0)	6 (1.9)	12 (2.0)
Withdrawal	7 (2.4)	8 (2.5)	15 (2.5)
Total	297 (100)	314 (100)	611 (100)
Time from randomisation to censoring (months), *n* (%)			
0	2 (0.7)	2 (0.6)	4 (0.7)
< 3	0 (0.0)	1 (0.3)	1 (0.2)
≤ 3 and < 6	1 (0.3)	0 (0.0)	1 (0.2)
≤ 6 and < 9	1 (0.3)	1 (0.3)	2 (0.3)
≤ 9 and < 12	0 (0.0)	1 (0.3)	1 (0.2)
≤ 12 and < 15	4 (1.3)	2 (0.6)	6 (1.0)
≤ 15 and < 18	6 (2.0)	7 (2.2)	13 (2.1)
18	283 (95.3)	300 (95.5)	583 (95.4)
Total	297 (100)	314 (100)	611 (100)
Months from randomisation to censoring			
N	297	314	611
Mean (SD)	17.7 (1.86)	17.7 (1.90)	17.7 (1.88)
Median	18.0 (0.0–18.0)	18.0 (0.0–18.0)	18.0 (0.0–18.0)

attendance occurred over the entire 18-month follow-up period; however, 15 (2.5%) participants were censored because of their withdrawal from further clinical data collection, as no events were observed prior to their withdrawal; and a further 13 (2.2%) participants were censored either because they were not linked to HES data and relied on researcher follow-up alone or because their HES data did not include their full 18-month follow-up period; in both cases researcher follow-up was available only for the periods not covered via HES data. Similar patterns of censoring are observed between trial arms.

Primary analysis: Cox proportional hazards regression

The results of the Cox proportional hazards regression modelling, adjusted for covariates, are displayed in *Table 48*. A HR of > 1 indicates an increased rate of self-harm in the variable listed before 'vs.' in the table (e.g. participants with three or more previous self-harm episodes repeat self-harm at 1.22 times the rate of participants with just two previous self-harm episodes).

TABLE 48 Primary analysis: Cox proportional hazards model (adjusted for covariates)

Covariate	Parameter estimate	SE	HR	95% CI	df	*p*-value
Treatment: FT (vs. TAU)	0.13	0.14	1.14	0.87 to 1.49	1	0.3349[a]
Gender: female (vs. male)	0.47	0.25	1.60	0.98 to 2.61	1	0.0589
Age group: 15–17 (vs. 11–14)	–0.36	0.14	0.70	0.53 to 0.92	1	0.0106
Number of previous self-harm episodes: ≥ 3 (vs. 2)	0.20	0.23	1.22	0.78 to 1.92	1	0.3900
Type of index episode					2	0.0333
Combined (vs. self-injury)	0.61	0.24	1.83	1.14 to 2.96	1	–
Self-poisoning (vs. self-injury)	0.03	0.20	1.03	0.69 to 1.54	1	–
Referred via hospital: yes (vs. no)	0.27	0.18	1.31	0.93 to 1.86	1	0.1231
Trust (vs. Y trust 1)					14	0.0938
M trust 3	0.18	0.23	1.20	0.77 to 1.86	1	–
L trust 5	–0.58	0.44	0.56	0.24 to 1.34	1	–
Y trust 3	–0.66	0.37	0.52	0.25 to 1.07	1	–
Y trust 2	–0.22	0.37	0.80	0.39 to 1.67	1	–
M trust 1	0.13	0.28	1.14	0.65 to 1.98	1	–
L trust 6	–0.30	0.73	0.74	0.18 to 3.08	1	–
Y trust 4	–1.18	0.53	0.31	0.11 to 0.87	1	–
L trust 2	–0.70	0.32	0.50	0.26 to 0.93	1	–
M trust 2	–0.23	0.25	0.79	0.48 to 1.30	1	–
L trust 1	–0.35	0.26	0.70	0.42 to 1.17	1	–
Y trust 6	0.19	0.33	1.21	0.64 to 2.29	1	–
L trust 4	–0.59	1.02	0.55	0.08 to 4.08	1	–
L trust 3	–0.29	0.53	0.75	0.26 to 2.11	1	–
Y trust 5	–0.23	0.41	0.80	0.35 to 1.79	1	–

df, degrees of freedom.
a Reduction in –2logL between Cox proportional hazards model with and without treatment effect: 0.932 ($p = 0.3344$, 1 degree of freedom).

There was no evidence to suggest a statistically significant difference in self-harm repetition rates between treatment groups. The HR for FT compared with TAU was 1.14 (95% CI 0.87 to 1.49; $p = 0.3349$). *Figure 17* presents the Kaplan–Meier curve for time to self-harm by randomised treatment arm. There is crossover between treatment groups until around month 11, after which the curves diverge and the lower curve represents slightly more FT participants with a primary outcome self-harm event. Further estimates of the self-harm repetition rate, with 95% CIs, at 1, 3, 6, 12 and 18 months post randomisation, are presented in *Table 49* by arm and for the difference between arms, with a positive difference indicating a higher repetition rate in the FT arm. As there were few events, it was not possible to obtain the Kaplan–Meier estimates of the median time to self-harm.

Covariates age group ($p = 0.0106$) and the type of index self-harm episode ($p = 0.0333$) were found to have a significant effect on the risk of self-harm at the 5% level. The rate of self-harm was lower in older participants (aged 15–17 years) than in those aged 11–14 years (HR 0.7, 95% CI 0.53 to 0.92) and higher in participants whose index episode combined self-injury and poisoning than in those with self-injury only (HR 1.83, 95% CI 1.14 to 2.96). There was also weak evidence (at the 10% level) of an effect for gender and trust, with a higher rate of self-harm in females than in males (HR 1.6, 95% CI 0.98 to 2.61; $p = 0.0589$)

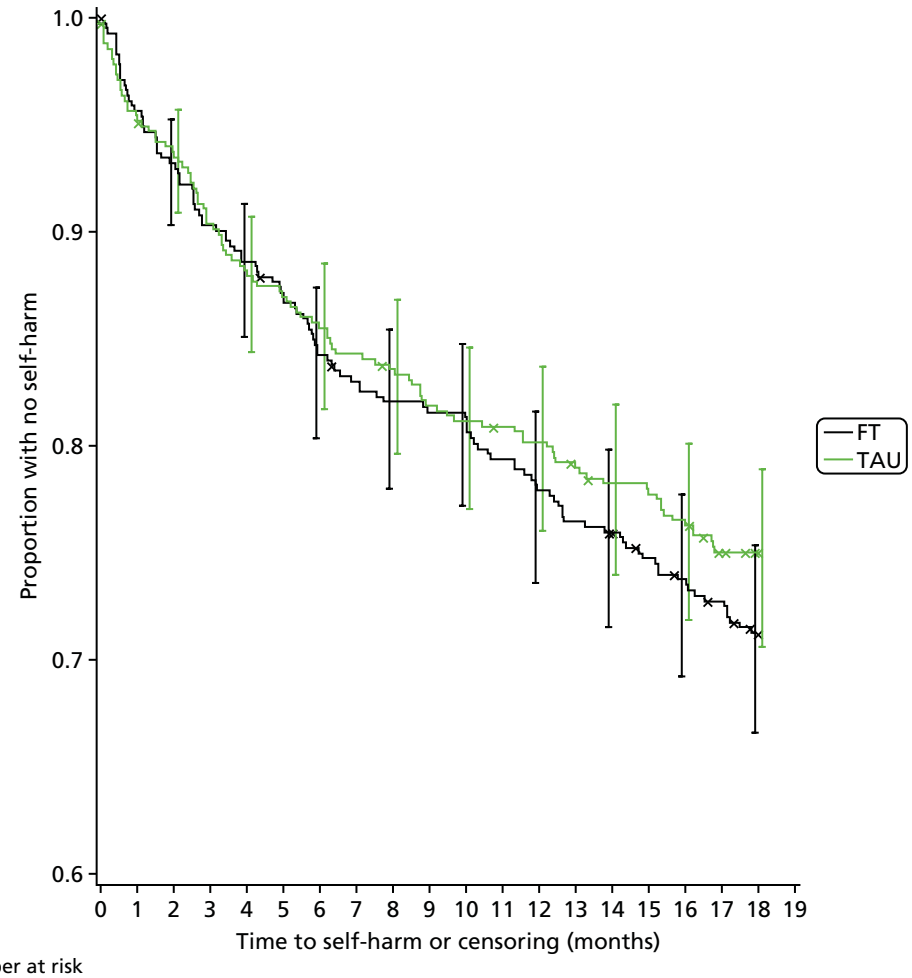


FIGURE 17 Kaplan–Meier plot of time to self-harm by randomised treatment group with 95% CIs. Reproduced from Cottrell *et al.*[70] © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

TABLE 49 Self-harm repetition rates during follow-up

Time from randomisation	FT (*N* = 415)	TAU (*N* = 417)	Difference (FT – TAU) (*N* = 832)
Month 1			
Repetition estimate (95% CI)	4.36% (2.39% to 6.33%)	4.82% (2.76% to 6.88%)	–0.46% (–3.31% to 2.39%)
Number with an event	18	20	–2
Number event free	395	395	0
Month 3			
Repetition estimate (95% CI)	9.69% (6.83% to 12.54%)	9.65% (6.81% to 12.49%)	0.04% (–3.99% to 4.06%)
Number with an event	40	40	0
Number event free	373	374	–1
Month 6			
Repetition estimate (95% CI)	15.75% (12.23% to 19.26%)	14.48% (11.09% to 17.87%)	1.27% (–3.61% to 6.15%)
Number with an event	65	60	5
Number event free	347	354	–7
Month 12			
Repetition estimate (95% CI)	22.08% (18.07% to 26.08%)	19.81% (15.97% to 23.65%)	2.27% (–3.28% to 7.82%)
Number with an event	91	82	9
Number event free	320	330	–10
Month 18			
Repetition estimate (95% CI)	28.74% (24.36% to 33.12%)	24.94% (20.77% to 29.12%)	3.80% (–2.26% to 9.85%)
Number with an event	118	103	15
Number event free	283	300	–17

and substantial variation across centres ($p = 0.0938$). The occurrence of primary outcome events, overall and within each treatment arm, for each covariate is further summarised in *Tables 50* and *51*.

In addition to the primary analysis model, further analysis excluding centre as a fixed effect (see model checking, *Table 52*) finds an increased risk of self-harm in participants referred to CAMHS from hospital, and summary statistics suggest a differential treatment effect based on participants' source of referral (see *Table 50*). The rate of self-harm among participants referred via hospital was higher in the FT arm than in the TAU arm (37.8% vs. 25.7%), an absolute increase in risk of 12.1% and a relative increase of 47.3%

TABLE 50 Summary of self-harm event by covariates (and hub) and treatment arm

Covariate	Primary outcome event reported (yes), *n* (%)		
	FT	TAU	Total
Age group (years)			
11–14 (*N* = 441)	66 (30.0)	62 (28.1)	128 (29.0)
15–17 (*N* = 391)	52 (26.7)	41 (20.9)	93 (23.8)
Gender			
Male (*N* = 95)	11 (23.4)	7 (14.6)	18 (18.9)
Female (*N* = 737)	107 (29.1)	96 (26.0)	203 (27.5)
Baseline number of self-harm episodes			
2 (*N* = 93)	10 (21.7)	12 (25.5)	22 (23.7)
≥ 3 (*N* = 739)	108 (29.3)	91 (24.6)	199 (26.9)
Type of index self-harm episode			
Self-poisoning (*N* = 184)	32 (34.4)	20 (22.0)	52 (28.3)
Self-injury (*N* = 594)	74 (24.9)	71 (23.9)	145 (24.4)
Combined (*N* = 54)	12 (48.0)	12 (41.4)	24 (44.4)
Referred into CAMHS via A&E?			
Yes (*N* = 304)	59 (37.8)	38 (25.7)	97 (31.9)
No (*N* = 528)	59 (22.8)	65 (24.2)	124 (23.5)
Centre hub (non-covariate)			
Yorkshire (*N* = 300)	42 (28.2)	37 (24.5)	79 (26.3)
Manchester (*N* = 286)	49 (34.3)	42 (29.4)	91 (31.8)
London (*N* = 246)	27 (22.0)	24 (19.5)	51 (20.7)
Total (*N* = 832)	118 (28.4)	103 (24.7)	221 (26.6)

(see *Table 46*), while there was little difference between groups in the case of participants referred to CAMHS through the community, 22.8% in the FT arm compared with 24.2% in the TAU arm, an absolute decrease in risk of 1.4% and a relative decrease of 5.7%. The presence of a moderating effect by source of referral was investigated in the moderator analysis; however, although there was an indication of an increased hazard in the FT arm in participants referred via hospital, the interaction between treatment and referral source was not found to be statistically significant (see *Moderator variables*).

Model checking

Graphical and statistical tests of the adequacy of the Cox proportional hazards regression model for treatment and covariates were generally satisfactory and are presented in *Appendix 5*, along with Kaplan–Meier curves for time to self-harm for each covariate.

There was, however, evidence of non-proportional hazards for a number of trusts, and further Cox proportional hazards regression models were fitted to investigate the effect on the treatment estimate

TABLE 51 Summary of self-harm events by recruiting trust (covariate) and treatment arm[a]

	Primary outcome event reported (yes), *n* (%)		
Centre trust	FT	TAU	Total
Yorkshire			
Y trust 1 (*N* = 114)	22 (37.9)	15 (26.8)	37 (32.5)
Y trust 2 (*N* = 34)	5 (27.8)	4 (25.0)	9 (26.5)
Y trust 3 (*N* = 52)	3 (11.1)	6 (24.0)	9 (17.3)
Y trust 4 (*N* = 37)	1 (5.9)	3 (15.0)	4 (10.8)
Y trust 5 (*N* = 29)	5 (35.7)	2 (13.3)	7 (24.1)
Y trust 6 (*N* = 34)	6 (40.0)	7 (36.8)	13 (38.2)
Manchester			
M trust 1 (*N* = 53)	11 (40.7)	8 (30.8)	19 (35.8)
M trust 2 (*N* = 105)	15 (28.3)	13 (25.0)	28 (26.7)
M trust 3 (*N* = 128)	23 (36.5)	21 (32.3)	44 (34.4)
London			
L trust 1 (*N* = 107)	16 (29.6)	9 (17.0)	25 (23.4)
L trust 2 (*N* = 73)	8 (21.6)	5 (13.9)	13 (17.8)
L trust 3 (*N* = 16)	1 (11.1)	3 (42.9)	4 (25.0)
L trust 4 (*N* = 6)	0 (0.0)	1 (33.3)	1 (16.7)
L trust 5 (*N* = 34)	2 (11.8)	4 (23.5)	6 (17.6)
L trust 6 (*N* = 10)	0 (0.0)	2 (28.6)	2 (20.0)
Total (*N* = 832)	118 (28.4)	103 (24.7)	221 (26.6)

a Shading is used to represent trusts with the same family team.

(*Table 52*). In all three alternative models, the estimated treatment effect matched very closely with that of the primary analysis, confirming the robustness of the treatment effect to variations in modelling and assumptions around trust.

Unlike the primary analysis model, there was evidence of an increased risk of self-harm in participants referred into CAMHS via hospital, with estimates from the frailty model suggesting a HR of 1.39 (95% CI 0.99 to 1.95) with $p = 0.0598$. A chi-squared test detected a highly significant ($p < 0.0001$) association between trust and referral source (see *Appendix 5*); thus, referral source is not identified as a significant effect in the primary model because of a lack of independence and collinearity with trust, with each providing 'overlapping' information.

Differences were also found for gender, which was no longer statistically significant when centre was excluded, with only weak evidence of an effect in the model, including hub and the shared frailty model. There was no evidence of an overall association between trust and gender (see *Appendix 5*); however, in the two trusts with the lowest risk of self-harm the case mix also included the highest proportion of females. Thus, in excluding centre, the reduced self-harm observed in these trusts is no longer explained by trust and is instead incorporated into the other covariates for these participants, including gender, thus diluting the effect of gender overall.

TABLE 52 Cox proportional hazards model: excluding centre, including hub, with random centre effects

	Primary model (excluding centre)					Primary model (including hub)					Primary model (with random centre effect)				
Covariate	Parameter estimate	SE	HR	95% CI	*p*-value	Parameter estimate	Standard error	HR	95% CI	*p*-value	Parameter estimate	Standard error	HR	95% CI	*p*-value
Treatment: FT (vs. TAU)	0.13	0.14	1.14	0.88 to 1.49	0.3214	0.13	0.14	1.14	0.87 to 1.48	0.3382	0.13	0.14	1.14	0.87 to 1.48	0.3343
Gender: female (vs. male)	0.38	0.25	1.47	0.90 to 2.38	0.1224	0.42	0.25	1.52	0.93 to 2.47	0.0925	0.43	0.25	1.54	0.94 to 2.50	0.0835
Age group (years): 15–17 (vs. 11–14)	–0.33	0.14	0.72	0.54 to 0.94	0.0171	–0.35	0.14	0.71	0.54 to 0.93	0.0133	–0.35	0.14	0.70	0.53 to 0.93	0.0123
Number of previous self-harm episodes: ≥ 3 (vs. 2)	0.20	0.23	1.22	0.78 to 1.91	0.3916	0.20	0.23	1.22	0.78 to 1.91	0.3944	0.20	0.23	1.22	0.78 to 1.92	0.3818
Type of index episode					0.0208					0.0173					0.0232
Combined (vs. self-injury)	0.60	0.24	1.83	1.14 to 2.92		0.63	0.24	1.87	1.17 to 2.99		0.62	0.24	1.85	1.15 to 2.98	
Self-poisoning (vs. self-injury)	–0.04	0.20	0.96	0.65 to 1.42		–0.03	0.20	0.98	0.66 to 1.44		0.00	0.20	1.00	0.67 to 1.49	
Referred via hospital: yes (vs. no)	0.38	0.17	1.47	1.05 to 2.06	0.0265	0.34	0.17	1.41	1.01 to 1.98	0.0457	0.33	0.17	1.39	0.99 to 1.95	0.0598
Recruiting hub										0.0169					
London (vs. Yorkshire)						–0.26	0.18	0.77	0.54 to 1.10						
Manchester (vs. Yorkshire)						0.24	0.15	1.27	0.94 to 1.72						
Trust (random effect)											0.06	0.05			0.0646

Figure 18 presents the estimated random centre effect in the primary Cox proportional hazards model with shared frailty (random centre effect). The impact of these fluctuations across trusts does not, however, have any bearing on the estimated treatment effect. The model including trust is, therefore, retained for the primary analysis purposes; however, although referral via hospital does not appear to be statistically significant, this is an important factor in the repetition of self-harm and retained in the model.

Sensitivity analysis for missing presenting details

Summary of non-primary outcome self-harm hospital events and events missing presenting details

Hospital attendances reported by NHS Digital but for which it could not be determined whether or not the reason was self-harm and which remained unclassified at the time of analysis are summarised and the potential impact of these episodes on the overall event rate is explored. The number of self-harm events leading to a MIU or WIC attendance, rather than hospital attendance, is also explored, as these events were not included in the definition of the primary outcome.

Only three MIU or WIC self-harm attendances were reported in three participants within 18 months of randomisation (*Table 53*). One of these three participants attended hospital attendance for self-harm before their MIU attendance, and hence there was no change to their primary outcome or to its timing, while no other self-harm hospital events were reported for the other two participants (one from each treatment arm) and, therefore, there was no primary outcome event. The impact of the exclusion of MIU and WIC attendances from the primary outcome is, therefore, minimal.

Table 54 presents, by randomised treatment group, the overall number of hospital attendances within 18 months of randomisation for which it could not be determined whether or not the reason was self-harm. There were a total of 47 unclassified hospital attendances in 41 participants: 26 episodes in 22 participants randomised to receive FT and 21 episodes in 19 participants randomised to receive TAU. There were also a further 12 unclassified MIU attendances in 10 participants and 39 unclassified WIC attendances

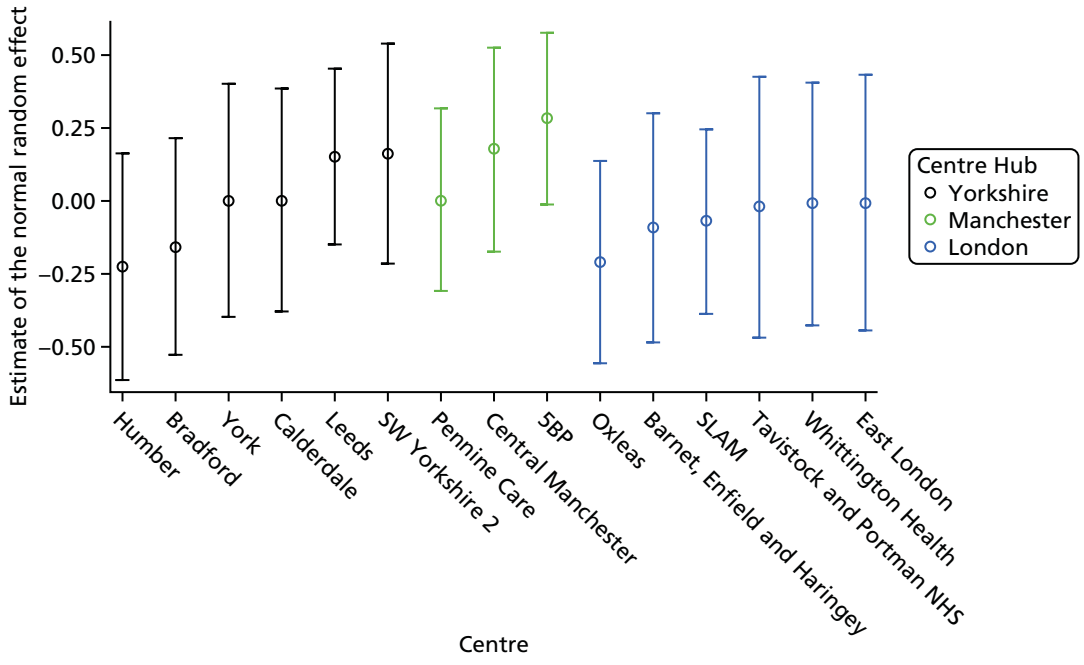


FIGURE 18 Estimated random centre effect in the primary Cox proportional hazards model with shared frailty (random centre effect). Estimated random effects > 0 represent higher frailty, that is, shorter time to and higher rate of self-harm, while estimates < 0 represent lower frailty, that is, longer time to and lower rate of self-harm.

TABLE 53 Minor injury unit or WIC attendances caused by self-harm

MIU or WIC attendances	FT (*n*)	TAU (*n*)	Total (*n*)
Number of participants with a MIU or WIC self-harm event			
MIU only	1	2	3
Other additional hospital self-harm events?			
Hospital self-harm event before MIU attendance for self-harm	0	1	1
No other self-harm events	1	1	2

TABLE 54 Unclassified hospital, MIU and WIC attendances

Unclassified attendances	FT (*n*)	TAU (*n*)	Total (*n*)
Number of unclassified attendances			
Hospital attendance	26	21	47[a]
MIU only	3	9	12
WIC only	12	27	39
Number of participants with an unclassified attendance			
Hospital attendance	22	19	41
MIU only	3	7	10
WIC only	7	15	22
Number of participants with an unclassified hospital attendance			
With a confirmed primary outcome event (*unclassified attendance prior to primary outcome event*)	12 (4)	11 (5)	23 (9)
Without a confirmed primary outcome event	10	8	18

a Of the 47 unclassified hospital attendances, 41 were reported by HES only, 5 from HES and the researcher and 1 by the researcher alone.

in 22 participants; however, the impact of these events is not explored further as they did not meet the definition of the primary outcome.

To assess the potential impact of unclassified hospital attendances on the overall self-harm event rate, *Table 54* further presents these in relation to whether or not the participant had already contributed to the primary outcome event rate. Of the 41 participants with at least one unclassified hospital episode, only 18 had no self-harm-related hospital attendances and thus had the potential to increase the event rate. Additionally, for 9 of the remaining 23 participants the unclassified attendance occurred prior to the known self-harm-related hospital attendance and, thus, had the potential to impact on the timing of the primary outcome. Thus, in summary, a total of 27 participants (14 in the FT arm and 13 in the TAU arm) had the potential to influence either the primary event rate or the timing of primary events.

Sensitivity analysis

Assuming that hospital attendances that were unclassified at the time of analysis were self-harm related, the recalculated HR for treatment was 1.15 (95% CI 0.89 to 1.49) with $p = 0.2736$ (see *Appendix 5*). Hence, as per the primary first event model, there is no evidence to suggest a difference in self-harm repetition rates between treatment groups and minimal difference for covariates in the model.

Sensitivity analysis for clustering

Clustering because of therapists

A total of 286 different therapists were identified as the 'main' therapist delivering the greatest number of sessions (comprising assessment and therapy sessions) within CAMHS per participant (*Table 55* and see *Figure 19*).

In the FT arm, 47 main therapists delivered treatment to 398 participants; 29 SHIFT family therapists led therapy for 380 participants, with the remaining 18 participants receiving other treatment within CAMHS (i.e. as part of a referral) led by 18 TAU CAMHS clinicians (10 of whom were also the main therapist for at least one TAU participant).

In the TAU arm, many more therapists were involved in delivering treatment, with a total of 250 main therapists across 333 participants (one of which was a SHIFT FT therapist prior to their involvement in the trial).

TABLE 55 Main therapist per participant and number of participants seen per therapist

Main therapist	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Type of main therapist per participant, *n* (%)			
SHIFT FT therapist	380 (91.6)	1 (0.2)[a]	381 (45.8)
TAU clinician	18 (4.3)	332 (79.6)	350 (42.1)
Clinician name missing	1 (0.2)	6 (1.4)	7 (0.8)
N/A – young person attended no sessions	16 (3.9)	33 (7.9)	49 (5.9)
Unknown – young person missing all treatment data	0 (0.0)	45 (10.8)	45 (5.4)
Overall number of different main therapists, *n* (%)			
SHIFT FT therapist	29 (61.7)	1 (0.4)	29 (10.1)
TAU clinician	18 (38.3)	249 (99.6)	257 (89.9)
Total	47 (100.0)	250 (100.0)	286 (100.0)
Number of participants per main therapist			
n	47	250	286
Mean (SD)	8.5 (10.19)	1.3 (0.81)	2.6 (4.93)
Median (range)	5.0 (1–46)	1.0 (1–6)	1.0 (1–46)
Number of participants per main therapist, *n* (%)			
1	21 (44.7)	198 (79.2)	202 (70.6)
≥ 2	26 (55.3)	52 (20.8)	84 (29.4)
Overall number of sessions delivered per main therapist			
n	47	250	286
Mean (SD)	54.7 (53.97)	9.5 (10.83)	17.3 (29.25)
Median (range)	37.0 (1–235)	6.0 (1–96)	7.0 (1–235)

N/A, not applicable.
a One TAU participant was treated by a SHIFT FT therapist before they were trained in SHIFT FT and before they became involved in the trial.

For seven participants the name of the clinician delivering treatment was missing. In addition, 49 participants were known to have been represented at no sessions (i.e. neither the participant nor family member attended) and for 45 participants all treatment data were missing. These participants were each classed as belonging to their own cluster, resulting in a total of 387 clusters, 286 main therapist clusters and 101 individual participant clusters (comprising the seven participants with no clinician name, the 49 participants who attended no sessions and the 45 participants for whom all treatment data were missing).

Table 56 presents the Cox proportional hazards model, adjusted for covariates, with shared frailty for therapist effect and *Figure 19* presents estimates of the random therapist effects. The random therapist effect (estimate of covariance: 0.08535), that is, the between-cluster variance, was found to be not statistically significant ($p = 0.3707$) and the results of the fixed effects are very similar to those based on

TABLE 56 Cox proportional hazards model with shared frailty for therapist (adjusted for covariates)

Covariate	Parameter estimate	SE	HR (95% CI)	df	Adjusted df	Adjusted *p*-value
Treatment: FT (vs. TAU)	0.08	0.15	1.09 (0.81 to 1.46)	1	0.8230	0.4984
Gender: female (vs. male)	0.48	0.25	1.61 (0.98 to 2.63)	1	0.9879	0.0570
Age group (years): 15–17 (vs. 11–14)	–0.37	0.14	0.69 (0.52 to 0.91)	1	0.9862	0.0095
Number of previous self-harm episodes: ≥ 3 (vs. 2)	0.19	0.23	1.21 (0.77 to 1.91)	1	0.9812	0.4071
Type of index episode				2	1.9606	0.0347
Combined (vs. self-injury)	0.60	0.25	1.83 (1.13 to 2.98)	1	–	–
Self-poisoning (vs. self-injury)	0.02	0.20	1.02 (0.68 to 1.53)	1	–	–
Referred via hospital: yes (vs. no)	0.28	0.18	1.33 (0.93 to 1.88)	1	0.9821	0.1115
Trust (vs. Y trust 1)				14	12.9124	0.1452
M trust 3	0.18	0.25	1.19 (0.73 to 1.94)	1	–	–
L trust 5	–0.50	0.46	0.61 (0.24 to 1.51)	1	–	–
Y trust 3	–0.64	0.39	0.53 (0.25 to 1.13)	1	–	–
Y trust 2	–0.19	0.39	0.83 (0.38 to 1.78)	1	–	–
M trust 1	0.13	0.30	1.14 (0.63 to 2.05)	1	–	–
L trust 6	–0.27	0.73	0.76 (0.18 to 3.22)	1	–	–
Y trust 4	–1.14	0.55	0.32 (0.11 to 0.93)	1	–	–
L trust 2	–0.68	0.34	0.51 (0.26 to 0.99)	1	–	–
M trust 2	–0.24	0.27	0.79 (0.46 to 1.34)	1	–	–
L trust 1	–0.37	0.28	0.69 (0.39 to 1.20)	1	–	–
Y trust 6	0.24	0.36	1.28 (0.63 to 2.60)	1	–	–
L trust 4	–0.42	1.00	0.66 (0.09 to 4.64)	1	–	–
L trust 3	–0.18	0.55	0.83 (0.28 to 2.47)	1	–	–
Y trust 5	–0.25	0.44	0.78 (0.33 to 1.86)	1	–	–
Main therapist (random effect)[a]				–	13.6705	0.3707

df, degrees of freedom.
a Covariance parameter estimates for the random therapist effect are: restricted maximum likelihood estimate 0.08535 with SE 0.09254.

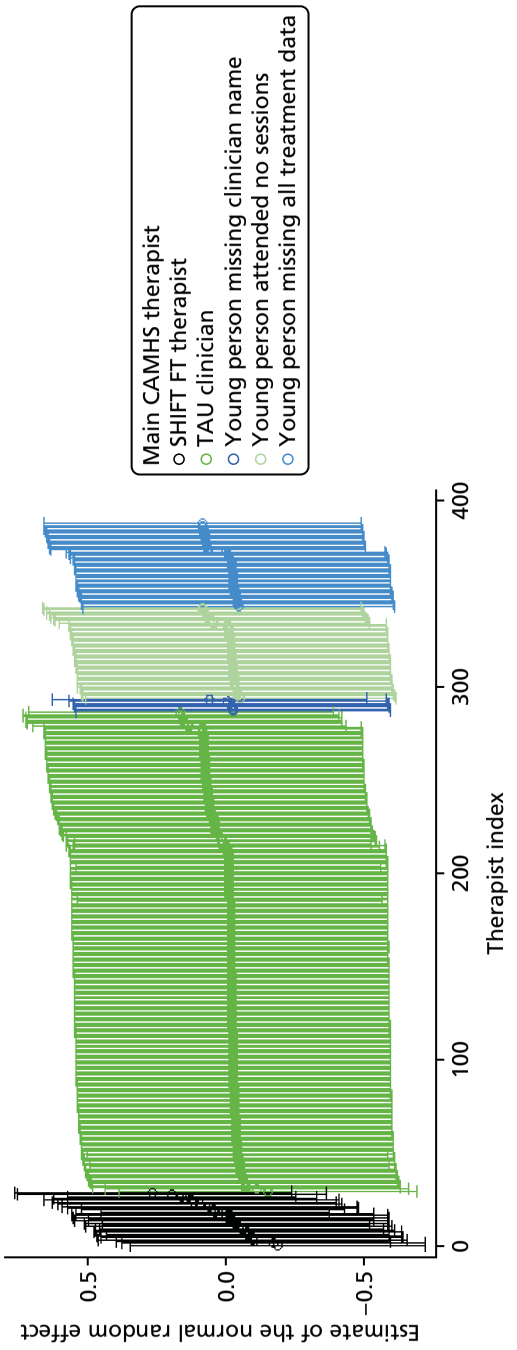


FIGURE 19 Estimated random therapist effect in the Cox proportional hazards model with shared frailty. Estimated random effects > 0 represent higher frailty, that is, shorter time to and higher rate of self-harm, while estimates < 0 represent lower frailty, that is, longer time to and lower rate of self-harm.

the primary analysis without accounting for shared frailty, leading to the same conclusion as the primary analysis of no evidence to suggest a difference between FT and TAU in delaying the time to, and occurrence of, self-harm leading to hospital attendance. See *Appendix 5* for further discussion and results relating to the estimated intracluster correlation for the therapist effect.

Secondary end points

Repetition of self-harm leading to hospital attendance within 12 months of randomisation

By 12 months post randomisation, self-harm events leading to hospital attendance were reported for 173 participants, 91 in the FT arm and 82 in the TAU; this gave a repetition rate of 22.1% (95% CI 18.1% to 26.1%) and 19.8% (95% CI 16.0% to 23.7%), respectively, with a difference of 2.3% (95% CI –3.3% to 7.8%) (see *Table 48*).

Table 57 presents results of the Cox proportional hazards regression model (adjusted for covariates) for the first event within 12 months of randomisation. The HR for treatment is 1.09 (95% CI 0.81 to 1.48) with a

TABLE 57 Cox proportional hazards model with 12 months follow-up (adjusted for covariates)[a]

Covariate	Parameter estimate	SE	HR (95% CI)	df	*p*-value
Treatment: FT (vs. TAU)	0.09	0.15	1.09 (0.81 to 1.48)	1	0.5592
Gender: female (vs. male)	0.47	0.28	1.60 (0.92 to 2.79)	1	0.0938
Age group (years): 15–17 (vs. 11–14)	–0.32	0.16	0.72 (0.53 to 0.99)	1	0.0432
Number of previous self-harm episodes: ≥ 3 (vs. 2)	0.27	0.27	1.31 (0.77 to 2.22)	1	0.3194
Type of index episode				2	0.0712
Combined (vs. self-injury)	0.59	0.28	1.80 (1.05 to 3.09)	1	–
Self-poisoning (vs. self-injury)	–0.00	0.23	1.00 (0.63 to 1.57)	1	–
Referred via hospital: yes (vs. no)	0.24	0.20	1.27 (0.86 to 1.88)	1	0.2326
Y trust 1				14	0.1354
M trust 3	0.36	0.26	1.44 (0.86 to 2.40)	1	–
L trust 5	–0.36	0.49	0.70 (0.27 to 1.84)	1	–
Y trust 3	–0.66	0.46	0.52 (0.21 to 1.27)	1	–
Y trust 2	0.07	0.41	1.07 (0.48 to 2.38)	1	–
M trust 1	0.35	0.32	1.42 (0.76 to 2.66)	1	–
L trust 6	0.06	0.74	1.06 (0.25 to 4.51)	1	–
Y trust 4	–0.76	0.54	0.47 (0.16 to 1.35)	1	–
L trust 2	–0.77	0.41	0.46 (0.21 to 1.02)	1	–
M trust 2	–0.08	0.29	0.92 (0.52 to 1.63)	1	–
L trust 1	–0.24	0.30	0.79 (0.43 to 1.43)	1	–
Y trust 6	0.38	0.37	1.47 (0.72 to 3.01)	1	–
L trust 4	–0.15	1.03	0.86 (0.12 to 6.44)	1	–
L trust 3	–0.19	0.61	0.83 (0.25 to 2.76)	1	–
Y trust 5	0.18	0.43	1.19 (0.51 to 2.77)	1	–

df, degrees of freedom.
a Reduction in –2logL between Cox proportional hazards model with and without Treatment effect: 0.342 ($p = 0.5589$, 1 degree of freedom).

p-value of 0.5592; hence, as per the primary first event model, albeit with a HR closer to 1, there is no evidence to suggest a difference in self-harm repetition rates between treatment groups.

There is no change in the direction of effects for any covariate with the exception of three centres (two in Yorkshire and one in London); however, these factors were found to be not significant. Results from model checks, similar to those for the primary analysis, similarly demonstrated the adequacy of the model.

Characteristics of further episodes of self-harm leading to hospital attendance

Summary of all self-harm events leading to hospital attendance

A total of 401 self-harm events leading to hospital attendance were observed in 221 (26.6%) participants within 18 months of randomisation: 206 events in 118 (28.4%) participants randomised to FT and 195 events in 103 (24.7%) participants randomised to TAU (*Table 58*). The overall number of events per participant ranged from 0 to 18. *Figure 20* presents a needle plot of all repeat events per participant.

TABLE 58 Overall number of repeat self-harm events

Repeat self-harm events	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
Number of participants with one or more event, *n* (%)	118 (28.4)	103 (24.7)	221 (26.6)
Number of events reported	206	195	401
Number of events per participant			
n	415	417	832
Mean (SD)	0.5 (1.05)	0.5 (1.26)	0.5 (1.16)
Median (range)	0.0 (0–9)	0.0 (0–18)	0.0 (0–18)
Number of events per participant (participants with an event)			
n	118	103	221
Mean (SD)	1.7 (1.30)	1.9 (1.93)	1.8 (1.62)
Median (range)	1.0 (1–9)	1.0 (1–18)	(1–18)
Number of events per participant, *n* (%)			
0	297 (71.6)	314 (75.3)	611 (73.4)
1	74 (17.8)	60 (14.4)	134 (16.1)
2	22 (5.3)	21 (5.0)	43 (5.2)
3	11 (2.7)	12 (2.9)	23 (2.8)
4	5 (1.2)	7 (1.7)	12 (1.4)
5	4 (1.0)	1 (0.2)	5 (0.6)
6	1 (0.2)	1 (0.2)	2 (0.2)
9	1 (0.2)	0 (0.0)	1 (0.1)
18	0 (0.0)	1 (0.2)	1 (0.1)
Total	415 (100)	417 (100)	832 (100)

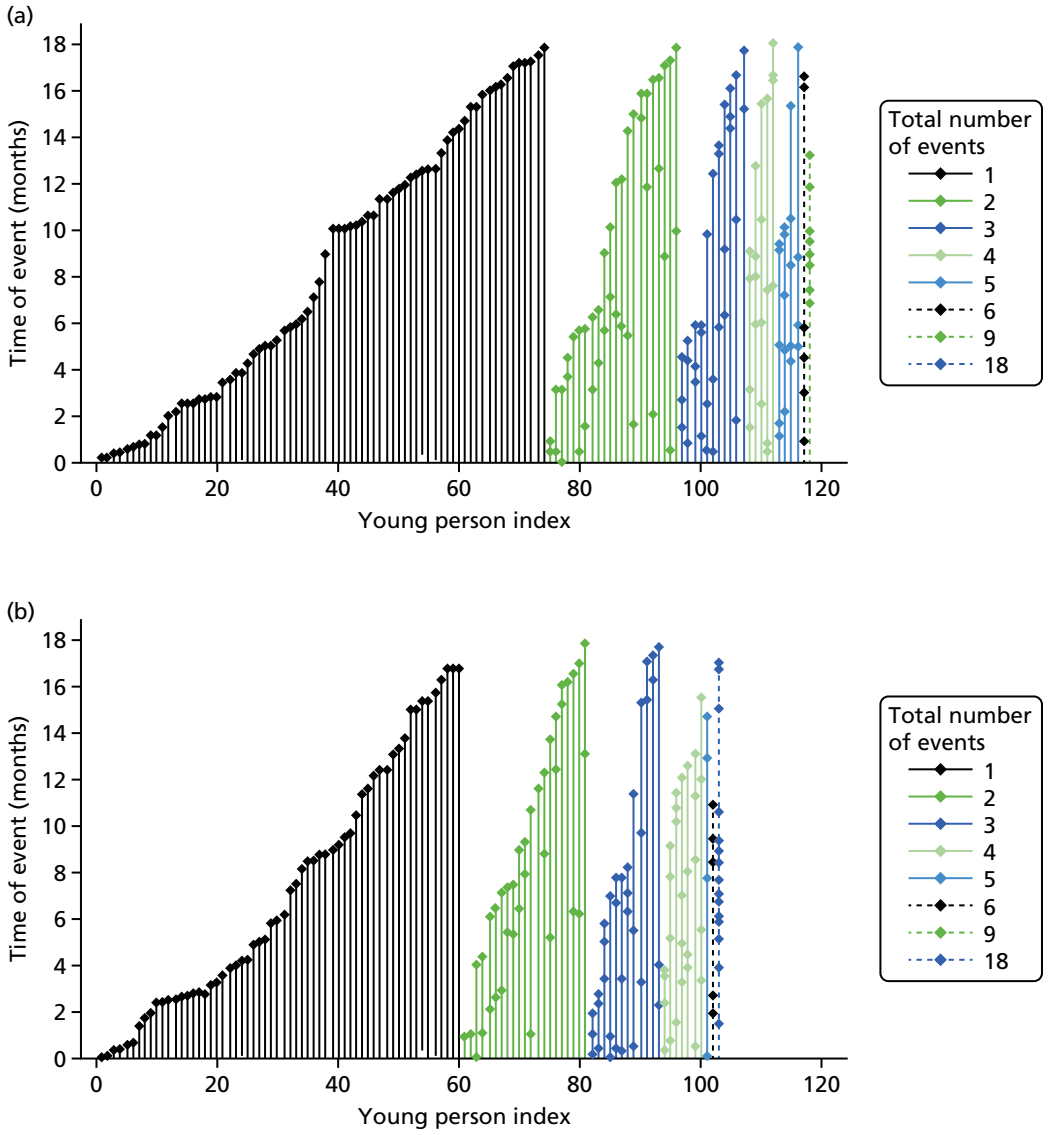


FIGURE 20 Needle plot of all recurrent events (self-harm leading to hospital attendance) ordered by total number of events and time of last event. (a) FT; and (b) TAU.

The most frequent reason for hospital attendance was self-poisoning, accounting for 209 (52.1%) events, with a further 93 (23.2%) attendances for cutting (*Table 59*). The most common outcome was admission to a hospital ward, which was the outcome of 265 (66.1%) events, with similar admission rates between trial arms. *Table 60* further details the reasons for attendance (type of self-harm) by outcome, which shows that self-poisoning was more likely than other reasons for attendance to result in admission.

The majority of young people discharged from A&E (90, 66.2%) were referred to their GP or to outpatient follow-up or did not require follow-up treatment (36, 26.5%). Ten admission events (3.8%) led to a psychiatric inpatient stay, with overall length of admission ranging from 0 to 91 days and a mean of 3.1 (SD 10.56) days (*Table 61*).

Recurrent event analysis

The results of the recurrent event analysis, using a counting process model with robust sandwich variance estimator for the rate of recurrent events, are presented in *Table 62*. The HR for treatment is 1.05 (95% CI 0.76s to 1.44) with a *p*-value of 0.7847, very similar to the primary first event model, albeit with a HR closer to 1. There is no evidence to suggest a difference in self-harm repetition rates between treatment groups.

TABLE 59 Type and outcome of all self-harm events

Type and outcome of all self-harm events	FT	TAU	Total
Number of events reported	206	195	401
Type of self-harm events, *n* (%)			
Poisoning	111 (53.9)	98 (50.3)	209 (52.1)
Cutting	42 (20.4)	51 (26.2)	93 (23.2)
Poisoning and cutting	10 (4.9)	15 (7.7)	25 (6.2)
Other self-injury	21 (10.2)	12 (6.2)	33 (8.2)
Other violent method	3 (1.5)	6 (3.1)	9 (2.2)
Other[a]	18 (8.7)	12 (6.2)	30 (7.5)
Missing	1 (0.5)	1 (0.5)	2 (0.5)
Total	206 (100)	195 (100)	401 (100)
Outcome of self-harm events, *n* (%)			
Discharged from A&E	72 (35.0)	64 (32.8)	136 (33.9)
Admitted to hospital ward	134 (65.0)	131 (67.2)	265 (66.1)
Total	206 (100)	195 (100)	401 (100)
Treatment received, *n* (%)			
Yes	185 (89.8)	180 (92.3)	365 (91.0)
No	18 (8.7)	13 (6.7)	31 (7.7)
Missing	3 (1.5)	2 (1.0)	5 (1.2)
Total	206 (100)	195 (100)	401 (100)
Level of treatment received, *n* (%)			
Minimal	106 (57.3)	113 (62.8)	219 (60.0)
Significant intervention	37 (20.0)	45 (25.0)	82 (22.5)
Missing	42 (22.7)	22 (12.2)	64 (17.5)
Total	185 (100)	180 (100)	365 (100)

a Other self-injury includes burning (4), head banging (1), punching (23), other (5); other violent methods include drowning (3), hanging/strangulation (1), jumping (4), other (1); other includes threat/thought of suicide only (17), psychiatric conditions (11), other (2).

Compared with the primary analysis of the first event, the effect of gender is no longer significant (HR 1.27, 95% CI 0.77 to 2.10; $p = 0.3407$), while the effect of referral via hospital is no longer masked by the centre effect and is highly significant, with a higher rate of self-harm in those referred via hospital (HR 1.98, 95% CI 1.18 to 3.32; $p = 0.0096$) and the centre effect is further enhanced, increasing in significance from the 10% level to the 5% level. Furthermore, the effect of the number of previous self-harm episodes increased compared with the first event model, but remained only borderline significant at the 10% level, while the overall effect of the type of index episode fell ($p = 0.0641$) and the direction of effect switched for self-poisoning versus self-injury.

Model checking identified that the model fit was adequate. Details are provided in *Appendix 6* (see *Figure 65*). Results of the sensitivity recurrent event analysis using the conditional, restricted gap-time Prentice–Williams–Peterson model[85] are also presented; although treatment and covariate estimates vary slightly, the overall conclusions that can be drawn from these remain similar to those based on the counting process model with robust sandwich variance estimator.

TABLE 60 Details of self-harm events by outcome

Details of self-harm events, *n* (%)	Discharged from A&E			Admitted to hospital ward		
	FT	TAU	Total	FT	TAU	Total
Type of self-harm event						
Poisoning	19 (26.4)	14 (21.9)	33 (24.3)	92 (68.7)	84 (64.1)	176 (66.4)
Cutting	20 (27.8)	29 (45.3)	49 (36.0)	22 (16.4)	22 (16.8)	44 (16.6)
Poisoning and cutting	1 (1.4)	1 (1.6)	2 (1.5)	9 (6.7)	14 (10.7)	23 (8.7)
Other self-injury	19 (26.4)	11 (17.2)	30 (22.1)	2 (1.5)	1 (0.8)	3 (1.1)
Other violent method	3 (4.2)	3 (4.7)	6 (4.4)	0 (0.0)	3 (2.3)	3 (1.1)
Other	9 (12.5)	5 (7.8)	14 (10.3)	9 (6.7)	7 (5.3)	16 (6.0)
Missing	1 (1.4)	1 (1.6)	2 (1.5)	0 (0.0)	0 (0.0)	0 (0.0)
Total	72 (100)	64 (100)	136 (100)	134 (100)	131 (100)	265 (100)
Treatment received						
Yes	51 (70.8)	49 (76.6)	100 (73.5)	134 (100.0)	131 (100.0)	265 (100.0)
No	18 (25.0)	13 (20.3)	31 (22.8)	0 (0.0)	0 (0.0)	0 (0.0)
Missing	3 (4.2)	2 (3.1)	5 (3.7)	0 (0.0)	0 (0.0)	0 (0.0)
Total	72 (100)	64 (100)	136 (100)	134 (100)	131 (100)	265 (100)
Level of treatment received						
Minimal	38 (74.5)	37 (75.5)	75 (75.0)	68 (50.7)	76 (58.0)	144 (54.3)
Significant intervention	12 (23.5)	12 (24.5)	24 (24.0)	25 (18.7)	33 (25.2)	58 (21.9)
Missing	1 (2.0)	0 (0.0)	1 (1.0)	41 (30.6)	22 (16.8)	63 (23.8)
Total	51 (100)	49 (100)	100 (100)	134 (100)	131 (100)	265 (100)

TABLE 61 Details of discharge from A&E and admissions for self-harm events

Discharge and admission details, *n* (%)	FT	TAU	Total
Discharged form A&E			
Type of discharge			
Self-discharge	3 (4.2)	5 (7.8)	8 (5.9)
Discharged with referral to GP or outpatient follow-up	51 (70.8)	39 (60.9)	90 (66.2)
Discharged – did not require any follow-up treatment	16 (22.2)	20 (31.3)	36 (26.5)
Transferred to other health-care professional/provider	2 (2.8)	0 (0.0)	2 (1.5)
Total	72 (100)	64 (100)	136 (100)
Admitted			
Type of ward			
Paediatric ward	45 (33.6)	34 (26.0)	79 (29.8)
Assessment unit	2 (1.5)	3 (2.3)	5 (1.9)
Adult ward	0 (0.0)	1 (0.8)	1 (0.4)
Missing	87 (64.9)	93 (71.0)	180 (67.9)
Total	134 (100)	131 (100)	265 (100)

TABLE 61 Details of discharge from A&E and admissions for self-harm events (*continued*)

Discharge and admission details, *n* (%)	FT	TAU	Total
Psychiatric inpatient stay?			
Yes	6 (4.5)	4 (3.1)	10 (3.8)
No	128 (95.5)	127 (96.9)	255 (96.2)
Total	134 (100)	131 (100)	265 (100)
Length of admission (days)			
n	129	129	258
n missing	5	2	7
Mean (SD)	4.5 (14.42)	1.7 (3.52)	3.1 (10.56)
Median (range)	1.0 (0–91)	1.0 (0–24)	1.0 (0–91)
Length of admission (days)			
0	22 (16.4)	29 (22.1)	51 (19.2)
1	66 (49.3)	69 (52.7)	135 (50.9)
2	19 (14.2)	15 (11.5)	34 (12.8)
3–9	15 (11.2)	12 (9.2)	27 (10.2)
10–30	1 (0.7)	4 (3.1)	5 (1.9)
> 30	6 (4.5)	0 (0.0)	6 (2.3)
Missing	5 (3.7)	2 (1.5)	7 (2.6)
Total	134 (100)	131 (100)	265 (100)

Characteristics of all further self-reported episodes of self-harm (SASII)

Characteristics of the first self-reported self-harm episode

Details of the young person's first episode of self-harm post-randomisation, regardless of whether or not it led to hospital attendance, as self-reported via the SASII questionnaire, are presented in *Table 63*.

Self-harm was reported during the first 12–18 months of follow-up for 202 (75.4%) in the FT arm and 147 (70.0%) in the TAU arm (see *Table 63*) for those with sufficient SASII completion from baseline to 12 or 18 months (completed at 12 months, 12 and 18 months or 18 months only but with a report of self-harm). Self-harm was reported during the first 12 months for 186 (71.5%) in the FT arm and 131 (65.5%) in the TAU arm (*Table 64*).

Figure 21 presents the time to first episode. The researcher rated as a suicide attempt self-harm by 44 (13%) participants, with similar rates across trial arms. Forty-three participants (21%) in the FT arm and 41 (28%) participants in the TAU arm were considered to have some intent to die; however, in the majority of these cases (82%), self-harm was deemed low risk (as defined by the SASII), with medical treatment sought or received by 28 participants (14%) in the FT arm and 23 participants (16%) in the TAU arm. Timing and details of the first self-harm episode are missing for 18 participants in whom self-harm was reported at 18 months but for whom there were no 12-month data to indicate the first post-baseline episode.

Suicide Attempt Self-Injury Interview timeline

Summaries of the number, methods and treatment required for self-harm episodes reported on the SASII timeline are presented in *Tables 64–68*.

TABLE 62 Recurrent event analysis estimates[a]

	Primary recurrent event analysis[b]				Sensitivity analysis[c]			
Covariate	Parameter Estimate	SE	HR (95% CI)	*p*-value	Parameter estimate	SE	HR (95% CI)	*p*-value
Treatment: FT (vs. TAU)	0.04	0.16	1.05 (0.76 to 1.44)	0.7847	0.08	0.10	1.08 (0.88 to 1.32)	0.4565
Gender: female (vs. male)	0.24	0.25	1.27 (0.77 to 2.10)	0.3407	0.07	0.17	1.07 (0.76 to 1.51)	0.6883
Age group (years): 15–17 (vs. 11–14)	–0.39	0.16	0.67 (0.50 to 0.92)	0.0120	–0.26	0.11	0.77 (0.63 to 0.96)	0.0178
Number of previous self-harm episodes: ≥ 3 (vs. 2)	0.42	0.25	1.52 (0.92 to 2.49)	0.1008	0.28	0.19	1.32 (0.91 to 1.92)	0.1460
Type of index episode				0.0641				0.1861
Combined (vs. self-injury)	0.19	0.30	1.20 (0.66 to 2.18)	–	0.19	0.18	1.21 (0.85 to 1.72)	–
Self-poisoning (vs. self-injury)	–0.33	0.24	0.72 (0.45 to 1.16)	–	–0.16	0.15	0.85 (0.63 to 1.15)	–
Referred via hospital: yes (vs. no)	0.68	0.26	1.98 (1.18 to 3.32)	0.0096	0.38	0.13	1.46 (1.14 to 1.88)	0.0031
Trust (vs. Y trust 1)				0.0493				0.1353
M trust 3	0.44	0.24	1.56 (0.97 to 2.50)	–	0.29	0.17	1.33 (0.95 to 1.88)	
L trust 5	–0.13	0.52	0.88 (0.32 to 2.44)	–	–0.07	0.32	0.94 (0.50 to 1.77)	
Y trust 3	–0.36	0.41	0.70 (0.31 to 1.57)	–	–0.23	0.29	0.79 (0.44 to 1.41)	
Y trust 2	0.07	0.33	1.07 (0.56 to 2.06)	–	–0.01	0.28	0.99 (0.57 to 1.70)	
M trust 1	0.11	0.29	1.12 (0.63 to 1.97)	–	0.04	0.24	1.04 (0.66 to 1.65)	
L trust 6	0.19	0.63	1.21 (0.36 to 4.14)	–	0.06	0.50	1.06 (0.40 to 2.80)	
Y trust 4	–0.94	0.58	0.39 (0.12 to 1.21)	–	–0.77	0.43	0.46 (0.20 to 1.09)	
L trust 2	–0.38	0.33	0.68 (0.35 to 1.31)	–	–0.24	0.25	0.79 (0.48 to 1.29)	
M trust 2	0.08	0.27	1.09 (0.64 to 1.84)	–	0.08	0.20	1.08 (0.73 to 1.59)	
L trust 1	–0.39	0.26	0.68 (0.41 to 1.13)	–	–0.29	0.22	0.74 (0.48 to 1.15)	

	Primary recurrent event analysis[b]				Sensitivity analysis[c]			
Covariate	Parameter Estimate	SE	HR (95% CI)	*p*-value	Parameter estimate	SE	HR (95% CI)	*p*-value
Y trust 6	0.87	0.51	2.40 (0.88 to 6.54)	–	0.44	0.24	1.56 (0.98 to 2.48)	
L trust 4	–0.66	1.08	0.52 (0.06 to 4.26)	–	–0.47	1.01	0.63 (0.09 to 4.55)	
L trust 3	0.11	0.50	1.11 (0.42 to 2.97)	–	0.02	0.38	1.02 (0.48 to 2.17)	
Y trust 5	0.16	0.47	1.17 (0.47 to 2.94)	–	0.08	0.30	1.09 (0.60 to 1.97)	

a Four participants were excluded as a result of lack of any follow-up leading to equal start and stop times, invalid under the counting process model. A single event was also excluded where two events took place on the same day for the same participant (again leading to equal start and stop times). A further participant had an event on the same day as randomisation (leading to equal start and stop times) and to include this event the time of event was adjusted to 1 day post randomisation.
b Andersen–Gill model with robust sandwich variance estimator: unrestricted counting process.
c Prentice–Williams–Peterson model:[85] conditional restricted gap time.

TABLE 63 Characteristics of the first self-harm episode post randomisation (SASII)

Characteristics of the first self-harm episode	FT (*N* = 415)	TAU (*N* = 417)	Total (*N* = 832)
SASII completed?			
Yes at 12 and 18 months, *n* (%)	187 (45.1)	144 (34.5)	331 (39.8)
Partly at 12 months, *n* (%)	73 (17.6)	56 (13.4)	129 (15.5)
Partly at 18 months, *n* (%)	17 (4.1)	21 (5.0)	38 (4.6)
No, *n* (%)	138 (33.3)	196 (47.0)	334 (40.1)
SASII sufficiently completed,[a] *n* (%)	268 (64.6)	210 (50.4)	478 (57.5)
Deliberately harmed self, *n* (%)			
Yes	202 (75.4)	147 (70.0)	349 (73.0)
No	66 (24.6)	63 (30.0)	129 (27.0)
Total	268 (100.0)	210 (100.0)	478 (100.0)
Time of self-harm episode (months)			
n	192	136	328
Missing	10	11	21
Mean (SD)	3.2 (4.42)	3.2 (4.16)	3.2 (4.31)
Median (range)	1.1 (0.0–17.9)	1.0 (0.0–16.6)	1.1 (0.0–17.9)
Type of behaviour,[b] *n* (%)			
Suicide attempt	25 (12.4)	19 (12.9)	44 (12.6)
NSSI	158 (78.2)	113 (76.9)	271 (77.7)
Missing	19 (9.4)	15 (10.2)	34 (9.7)
Intent to die[b] (some intent to die)	43 (21.3)	41 (27.9)	84 (24.1)
Medical risk of death[b] (low)	164 (81.2)	121 (82.3)	285 (81.7)
Impulsivity[b] (planned)	41 (20.3)	26 (17.7)	67 (19.2)
Communication/threat of suicide intent (yes)	22 (10.9)	28 (19.0)	50 (14.3)
Interpersonal influence (mentioned)	63 (31.2)	51 (34.7)	114 (32.7)
Emotion relief (mentioned)	170 (84.2)	127 (86.4)	297 (85.1)
Medical treatment sought/received (yes)	28 (13.9)	23 (15.6)	51 (14.6)
Physical condition following episode,[b] *n* (%)			
No effect	27 (13.4)	22 (15.0)	49 (14.0)
Mild effect	145 (71.8)	105 (71.4)	250 (71.6)
Moderate effect	8 (4.0)	4 (2.7)	12 (3.4)
Severe effect	1 (0.5)	1 (0.7)	2 (0.6)
Missing	21 (10.4)	15 (10.2)	36 (10.3)
Total	202 (100.0)	147 (100.0)	349 (100.0)

a At 12 months, 12 and 18 months or 18 months with self-harm.
b Interviewer-rated items.

TABLE 64 Self-reported self-harm episodes (SASII timeline)[a]

Time of self-harm episode	FT	TAU	Total
Proportion with self-harm during, *n* (%)			
0–3 months	136 (52.3)	92 (46.0)	228 (49.6)
> 3–6 months	137 (52.7)	90 (45.0)	227 (49.3)
> 6–9 months	116 (44.6)	86 (43.0)	202 (43.9)
> 9–12 months	125 (48.1)	79 (39.5)	204 (44.3)
Total	260 (100.0)	200 (100.0)	460 (100.0)
> 12–15 months	71 (34.8)	57 (34.5)	128 (34.7)
> 15–18 months	75 (36.8)	54 (32.7)	129 (35.0)
Total	204 (100.0)	165 (100.0)	369 (100.0)
Cumulative proportion with self-harm (forward), *n* (%)			
≤ 3 months	136 (52.3)	92 (46.0)	228 (49.6)
≤ 6 months	164 (63.1)	113 (56.5)	277 (60.2)
≤ 9 months	171 (65.8)	124 (62.0)	295 (64.1)
≤ 12 months	186 (71.5)	131 (65.5)	317 (68.9)
Total	260 (100.0)	200 (100.0)	460 (100.0)
Cumulative proportion with self-harm (backward, up to 12 months), *n* (%)			
> 9 months	125 (48.1)	79 (39.5)	204 (44.3)
> 6 months	148 (56.9)	106 (53.0)	254 (55.2)
> 3 months	168 (64.6)	119 (59.5)	287 (62.4)
> 0 months	186 (71.5)	131 (65.5)	317 (68.9)
Total	260 (100.0)	200 (100.0)	460 (100.0)

a Timing of self-harm unknown for four young people, with young people therefore classed as having self-harm at all time points.

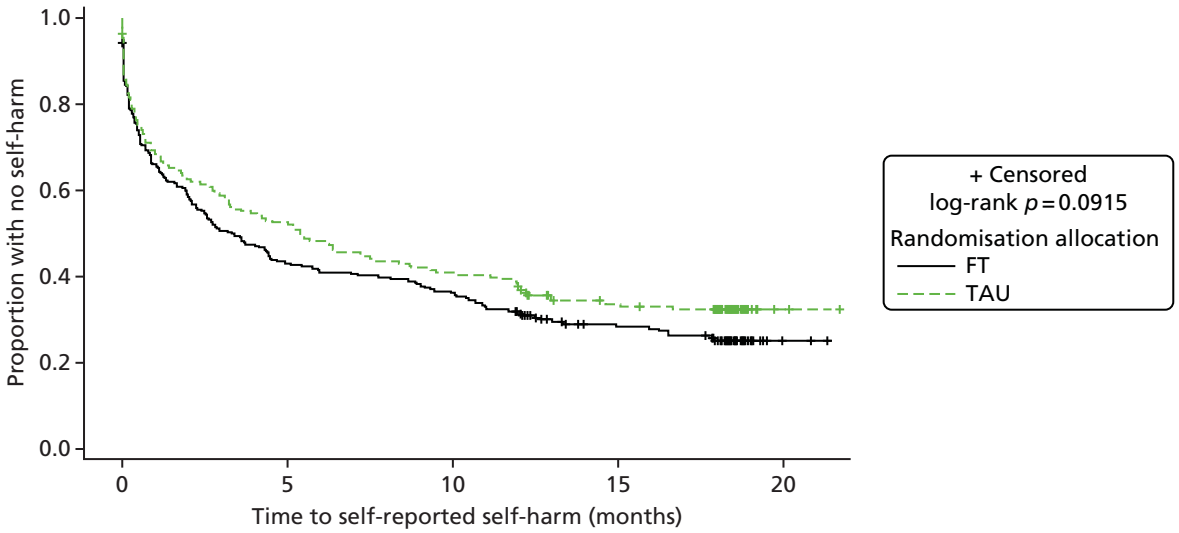


FIGURE 21 Time to first self-reported self-harm (SASII). Participants with missing data are censored at randomisation, as are participants known to have self-harmed but for whom date of first self-harm is missing.

Overall, self-harm was reported during the first 3 months post randomisation by almost half the young people completing the SASII (49.6%; see *Table 64*). Self-harm rates decreased slightly in the following months, 49.3% over 3–6 months post randomisation, 43.9% over 6–9 months and 44.3% over 9–12 months, with higher rates in the FT arm than in the TAU arm. Rates decreased further but were more similar across arms post 12 months: 34.7% over 12–15 months and 35% over 15–18 months; however, missing data increased at 18 months.

Cumulatively, the overall proportion who self-harmed increased from 49.6% by 3 months post randomisation to 60.2% by 6 months, 64.1% by 9 months and 68.9% by 12 months. Furthermore, in the period up to 12 months post randomisation, self-harm was still being reported for 44.3% participants from 9 months post randomisation, for 55.2% from 6 months, for 62.4% from 3 months and again for 68.9% from 12 months post randomisation. *Figure 22* further depicts the cumulative proportion of participants reporting self-harm in 3-month intervals.

The number of self-harm episodes per participant is highly skewed, with a large proportion of participants reporting no self-harm; therefore, the median number of episodes over all participants varies between 0 and 1. However, the mean number of self-harm episodes (*Table 65*) in the first 3 months post randomisation was 7.1 (SD 16.22) in the FT arm and 9 (SD 19.96) in the TAU arm; these figures were generally stable over time, but reduced to a mean of 5.0 (SD 17.23) in the FT arm and 4.5 (SD 12.78) in the TAU arm by 9–12 months. The number of episodes fell further during later months; however, missing data increased at 18 months. The mean number of self-harm episodes between 12 and 15 months was 3.6 (SD 11.89) in the FT arm and 2.5 (SD 7.84) in the TAU arm and between 15 and 18 months post randomisation was 3.0 (SD 10.15) in the FT arm and 2.3 (SD 7.66) in the TAU arm.

The mean total number of self-harm episodes reported overall (of those with at least partial SASII completion at 12 and/or 18 months) was 29.0 (SD 76.07) in the FT arm and 27.7 (SD 57.67) in the TAU arm (see *Table 65*); however, this varied further when broken down by participants with full or partial SASII completion at 12 and 18 months (*Table 66*).

Hospital medical treatment was self-reported by a total of 95 (27.2%) participants reporting self-harm in the SASII timeline, 56 (27.7%) in the FT arm and 39 (26.5%) in the TAU arm, while no medical treatment was required for any self-harm reported by 225 (64.5%) participants, with similar proportions in each arm (*Table 67*).

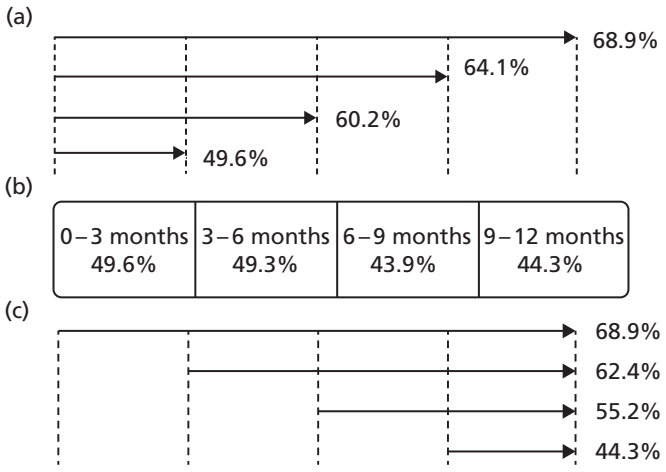


FIGURE 22 Proportion of participants who self-reported self-harm up to 12 months post randomisation at 3-month and cumulative intervals. (a) Forward cumulative proportion with self-harm; (b) proportion with self-harm; and (c) backward cumulative proportion with self-harm.

TABLE 65 Frequency of self-reported self-harm episodes (SASII)

Time period post randomisation/trial arm	Number of self-harm episodes				Frequency of self-harm,[a] *n* (%)						
	N	Missing	Mean (SD)	Median (range)	None	Less than once a month	Less than once a fortnight	Less than once a week	Once or more a week	Twice or more a week	Most days
0–3 months post randomisation											
FT	257	158	7.1 (16.22)	1.0 (0–162)	124 (47.7)	39 (15.0)	36 (13.8)	22 (8.5)	16 (6.2)	124 (47.7)	39 (15.0)
TAU	199	218	9.0 (19.96)	0.0 (0–166)	108 (54.0)	24 (12.0)	15 (7.5)	14 (7.0)	13 (6.5)	108 (54.0)	24 (12.0)
Total	456	376	7.9 (17.95)	0.0 (0–166)	232 (50.4)	63 (13.7)	51 (11.1)	36 (7.8)	29 (6.3)	232 (50.4)	63 (13.7)
> 3–6 months post randomisation											
FT	257	158	7.2 (20.72)	1.0 (0–214)	123 (47.3)	47 (18.1)	31 (11.9)	22 (8.5)	20 (7.7)	9 (3.5)	8 (3.1)
TAU	199	218	6.7 (16.05)	0.0 (0–95)	110 (55.0)	28 (14.0)	20 (10.0)	12 (6.0)	15 (7.5)	6 (3.0)	9 (4.5)
Total	456	376	7.0 (18.81)	0.0 (0–214)	233 (50.7)	75 (16.3)	51 (11.1)	34 (7.4)	35 (7.6)	15 (3.3)	17 (3.7)
> 6–9 months post randomisation											
FT	257	158	6.6 (21.59)	0.0 (0–221)	144 (55.4)	42 (16.2)	30 (11.5)	15 (5.8)	12 (4.6)	7 (2.7)	10 (3.8)
TAU	199	218	6.7 (21.61)	0.0 (0–185)	114 (57.0)	30 (15.0)	20 (10.0)	9 (4.5)	12 (6.0)	10 (5.0)	5 (2.5)
Total	456	376	6.6 (21.58)	0.0 (0–221)	258 (56.1)	72 (15.7)	50 (10.9)	24 (5.2)	24 (5.2)	17 (3.7)	15 (3.3)
> 9–12 months post randomisation											
FT	258	157	5.0 (17.23)	0.0 (0–186)	135 (51.7)	60 (23.0)	31 (11.9)	15 (5.7)	9 (3.4)	5 (1.9)	6 (2.3)
TAU	199	218	4.5 (12.78)	0.0 (0–92)	121 (60.5)	30 (15.0)	18 (9.0)	14 (7.0)	7 (3.5)	4 (2.0)	6 (3.0)
Total	456	375	4.8 (15.44)	0.0 (0–186)	256 (55.5)	90 (19.5)	49 (10.6)	29 (6.3)	16 (3.5)	9 (2.0)	12 (2.6)
> 12–15 months post randomisation											
FT	201	214	3.6 (11.89)	0.0 (0–91)	133 (65.2)	29 (14.2)	18 (8.8)	8 (3.9)	9 (4.4)	4 (2.0)	3 (1.5)
TAU	164	253	2.5 (7.84)	0.0 (0–58)	108 (65.5)	30 (18.2)	13 (7.9)	5 (3.0)	4 (2.4)	3 (1.8)	2 (1.2)
Total	365	467	3.1 (10.27)	0.0 (0–91)	241 (65.3)	59 (16.0)	31 (8.4)	13 (3.5)	13 (3.5)	7 (1.9)	5 (1.4)

continued

TABLE 65 Frequency of self-reported self-harm episodes (SASII) (*continued*)

	Number of self-harm episodes				Frequency of self-harm,[a] *n* (%)						
Time period post randomisation/trial arm	*N*	Missing	Mean (SD)	Median (range)	None	Less than once a month	Less than once a fortnight	Less than once a week	Once or more a week	Twice or more a week	Most days
> 15–18 months post randomisation											
FT	201	214	3.0 (10.15)	0.0 (0–92)	129 (63.2)	39 (19.1)	16 (7.8)	8 (3.9)	5 (2.5)	4 (2.0)	3 (1.5)
TAU	164	253	2.3 (7.66)	0.0 (0–60)	111 (67.3)	30 (18.2)	10 (6.1)	7 (4.2)	3 (1.8)	1 (0.6)	3 (1.8)
Total	365	467	2.7 (9.11)	0.0 (0–92)	240 (65.0)	69 (18.7)	26 (7.0)	15 (4.1)	8 (2.2)	5 (1.4)	6 (1.6)
Total self-harm episodes (out of all with at least partial completion)											
FT	275	140	29.0 (76.07)	4.0 (0–868)							
TAU	221	196	27.7 (57.67)	2.0 (0–396)							
Total	496	336	28.4 (68.42)	3.0 (0–868)							

a None (0 days), less than once a month (< 3 days), less than once a fortnight (3–6 days), less than once a week (7–12 days), once or more a week (13–25 days), twice or more a week (26–45 days), most days (46–91 days). When the timing of self-harm was unknown (for four young people) but self-harm was indicated, these young people were classed as having the lowest frequency of self-harm at all time points.

TABLE 66 Summary of missing SASII data and total number of self-reported episodes

Missing SASII data	FT	TAU	Total
SASII at 12 and 18 months: total self-harm episodes			
n	187	144	331
Mean (SD)	34.6 (82.07)	28.9 (56.17)	32.1 (71.91)
Median (range)	6.0 (0–868)	4.0 (0–396)	5.0 (0–868)
Number with self-harm, *n* (%)	140 (74.9)	94 (65.3)	234 (70.7)
SASII at 12 months only: total self-harm episodes			
n	73	56	129
Mean (SD)	20.0 (65.67)	33.6 (69.30)	25.9 (67.34)
Median (range)	2.0 (0–504)	2.5 (0–338)	2.0 (0–504)
Number with self-harm, *n* (%)	53 (72.6)	42 (75.0)	95 (73.6)
SASII at 18 months only: total self-harm episodes			
n	17	21	38
Mean (SD)	2.7 (5.74)	3.9 (10.40)	3.3 (8.55)
Median (range)	0.0 (0–23)	0.0 (0–47)	0.0 (0–47)
Number with self-harm, *n* (%)	8 (47.1)	9 (42.9)	17 (44.7)

TABLE 67 Most serious treatment required

Most serious medical treatment required, *n* (%)	FT	TAU	Total
Hospital medical treatment	56 (27.7)	39 (26.5)	95 (27.2)
Non-hospital medical treatment	15 (7.4)	10 (6.8)	25 (7.2)
No medical treatment	129 (63.9)	96 (65.3)	225 (64.5)
Missing	2 (1.0)	2 (1.4)	4 (1.1)
Total	202 (100.0)	147 (100.0)	349 (100.0)

The most frequent type of self-harm reported was scratching/cutting, with a reported 11,779 (83.6%) episodes in 312 (89.4%) participants who reported self-harm with sufficient SASII completion (*Table 68*). There were 601 (4.3%) episodes involving drugs/medications among 92 (26.4%) participants and 534 (3.8%) hitting episodes among 24 (6.9%) participants. Similar numbers were observed across trial arms. However, there were slightly fewer episodes of scratching/cutting in the FT arm than in the TAU arm: 6302 (79.1%) episodes in 178 (88.1%) participants and 5477 (89.4%) episodes in 134 (91.2%) participants, respectively. In contrast, there were more drugs/medication-related episodes and hitting episodes in the FT arm than in the TAU arm; however, the numbers of participants and the difference in numbers between arms are low for both types of episodes, especially for hitting.

TABLE 68 Methods of all self-harm and per participant and most serious treatment required

Self-harm method	Young person with self-harm (not mutually exclusive), *n* (%)			All episodes, *n* (%)		
	FT	TAU	Total	FT	TAU	Total
Scratch/cut	178 (88.1)	134 (91.2)	312 (89.4)	6302 (79.1)	5477 (89.4)	11,779 (83.6)
Drugs/medications	57 (28.2)	35 (23.8)	92 (26.4)	548 (6.9)	53 (0.9)	601 (4.3)
Hitting body	15 (7.4)	9 (6.1)	24 (6.9)	423 (5.3)	111 (1.8)	534 (3.8)
Burning	12 (5.9)	5 (3.4)	17 (4.9)	56 (0.7)	35 (0.6)	91 (0.6)
Strangling	4 (2.0)	1 (0.7)	5 (1.4)	6 (0.1)	24 (0.4)	30 (0.2)
Poison/caustic substance	3 (1.5)	1 (0.7)	4 (1.1)	3 (0.0)	1 (0.0)	4 (0.0)
Stepped into traffic	1 (0.5)	2 (1.4)	3 (0.9)	1 (0.0)	4 (0.1)	5 (0.0)
Hanging	1 (0.5)	1 (0.7)	2 (0.6)	1 (0.0)	1 (0.0)	2 (0.0)
Jumping	1 (0.5)	1 (0.7)	2 (0.6)	1 (0.0)	1 (0.0)	2 (0.0)
Stabbing, puncture	0 (0.0)	1 (0.7)	1 (0.3)	0 (0.0)	4 (0.1)	4 (0.0)
Drowning	0 (0.0)	1 (0.7)	1 (0.3)	0 (0.0)	4 (0.1)	4 (0.0)
Transportation-related injury	0 (0.0)	1 (0.7)	1 (0.3)	0 (0.0)	2 (0.0)	2 (0.0)
Method not specified	6 (3.0)	6 (4.1)	12 (3.4)	48 (0.6)	16 (0.3)	64 (0.5)
Other	17 (8.4)	13 (8.8)	30 (8.6)	583 (7.3)	392 (6.4)	975 (6.9)
Total	202 (100.0)	147 (100.0)	349 (100.0)	7972 (100.0)	6125 (100.0)	14097 (100.0)

Young person- and caregiver-completed questionnaires

Young person questionnaires

The summary statistics of complete data for the young person questionnaire outcomes are presented in *Table 69*. The results of the repeated measures modelling (adjusted for baseline scores and covariates), presenting mean total and subscale scores, differences between the two groups and 95% CIs at each time point, are shown in *Tables 70* and *71*. The results of the sensitivity analysis conducted using complete data were similar, with no change to conclusions. Missing data patterns used to determine the number of imputations and baseline characteristics of participants lost to follow-up at 12 and 18 months are presented in *Tables 109–112* and *Figure 66*, *Appendix 6*.

No significant differences between treatment groups were detected in young person questionnaire outcomes (CDRS-R, PQ-LES-Q, BSS, Hopelessness, McMaster FAD) with the exception of a number of subscales on the SDQ and for suicide ideation screening on the BSS.

On the SDQ, there was good evidence (p-value < 0.05) that participants had better outcomes on the prosocial subscale in the FT arm at both 12 and 18 months and on the impact subscale at 12 months only.

There was good evidence of reduced odds of suicide ideation (as reported on the Beck scale) in the FT arm at 12 months [odds ratio (OR) 0.64, 95% CI 0.44 to 0.94, $p = 0.0242$]. By 18 months, suicide ideation had decreased further in the TAU group, but it decreased to a lesser extent in the FT group, so the difference between treatment arms was no longer statistically significant.

Unadjusted mean estimates and 95% CIs of complete data are presented in *Figures 23*, *24* and *25* for questionnaires and subscales with good evidence of a treatment effect at 12 or 18 months ($p < 0.05$).

TABLE 69 Summary statistics for the young person questionnaire outcomes

Outcome	Baseline			12 months			18 months		
	FT	TAU	Total	FT	TAU	Total	FT	TAU	Total
CDRS-R									
Total CDRS-R score (17–113),[a] *N*	415	416	831	244	187	431	204	165	369
Mean (SD)	48.0 (14.19)	49.4 (13.29)	48.7 (13.76)	36.5 (14.33)	37.2 (13.09)	36.8 (13.80)	33.8 (14.77)	35.0 (14.39)	34.4 (14.59)
Not depressed (< 30), *n* (%)	45 (10.8)	24 (5.8)	69 (8.3)	94 (38.5)	63 (33.7)	157 (36.4)	96 (47.1)	76 (46.1)	172 (46.6)
Mild depression (30–42), *n* (%)	108 (26.0)	108 (26.0)	216 (26.0)	72 (29.5)	67 (35.8)	139 (32.3)	57 (27.9)	46 (27.9)	103 (27.9)
Moderate depression (43–57), *n* (%)	154 (37.1)	168 (40.4)	322 (38.7)	54 (22.1)	45 (24.1)	99 (23.0)	33 (16.2)	28 (17.0)	61 (16.5)
Severe depression (58–72), *n* (%)	91 (21.9)	102 (24.5)	193 (23.2)	22 (9.0)	10 (5.3)	32 (7.4)	15 (7.4)	10 (6.1)	25 (6.8)
Very severe depression (> 72), *n* (%)	17 (4.1)	14 (3.4)	31 (3.7)	2 (0.8)	2 (1.1)	4 (0.9)	3 (1.5)	5 (3.0)	8 (2.2)
PQ-LES-Q									
Total PQ-LES-Q score (14–70),[b] *N*	411	408	819	259	201	460	213	180	393
Mean (SD)	41.2 (9.37)	41.2 (9.45)	41.2 (9.41)	48.5 (10.57)	47.3 (10.26)	47.9 (10.44)	49.1 (11.14)	48.7 (11.25)	48.9 (11.18)
Beck									
Beck score (0–38),[c] *N*	408	408	816	257	202	459	212	180	392
Mean (SD)	10.8 (8.94)	10.4 (9.42)	10.6 (9.18)	4.6 (7.25)	5.7 (7.91)	5.1 (7.56)	4.6 (7.76)	5.2 (7.76)	4.9 (7.76)
Median	11.0	9.0	10.0	0.0	0.0	0.0	0.0	0.0	0.0
Suicide ideation (screening Q4 and Q5), *n* (%)	323 (79.2)	306 (75.0)	629 (77.1)	111 (43.2)	98 (48.5)	209 (45.5)	85 (40.1)	80 (44.4)	165 (42.1)
Total hopelessness score (0–17),[d] *N*	409	406	815	255	201	456	213	179	392
Mean (SD)	7.7 (4.30)	7.3 (4.19)	7.5 (4.25)	4.9 (4.14)	5.2 (4.14)	5.0 (4.14)	4.6 (4.26)	4.8 (4.00)	4.7 (4.14)
SDQ[e]									
Total difficulties score (0–40), *N*	413	415	828	261	204	465	213	182	395
Mean (SD)	19.6 (5.70)	20.1 (5.60)	19.8 (5.65)	15.5 (6.76)	16.8 (6.93)	16.1 (6.86)	14.2 (7.37)	15.5 (6.84)	14.8 (7.15)
Close to average (0–14), *n* (%)	84 (20.3)	70 (16.9)	154 (18.6)	123 (47.1)	77 (37.7)	200 (43.0)	117 (54.9)	79 (43.4)	196 (49.6)
Slightly raised (15–17), *n* (%)	58 (14.0)	68 (16.4)	126 (15.2)	41 (15.7)	29 (14.2)	70 (15.1)	27 (12.7)	30 (16.5)	57 (14.4)

continued

TABLE 69 Summary statistics for the young person questionnaire outcomes (*continued*)

Outcome	Baseline			12 months			18 months		
	FT	TAU	Total	FT	TAU	Total	FT	TAU	Total
High (18–19), *n* (%)	60 (14.5)	44 (10.6)	104 (12.6)	19 (7.3)	25 (12.3)	44 (9.5)	19 (8.9)	17 (9.3)	36 (9.1)
Very high (20–40), *n* (%)	211 (51.1)	233 (56.1)	444 (53.6)	78 (29.9)	73 (35.8)	151 (32.5)	50 (23.5)	56 (30.8)	106 (26.8)
Prosocial score (0–10), *N*	414	415	829	261	204	465	213	182	395
Mean (SD)	7.1 (1.84)	7.2 (1.87)	7.2 (1.85)	7.7 (1.81)	7.4 (2.07)	7.6 (1.93)	7.8 (1.86)	7.6 (2.00)	7.7 (1.92)
Emotional problems score (0–10), *N*	414	415	829	261	204	465	213	182	395
Mean (SD)	6.2 (2.34)	6.5 (2.31)	6.4 (2.33)	5.0 (2.73)	5.2 (2.61)	5.1 (2.68)	4.4 (2.97)	4.9 (2.71)	4.6 (2.86)
Conduct problems score (0–10), *N*	414	415	829	261	204	465	213	182	395
Mean (SD)	3.9 (2.11)	4.0 (1.94)	3.9 (2.02)	2.8 (1.98)	3.0 (2.12)	2.9 (2.05)	2.5 (2.01)	2.5 (1.74)	2.5 (1.89)
Hyperactivity score (0–10), *N*	413	415	828	261	204	465	213	182	395
Mean (SD)	6.3 (2.22)	6.2 (2.17)	6.2 (2.20)	4.9 (2.52)	5.3 (2.57)	5.1 (2.55)	4.5 (2.58)	4.9 (2.51)	4.7 (2.56)
Peer problems score (0–10), *N*	413	415	828	261	204	465	213	182	395
Mean (SD)	3.1 (1.87)	3.5 (2.08)	3.3 (1.98)	2.8 (1.95)	3.3 (2.09)	3.0 (2.03)	2.8 (1.87)	3.2 (2.23)	3.0 (2.05)
Impact score (0–10), *N*	413	414	827	180	147	327	153	127	280
Mean (SD)	3.3 (2.41)	3.5 (2.52)	3.4 (2.47)	2.3 (2.13)	2.9 (2.51)	2.6 (2.32)	2.1 (2.33)	2.9 (2.47)	2.5 (2.42)
Externalising score (0–20), *N*	413	415	828	261	204	465	213	182	395
Mean (SD)	10.2 (3.80)	10.1 (3.48)	10.2 (3.65)	7.7 (3.98)	8.3 (4.09)	7.9 (4.03)	7.0 (4.15)	7.4 (3.72)	7.2 (3.96)
Internalising score (0–20), *N*	413	415	828	261	204	465	213	182	395
Mean (SD)	9.4 (3.41)	10.0 (3.63)	9.7 (3.53)	7.8 (3.97)	8.5 (4.02)	8.1 (4.00)	7.2 (4.18)	8.1 (4.28)	7.6 (4.25)
McMaster FAD (1–4)[f]									
Overall FAD score, *N*	404	405	809	258	202	460	213	179	392
Mean (SD)	2.5 (0.33)	2.4 (0.36)	2.5 (0.35)	2.2 (0.41)	2.2 (0.39)	2.2 (0.40)	2.2 (0.45)	2.2 (0.42)	2.2 (0.44)
General functioning, *N*	410	408	818	260	203	463	213	181	394
Mean (SD)	2.5 (0.53)	2.5 (0.56)	2.5 (0.54)	2.2 (0.60)	2.2 (0.58)	2.2 (0.59)	2.1 (0.63)	2.1 (0.59)	2.1 (0.61)

Outcome	Baseline			12 months			18 months		
	FT	TAU	Total	FT	TAU	Total	FT	TAU	Total
Unhealthy (≥ 2.0), *n* (%)	354 (86.3)	339 (83.1)	693 (84.7)	164 (63.1)	130 (64.0)	294 (63.5)	129 (60.6)	107 (59.1)	236 (59.9)
Behaviour control subscale, *N*	413	409	822	260	204	464	213	181	394
Mean (SD)	2.1 (0.38)	2.1 (0.37)	2.1 (0.38)	1.9 (0.40)	1.9 (0.41)	1.9 (0.40)	1.8 (0.47)	1.9 (0.43)	1.8 (0.45)
Unhealthy (≥ 1.9), *n* (%)	319 (77.2)	311 (76.0)	630 (76.6)	149 (57.3)	120 (58.8)	269 (58.0)	111 (52.1)	108 (59.7)	219 (55.6)
Affective involvement subscale, *N*	412	409	821	260	204	464	213	180	393
Mean (SD)	2.5 (0.48)	2.5 (0.49)	2.5 (0.48)	2.3 (0.51)	2.3 (0.48)	2.3 (0.50)	2.2 (0.60)	2.3 (0.55)	2.3 (0.58)
Unhealthy (≥ 2.1), *n* (%)	345 (83.7)	341 (83.4)	686 (83.6)	173 (66.5)	140 (68.6)	313 (67.5)	129 (60.6)	118 (65.6)	247 (62.8)
Affective responsiveness subscale, *N*	410	408	818	259	203	462	213	180	393
Mean (SD)	2.6 (0.48)	2.6 (0.51)	2.6 (0.50)	2.4 (0.60)	2.4 (0.58)	2.4 (0.59)	2.3 (0.61)	2.3 (0.56)	2.3 (0.58)
Unhealthy (≥ 2.2), *n* (%)	343 (83.7)	335 (82.1)	678 (82.9)	177 (68.3)	145 (71.4)	322 (69.7)	136 (63.8)	129 (71.7)	265 (67.4)
Roles subscale, *N*	412	409	821	260	204	464	213	181	394
Mean (SD)	2.5 (0.35)	2.5 (0.37)	2.5 (0.36)	2.3 (0.41)	2.3 (0.40)	2.3 (0.40)	2.3 (0.47)	2.3 (0.42)	2.3 (0.45)
Unhealthy (≥ 2.3), *n* (%)	310 (75.2)	309 (75.6)	619 (75.4)	143 (55.0)	118 (57.8)	261 (56.3)	112 (52.6)	95 (52.5)	207 (52.5)
Communication subscale, *N*	413	409	822	260	203	463	213	181	394
Mean (SD)	2.6 (0.37)	2.6 (0.38)	2.6 (0.37)	2.4 (0.44)	2.3 (0.44)	2.4 (0.44)	2.3 (0.47)	2.3 (0.45)	2.3 (0.46)
Unhealthy (≥ 2.2), *n* (%)	360 (87.2)	343 (83.9)	703 (85.5)	190 (73.1)	136 (67.0)	326 (70.4)	142 (66.7)	122 (67.4)	264 (67.0)
Problem-solving subscale, *N*	410	409	819	259	204	463	213	180	393
Mean (SD)	2.5 (0.47)	2.5 (0.53)	2.5 (0.50)	2.3 (0.52)	2.3 (0.53)	2.3 (0.52)	2.2 (0.53)	2.2 (0.50)	2.2 (0.51)
Unhealthy (≥ 2.2), *n* (%)	331 (80.7)	306 (74.8)	637 (77.8)	164 (63.3)	128 (62.7)	292 (63.1)	117 (54.9)	108 (60.0)	225 (57.3)

Q, question.
a Higher scores represent greater levels of depression.
b Higher scores indicative of greater enjoyment and satisfaction.
c Higher scores indicate a higher level of suicide ideation. Median presented as Beck scores considerably skewed and zero inflated at follow-up.
d Higher scores reflect greater hopelessness or negative expectations towards the future.
e Higher scores in all but the prosocial score represent greater issues in that category. For the prosocial score, lower scores represent greater issues. See *Appendix 3* for categorisation of the total difficulties score (differs for young person and caregiver).
f Higher scores are indicative of poorer family functioning.

TABLE 70 Mean scores with 95% CIs for young person questionnaire outcomes (adjusted for baseline score and covariates) via multiple imputation ($n = 832$)

Outcome	12 months			18 months		
	FT, mean (95% CI), SE	TAU, mean (95% CI), SE	Difference,[f] mean (95% CI), SE; *p*-value	FT, mean (95% CI), SE	TAU, mean (95% CI), SE	Difference,[f] mean (95% CI), SE; *p*-value
CDRS-R[a]	***33.2 (30.4 to 36.1), SE 1.46***	***33.9 (30.8 to 37.0), SE 1.57***	***–0.6 (–3.1 to 1.9), SE 1.27; p = 0.6170***	***30.6 (27.6 to 33.6), SE 1.50***	***31.6 (28.7 to 34.5), SE 1.46***	***–1.0 (–3.5 to 1.5), SE 1.26; p = 0.4250***
PQ-LES-Q[b]	***49.9 (47.7 to 52.1), SE 1.12***	***48.8 (46.5 to 51.0), SE 1.13***	***1.1 (–0.5 to 2.7), SE 0.82; p = 0.1774***	***50.6 (48.4 to 52.8), SE 1.12***	***50.4 (48.1 to 52.8), SE 1.20***	***0.1 (–1.9 to 2.1), SE 1.02; p = 0.9013***
Hopelessness[c]	***4.8 (4.0 to 5.6), SE 0.40***	***5.1 (4.3 to 6.0), SE 0.43***	***–0.3 (–1.1 to 0.4), SE 0.37; p = 0.3809***	***4.4 (3.6 to 5.2), SE 0.42***	***4.6 (3.7 to 5.4), SE 0.43***	***–0.2 (–0.9 to 0.5), SE 0.36; p = 0.6327***
SDQ[d]						
Total difficulties	14.8 (13.4 to 16.1), SE 0.69	15.5 (14.1 to 16.9), SE 0.70	–0.7 (–1.8 to 0.4), SE 0.54; $p = 0.1885$	13.3 (12.0 to 14.6), SE 0.67	14.1 (12.7 to 15.5), SE 0.71	–0.8 (–2.0 to 0.4), SE 0.61; $p = 0.1845$
Prosocial score	7.7 (7.3 to 8.1), SE 0.19	7.3 (6. to, 7.7), SE 0.19	0.4 (0.1 to 0.7), SE 0.15; $p = 0.0064$	7.8 (7.4 to 8.1), SE 0.18	7.4 (7.1 to 7.8), SE 0.19	0.3 (0.0 to 0.7), SE 0.16; $p = 0.0337$
Emotional problems score	4.4 (3.8 to 4.9), SE 0.28	4.3 (3.7 to 4.8), SE 0.28	0.1 (–0.4 to 0.5), SE 0.22; $p = 0.8098$	3.7 (3.2 to 4.3), SE 0.27	4.0 (3.5 to 4.6), SE 0.28	–0.3 (–0.8 to 0.2), SE 0.24; $p = 0.2002$
Conduct problems score	2.7 (2.4 to 3.1), SE 0.19	2.9 (2.5 to 3.3), SE 0.19	–0.2 (–0.5 to 0.1), SE 0.15; $p = 0.2059$	2.4 (2.0 to 2.7), SE 0.18	2.4 (2.0 to 2.8), SE 0.19	0.0 (–0.3 to 0.3), SE 0.16; $p = 0.8964$
Hyperactivity score	5.0 (4.5 to 5.5), SE 0.25	5.3 (4.8 to 5.9), SE 0.26	–0.3 (–0.7 to 0.1), SE 0.20; $p = 0.0978$	4.5 (4.1 to 5.0), SE 0.24	4.9 (4.4 to 5.4), SE 0.26	–0.4 (–0.8 to 0.1), SE 0.22; $p = 0.1090$
Peer problems score	2.6 (2.2 to 3.0), SE 0.20	2.9 (2.5 to 3.3), SE 0.20	–0.3 (–0.6 to 0.0), SE 0.16; $p = 0.0501$	2.5 (2.1 to 2.9), SE 0.20	2.8 (2.4 to 3.2), SE 0.21	–0.3 (–0.6 to 0.1), SE 0.18; $p = 0.1346$
Impact score	2.0 (1.5 to 2.6), SE 0.27	2.7 (2.2 to 3.2), SE 0.25	–0.7 (–1.1 to –0.2), SE 0.22; $p = 0.0033$	1.9 (1.3 to 2.5), SE 0.29	2.2 (1.6 to 2.8), SE 0.30	–0.3 (–0.8 to 0.2), SE 0.26, $p = 0.2153$
Externalising score	7.7 (7.0 to 8.5), SE 0.38	8.2 (7.5 to 9.0), SE 0.39	–0.5 (–1.1 to 0.1), SE 0.30; $p = 0.0817$	6.9 (6.2 to 7.6), SE 0.37	7.2 (6.5 to 8.0), SE 0.39	–0.3 (–1.0 to 0.3), SE 0.33; $p = 0.3196$
Internalising score	7.0 (6.2 to 7.8), SE 0.40	7.2 (6.4 to 8.0), SE 0.41	–0.2 (–0.9 to 0.4), SE 0.32; $p = 0.4809$	6.3 (5.5 to 7.1), SE 0.39	6.8 (6.0 to 7.6), SE 0.41	–0.5 (–1.2 to 0.2), SE 0.36; $p = 0.1373$

Outcome	12 months			18 months		
	FT, mean (95% CI), SE	TAU, mean (95% CI), SE	Difference,[f] mean (95% CI), SE; *p*-value	FT, mean (95% CI), SE	TAU, mean (95% CI), SE	Difference,[f] mean (95% CI), SE; *p*-value
McMaster FAD[e]						
Overall FAD score	2.3 (2.2 to 2.3), SE 0.04	2.3 (2.2 to 2.4), SE 0.04	–0.0 (–0.1 to 0.0), SE 0.03; $p = 0.6720$	2.2 (2.1 to 2.3), SE 0.04	2.2 (2.1 to 2.3), SE 0.04	0.0 (–0.1 to 0.1), SE 0.03; $p = 0.9321$
General functioning	2.2 (2.1 to 2.3), SE 0.06	2.2 (2.1 to 2.4), SE 0.06	–0.0 (–0.1 to 0.1), SE 0.05; $p = 0.7364$	2.2 (2.1 to 2.3), SE 0.06	2.2 (2.0 to 2.3), SE 0.06	0.0 (–0.1 to 0.1), SE 0.05; $p = 0.8115$
Behaviour control	2.0 (1.9 to 2.0), SE 0.04	1.9 (1.9 to 2.0), SE 0.04	0.0 (–0.0 to 0.1), SE 0.03; $p = 0.6267$	1.9 (1.8 to 2.0), SE 0.04	1.9 (1.8 to 2.0), SE 0.05	–0.0 (–0.1 to 0.1), SE 0.04; $p = 0.5654$
Affective involvement	2.4 (2.3 to 2.5), SE 0.05	2.4 (2.3 to 2.5), SE 0.05	–0.1 (–0.1 to 0.0), SE 0.04; $p = 0.2246$	2.3 (2.2 to 2.4), SE 0.05	2.4 (2.3 to 2.5), SE 0.06	–0.1 (–0.2 to 0.0), SE 0.05; $p = 0.1632$
Affective responsiveness	2.4 (2.3 to 2.5), SE 0.06	2.4 (2.3 to 2.6), SE 0.06	–0.0 (–0.1 to 0.1), SE 0.05; $p = 0.6644$	2.4 (2.3 to 2.5), SE 0.06	2.4 (2.3 to 2.5), SE 0.06	–0.0 (–0.1 to 0.1), SE 0.05; $p = 0.9381$
Roles	2.3 (2.2 to 2.4), SE 0.04	2.3 (2.2 to 2.4), SE 0.04	–0.0 (–0.1 to 0.0), SE 0.03; $p = 0.4592$	2.2 (2.2 to 2.3), SE 0.04	2.2 (2.2 to 2.3), SE 0.04	–0.0 (–0.1 to 0.1), SE 0.04; $p = 0.8653$
Communication	2.4 (2.3 to 2.5), SE 0.04	2.4 (2.3 to 2.4), SE 0.04	0.0 (–0.0 to 0.1), SE 0.04; $p = 0.2562$	2.3 (2.2 to 2.4), SE 0.05	2.3 (2.2 to 2.4), SE 0.05	0.0 (–0.1 to 0.1), SE 0.04; $p = 0.7817$
Problem-solving	2.3 (2.2 to 2.4), SE 0.05	2.3 (2.2 to 2.4), SE 0.05	–0.0 (–0.1 to 0.1), SE 0.04; $p = 0.8091$	2.2 (2.1 to 2.3), SE 0.05	2.2 (2.1 to 2.3), SE 0.05	–0.0 (–0.1 to 0.1), SE 0.05; $p = 0.6052$

a CDRS-R: higher scores represent greater levels of depression.
b PQ-LES-Q: higher scores indicative of greater enjoyment and satisfaction.
c Hopelessness scale: higher scores reflect greater hopelessness or negative expectations towards the future.
d SDQ: higher scores in all but the prosocial score represent greater issues in that category. For the prosocial score, lower scores represent greater issues.
e McMaster FAD: higher scores are indicative of poorer family functioning.
f Difference: FT – TAU.

TABLE 71 Mean proportions, odds and odds ratios with 95% CIs for suicide ideation as screened by the young person Beck scale (adjusted for baseline score and covariates) via multiple imputation ($n = 832$)

	Mean proportion (95% CI), SE		Odds (95% CI)		OR (95% CI), *p*-value
Time point	FT	TAU	FT	TAU	FT vs. TAU
12 months	0.26 (0.17 to 0.36), SE = 0.05	0.36 (0.25 to 0.46), SE = 0.05	0.35 (0.22 to 0.58)	0.55 (0.35 to 0.87)	0.64 (0.44 to 0.94), $p = 0.0242$
18 months	0.22 (0.14 to 0.31), SE = 0.04	0.28 (0.18 to 0.37), SE = 0.05	0.29 (0.18 to 0.47)	0.38 (0.23 to 0.62)	0.76 (0.49 to 1.16), $p = 0.2009$

OR, odds ratio.

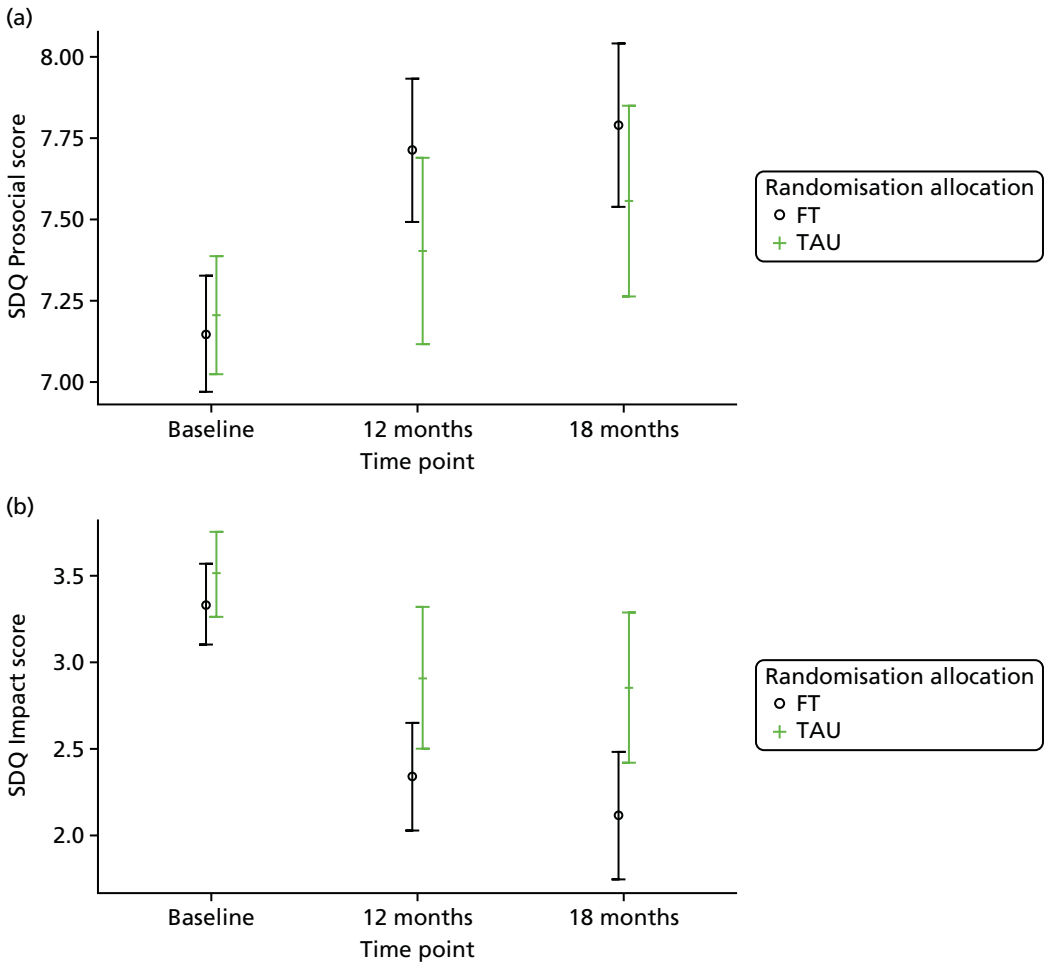


FIGURE 23 Unadjusted 95% CIs for young person SDQ prosocial and impact subscales. (a) prosocial and (b) impact.

Caregiver questionnaires

Summary statistics for the caregiver questionnaire outcomes are presented in *Table 72* and results of the repeated measures modelling (adjusted for baseline scores and covariates) in *Table 73*. The results of the sensitivity analysis conducted using complete data were similar and did not result in any change to the conclusions.

Caregiver questionnaire outcomes on the SDQ and the McMaster FAD differed significantly between the treatment groups. No significant differences between treatments were detected for the GHQ-12 or the Family Questionnaire.

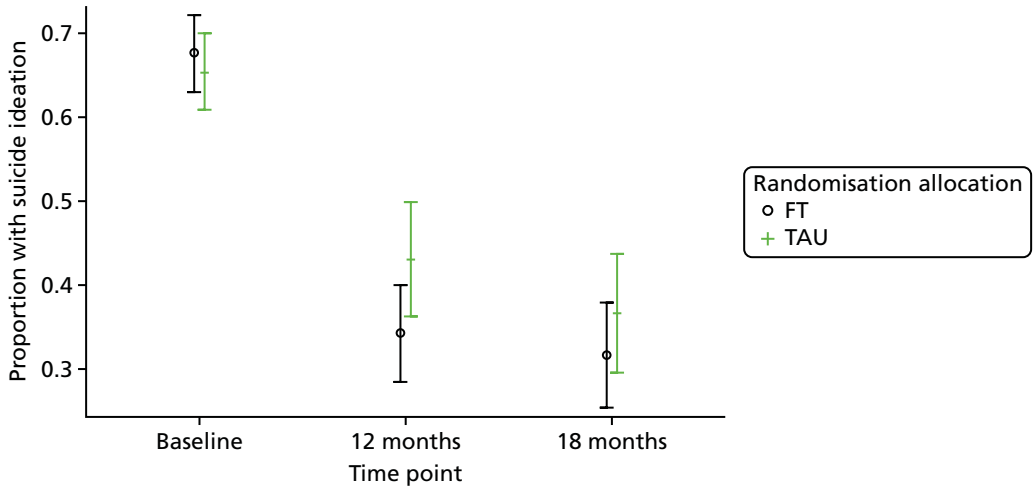


FIGURE 24 Unadjusted 95% CIs for suicide ideation as reported by the young person in the BSS.

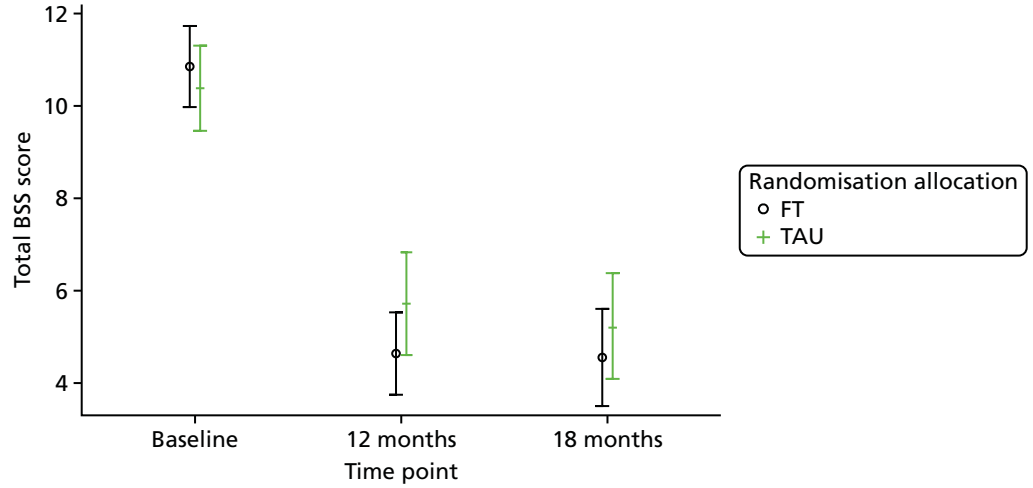


FIGURE 25 Unadjusted 95% CIs for the young person BSS score.

On the SDQ, there was good evidence (p-value < 0.05) of significantly better outcomes in the FT arm on the total difficulties score and on the emotional problems, peer problems and internalising subscores at both 12 and 18 months, and on the conduct problems and externalising subscores at 18 months only. There was also good evidence of better outcomes in the FT arm on the impact score at 12 months; however, by 18 months, scores in the FT arm had risen slightly so that the difference between treatment arms was no longer statistically significant.

On the McMaster FAD, there was good evidence (p-value < 0.05) of better family functioning on the roles subscale at 12 months; however, by 18 months, scores remained similar in the FT arm but decreased to a similar level in the TAU arm so that the difference between treatment arms was no longer statistically significant.

Unadjusted mean estimates and 95% CIs for complete data are presented in *Figures 26*, *27* and *28* for questionnaires and subscales with good evidence of a treatment effect at 12 or 18 months (p-value < 0.05).

TABLE 72 Summary statistics for the caregiver questionnaire outcomes

	Baseline			12 months			18 months		
Outcome	FT	TAU	Total	FT	TAU	Total	FT	TAU	Total
***GHQ-12 score. Likert scale (0–36),*[a] N**									
n	414	415	829	254	193	447	218	173	391
Mean (SD)	17.7 (7.06)	18.6 (7.24)	18.2 (7.16)	12.6 (6.93)	14.0 (6.15)	13.2 (6.63)	13.0 (6.48)	13.6 (6.75)	13.3 (6.60)
95% CI	(17.1 to 18.4)	(17.9 to 19.3)	(17.7 to 18.7)	(11.8 to 13.5)	(13.1 to 14.9)	(12.6 to 13.8)	(12.1 to 13.9)	(12.6 to 14.6)	(12.6 to 13.9)
***Family Questionnaire,*[b] N**	***415***	***416***	***831***	***241***	***198***	***439***	***210***	***150***	***360***
Total score (20–80), mean (SD)	52.9 (10.67)	52.9 (10.85)	52.9 (10.75)	49.4 (10.88)	48.3 (10.65)	48.9 (10.78)	45.5 (10.78)	46.9 (12.36)	46.1 (11.47)
Emotional overinvolvement (10–40), mean (SD)	27.2 (4.93)	27.2 (5.03)	27.2 (4.98)	25.3 (5.26)	24.9 (5.35)	25.1 (5.30)	23.1 (5.37)	23.6 (6.15)	23.3 (5.71)
Criticism (10–40), mean (SD)	25.7 (6.97)	25.7 (7.06)	25.7 (7.01)	24.2 (6.65)	23.5 (6.43)	23.8 (6.55)	22.4 (6.28)	23.2 (7.17)	22.8 (6.67)
***SDQ*[c]**									
Total difficulties score (0–40), N	*412*	*415*	*827*	*253*	*194*	*447*	*215*	*173*	*388*
Mean (SD)	19.4 (6.56)	19.8 (6.83)	19.6 (6.69)	14.2 (7.23)	15.3 (7.44)	14.7 (7.33)	13.0 (7.14)	15.0 (7.34)	13.9 (7.29)
Close to average (0–13), *n* (%)	78 (18.9)	86 (20.7)	164 (19.8)	125 (49.4)	81 (41.8)	206 (46.1)	115 (53.5)	75 (43.4)	190 (49.0)
Slightly raised (14–16), *n* (%)	51 (12.4)	36 (8.7)	87 (10.5)	36 (14.2)	31 (16.0)	67 (15.0)	29 (13.5)	34 (19.7)	63 (16.2)
High (17–19), *n* (%)	67 (16.3)	66 (15.9)	133 (16.1)	27 (10.7)	28 (14.4)	55 (12.3)	35 (16.3)	19 (11.0)	54 (13.9)
Very high (20–40), *n* (%)	216 (52.4)	227 (54.7)	443 (53.6)	65 (25.7)	54 (27.8)	119 (26.6)	36 (16.7)	45 (26.0)	81 (20.9)
Prosocial score (0–10), N	*414*	*416*	*830*	*254*	*194*	*448*	*216*	*173*	*389*
Mean (SD)	6.3 (2.33)	6.3 (2.30)	6.3 (2.31)	6.9 (2.22)	6.8 (2.21)	6.8 (2.22)	7.1 (2.20)	6.9 (2.40)	7.0 (2.29)
Emotional problems score (0–10), N	*414*	*415*	*829*	*253*	*194*	*447*	*217*	*173*	*390*
Mean (SD)	6.2 (2.39)	6.2 (2.60)	6.2 (2.49)	4.3 (2.72)	4.8 (2.76)	4.5 (2.75)	3.9 (2.74)	4.5 (2.66)	4.2 (2.72)
Conduct problems score (0–10), N	*413*	*416*	*829*	*254*	*194*	*448*	*216*	*174*	*390*
Mean (SD)	4.2 (2.42)	4.3 (2.47)	4.2 (2.44)	2.9 (2.21)	3.1 (2.23)	3.0 (2.22)	2.5 (2.12)	2.9 (2.21)	2.7 (2.17)

Outcome	Baseline			12 months			18 months		
	FT	TAU	Total	FT	TAU	Total	FT	TAU	Total
Hyperactivity score (0–10), N	*414*	*415*	*829*	*254*	*194*	*448*	*217*	*173*	*390*
Mean (SD)	5.5 (2.48)	5.7 (2.56)	5.6 (2.52)	4.1 (2.50)	4.3 (2.55)	4.2 (2.52)	4.1 (2.57)	4.5 (2.50)	4.3 (2.55)
Peer problems score (0–10), N	*413*	*415*	*828*	*253*	*194*	*447*	*215*	*173*	*388*
Mean (SD)	3.5 (2.09)	3.6 (2.14)	3.6 (2.11)	2.8 (2.06)	3.1 (2.06)	3.0 (2.06)	2.6 (1.79)	3.1 (2.27)	2.8 (2.03)
Impact score (0–10), N	*410*	*413*	*823*	*193*	*145*	*338*	*133*	*127*	*260*
Mean (SD)	4.4 (2.73)	4.3 (2.73)	4.4 (2.73)	2.6 (2.67)	3.2 (2.91)	2.8 (2.79)	2.8 (2.81)	3.2 (2.91)	3.0 (2.86)
Externalising score (0–20), N	*413*	*415*	*828*	*254*	*194*	*448*	*216*	*173*	*389*
Mean (SD)	9.7 (4.30)	10.0 (4.50)	9.9 (4.40)	7.0 (4.13)	7.4 (4.30)	7.2 (4.20)	6.5 (4.22)	7.4 (4.28)	6.9 (4.26)
Internalising score (0–20), N	*413*	*415*	*828*	*253*	*194*	*447*	*215*	*173*	*388*
Mean (SD)	9.8 (3.69)	9.8 (4.02)	9.8 (3.86)	7.2 (4.10)	7.9 (4.31)	7.5 (4.21)	6.5 (3.88)	7.6 (4.32)	7.0 (4.12)
McMaster FAD (1–4)[d]									
Overall FAD score, N	*408*	*415*	*823*	*251*	*193*	*444*	*215*	*174*	*389*
Mean (SD)	2.2 (0.36)	2.2 (0.36)	2.2 (0.36)	2.0 (0.36)	2.1 (0.35)	2.0 (0.36)	2.0 (0.38)	2.0 (0.38)	2.0 (0.38)
General functioning, N	*415*	*416*	*831*	*253*	*194*	*447*	*216*	*174*	*390*
Mean (SD)	2.3 (0.48)	2.3 (0.47)	2.3 (0.47)	2.0 (0.46)	2.1 (0.44)	2.1 (0.45)	2.0 (0.47)	2.0 (0.47)	2.0 (0.47)
Unhealthy (≥ 2.0), *n* (%)	314 (75.7)	316 (76.0)	630 (75.8)	148 (58.5)	115 (59.3)	263 (58.8)	120 (55.6)	96 (55.2)	216 (55.4)
Behaviour control, N	*411*	*416*	*827*	*251*	*194*	*445*	*215*	*174*	*389*
Mean (SD)	1.8 (0.41)	1.8 (0.41)	1.8 (0.41)	1.7 (0.40)	1.7 (0.39)	1.7 (0.40)	1.7 (0.42)	1.7 (0.42)	1.7 (0.42)
Unhealthy (≥ 1.9), *n* (%)	197 (47.9)	226 (54.3)	423 (51.1)	103 (41.0)	71 (36.6)	174 (39.1)	75 (34.9)	60 (34.5)	135 (34.7)
Affective involvement, N	*412*	*415*	*827*	*252*	*194*	*446*	*215*	*174*	*389*
Mean (SD)	2.2 (0.45)	2.3 (0.51)	2.2 (0.48)	2.1 (0.45)	2.1 (0.45)	2.1 (0.45)	2.1 (0.44)	2.1 (0.50)	2.1 (0.47)
Unhealthy (≥ 2.1), *n* (%)	270 (65.5)	272 (65.5)	542 (65.5)	135 (53.6)	107 (55.2)	242 (54.3)	119 (55.3)	92 (52.9)	211 (54.2)

continued

TABLE 72 Summary statistics for the caregiver questionnaire outcomes (*continued*)

	Baseline			12 months			18 months		
Outcome	FT	TAU	Total	FT	TAU	Total	FT	TAU	Total
Affective responsiveness, N	*411*	*415*	*826*	*252*	*194*	*446*	*216*	*174*	*390*
Mean (SD)	2.1 (0.55)	2.1 (0.57)	2.1 (0.56)	1.9 (0.54)	2.0 (0.54)	2.0 (0.54)	1.9 (0.53)	1.9 (0.51)	1.9 (0.52)
Unhealthy (≥ 2.2), *n* (%)	195 (47.4)	197 (47.5)	392 (47.5)	98 (38.9)	69 (35.6)	167 (37.4)	83 (38.4)	64 (36.8)	147 (37.7)
Roles, N	*415*	*415*	*830*	*254*	*194*	*448*	*217*	*174*	*391*
Mean (SD)	2.5 (0.42)	2.5 (0.42)	2.5 (0.42)	2.3 (0.41)	2.4 (0.40)	2.4 (0.41)	2.3 (0.43)	2.3 (0.43)	2.3 (0.43)
Unhealthy (≥ 2.3), *n* (%)	304 (73.3)	309 (74.5)	613 (73.9)	152 (59.8)	138 (71.1)	290 (64.7)	120 (55.3)	104 (59.8)	224 (57.3)
Communication, N	*413*	*415*	*828*	*254*	*193*	*447*	*215*	*174*	*389*
Mean (SD)	2.3 (0.44)	2.3 (0.41)	2.3 (0.43)	2.1 (0.41)	2.1 (0.42)	2.1 (0.42)	2.1 (0.43)	2.1 (0.44)	2.1 (0.44)
Unhealthy (≥ 2.2), *n* (%)	248 (60.0)	255 (61.4)	503 (60.7)	99 (39.0)	92 (47.7)	191 (42.7)	92 (42.8)	71 (40.8)	163 (41.9)
Problem-solving, N	*411*	*416*	*827*	*252*	*193*	*445*	*216*	*174*	*390*
Mean (SD)	2.2 (0.48)	2.2 (0.48)	2.2 (0.48)	2.0 (0.45)	2.1 (0.45)	2.0 (0.45)	2.0 (0.46)	2.0 (0.43)	2.0 (0.45)
Unhealthy (≥ 2.2), *n* (%)	234 (56.9)	249 (59.9)	483 (58.4)	94 (37.3)	84 (43.5)	178 (40.0)	84 (38.9)	76 (43.7)	160 (41.0)

a Higher scores are indicative of greater psychological distress.
b Higher scores indicate greater levels of expressed emotion.
c Higher scores in all but the prosocial score represent greater issues in that category. For the prosocial score, lower scores represent greater issues. See *Appendix 3* for categorisation of the total difficulties score (differs for young person and caregiver).
d Higher scores are indicative of poorer family functioning.

TABLE 73 Mean scores with 95% CIs for caregivers questionnaire outcomes (adjusted for baseline and covariates) via multiple imputation ($n = 832$)

Outcome	12 months			18 months		
	FT, mean (95% CI), SE	TAU, mean (95% CI), SE	Difference,[e] mean (95% CI), SE; *p*-value	FT, mean (95% CI), SE	TAU, mean (95% CI), SE	Difference,[e] mean (95% CI), SE; *p*-value
GHQ-12[a]	***12.8 (11.6 to 14.0), SE = 0.61***	***13.5 (12.3 to 14.8), SE = 0.65***	***–0.7 (–1.8 to 0.3), SE = 0.54; p = 0.1870***	***13.0 (11.8 to 14.2), SE = 0.62***	***13.2 (11.8 to 14.6), SE = 0.71***	***–0.2 (–1.3 to 0.9), SE = 0.57; p = 0.7307***
Family Questionnaire (3 and 6 months)[b]						
Total score	50.9 (49.0 to 52.7), SE = 0.96	50.2 (48.3 to 52.0), SE = 0.95	0.7 (–0.7 to 2.1), SE = 0.73; *p* = 0.3400	47.4 (45.3 to 49.4), SE = 1.06	48.8 (46.7 to 50.9), SE = 1.07	–1.4 (–3.3 to 0.5), SE = 0.94; *p* = 0.1369
Emotional subscore	25.9 (24.9 to 26.9), SE = 0.50	25.5 (24.6 to 26.5), SE = 0.50	0.4 (–0.4 to 1.1), SE = 0.38; *p* = 0.3533	23.8 (22.7 to 24.9), SE = 0.55	24.4 (23.3 to 25.5), SE = 0.56	–0.6 (–1.6 to 0.4), SE = 0.50; *p* = 0.2614
Criticism subscore	25.0 (23.9 to 26.1), SE = 0.55	24.7 (23.6 to 25.8), SE = 0.55	0.3 (–0.5 to 1.2), SE = 0.43; *p* = 0.4385	23.6 (22.4 to 24.8), SE = 0.61	24.4 (23.2 to 25.6), SE = 0.61	–0.9 (–1.9 to 0.2), SE = 0.54; *p* = 0.1150
SDQ[c]						
Total difficulties	14.1 (12.7 to 15.5), SE = 0.72	15.4 (14.0 to 16.8), SE = 0.71	–1.3 (–2.4 to –0.2), SE = 0.56; *p* = 0.0260	13.2 (11.9 to 14.6), SE = 0.68	14.9 (13.3 to 16.4), SE = 0.79	–1.6 (–2.9 to –0.4), SE = 0.65; *p* = 0.0131
Prosocial score	6.9 (6.5 to 7.3), SE = 0.21	6.8 (6.4 to 7.2), SE = 0.22	0.1 (–0.3 to 0.4), SE = 0.17; *p* = 0.6005	7.1 (6.6 to 7.5), SE = 0.22	6.9 (6.5 to 7.4), SE = 0.22	0.2 (–0.2 to 0.5), SE = 0.18; *p* = 0.3877
Emotional problems score	4.0 (3.4 to 4.5), SE = 0.28	4.5 (3.9 to 5.1), SE = 0.29	–0.5 (–1.0 to –0.1), SE = 0.23; *p* = 0.0166	3.6 (3.1 to 4.2), SE = 0.28	4.2 (3.6 to 4.8), SE = 0.30	–0.6 (–1.1 to –0.1), SE = 0.25; *p* = 0.0218
Conduct problems score	3.1 (2.7 to 3.4), SE = 0.20	3.3 (2.9 to 3.7), SE = 0.20	–0.3 (–0.6 to 0.0), SE = 0.17; *p* = 0.0925	2.8 (2.4 to 3.2), SE = 0.20	3.1 (2.7 to 3.5), SE = 0.21	–0.3 (–0.6 to -0.0), SE = 0.16; *p* = 0.0499
Hyperactivity score	4.3 (3.8 to 4.7), SE = 0.24	4.4 (3.9 to 4.9), SE = 0.25	–0.1 (–0.5 to 0.3), SE = 0.19; *p* = 0.5494	4.3 (3.8 to 4.8), SE = 0.24	4.5 (4.0 to 5.1), SE = 0.28	–0.2 (–0.6 to 0.2), SE = 0.22; *p* = 0.3536
Peer problems score	2.9 (2.5 to 3.3), SE = 0.21	3.2 (2.8 to 3.6), SE = 0.21	–0.3 (–0.7 to 0.0), SE = 0.17; *p* = 0.0366	2.6 (2.2 to 3.0), SE = 0.20	3.1 (2.7 to 3.6), SE = 0.23	–0.5 (–0.9 to –0.1), SE = 0.20; *p* = 0.0092
Impact score	2.3 (1.7 to 2.9), SE = 0.31	2.9 (2.2 to 3.7), SE = 0.37	–0.7 (–1.3 to -0.1), SE = 0.30; *p* = 0.0309	2.4 (1.7 to 3.0), SE = 0.32	2.6 (1.9 to 3.3), SE = 0.36	–0.3 (–0.9 to 0.3), SE = 0.31; *p* = 0.3844
Externalising score	7.3 (6.6 to 8.0), SE = 0.36	7.7 (6.9 to 8.4), SE = 0.38	–0.4 (–1.0 to 0.2), SE = 0.31; *p* = 0.1827	7.0 (6.2 to 7.8), SE = 0.39	7.6 (6.9 to 8.4), SE = 0.39	–0.7 (–1.3 to 0.0), SE = 0.32; *p* = 0.0446

continued

TABLE 73 Mean scores with 95% CIs for caregivers questionnaire outcomes (adjusted for baseline and covariates) via multiple imputation ($n = 832$) (*continued*)

Outcome	12 months			18 months		
	FT, mean (95% CI), SE	TAU, mean (95% CI), SE	Difference,[e] mean (95% CI), SE; *p*-value	FT, mean (95% CI), SE	TAU, mean (95% CI), SE	Difference,[e] mean (95% CI), SE; *p*-value
Internalising score	6.8 (6.0 to 7.7), SE = 0.43	7.7 (6.9 to 8.6), SE = 0.43	–0.9 (–1.5 to –0.2), SE = 0.34; $p = 0.0111$	6.2 (5.4 to 7.0), SE = 0.41	7.3 (6.3 to 8.2), SE = 0.48	–1.1 (–1.9 to –0.3), SE = 0.39; $p = 0.0074$
McMaster FAD[d]						
Overall FAD score	2.1 (2.0 to 2.1), SE = 0.03	2.1 (2.0 to 2.2), SE = 0.03	–0.0 (–0.1 to 0.0), SE = 0.03; $p = 0.0797$	2.0 (2.0 to 2.1), SE = 0.03	2.1 (2.0 to 2.1), SE = 0.03	–0.0 (–0.1 to 0.0), SE = 0.03; $p = 0.6436$
General functioning	2.1 (2.0 to 2.2), SE = 0.04	2.1 (2.0 to 2.2), SE = 0.04	–0.0 (–0.1 to 0.0), SE = 0.04; $p = 0.2444$	2.0 (1.9 to 2.1), SE = 0.05	2.1 (2.0 to 2.2), SE = 0.04	–0.0 (–0.1 to 0.0), SE = 0.04; $p = 0.2085$
Behaviour control	1.7 (1.6 to 1.8), SE = 0.04	1.7 (1.6 to 1.8), SE = 0.04	–0.0 (–0.1 to 0.1), SE = 0.03; $p = 0.8728$	1.7 (1.6 to 1.7), SE = 0.04	1.7 (1.6 to 1.8), SE = 0.04	–0.0 (–0.1 to 0.1), SE = 0.03; $p = 0.8188$
Affective involvement	2.1 (2.0 to 2.2), SE = 0.04	2.1 (2.1 to 2.2), SE = 0.04	–0.1 (–0.1 to 0.0), SE = 0.03; $p = 0.0648$	2.1 (2.0 to 2.2), SE = 0.04	2.1 (2.0 to 2.2), SE = 0.05	–0.0 (–0.1 to 0.1), SE = 0.05; $p = 0.5652$
Affective responsiveness	2.0 (1.9 to 2.1), SE = 0.04	2.0 (1.9 to 2.1), SE = 0.05	–0.0 (–0.1 to 0.1), SE = 0.04; $p = 0.6716$	2.0 (1.9 to 2.1), SE = 0.05	1.9 (1.9 to 2.0), SE = 0.05	0.0 (–0.1 to 0.1), SE = 0.04; $p = 0.4294$
Roles	2.4 (2.3 to 2.4), SE = 0.03	2.5 (2.4 to 2.5), SE = 0.04	–0.1 (–0.2 to 0.0), SE = 0.03; $p = 0.0020$	2.4 (2.3 to 2.4), SE = 0.04	2.4 (2.3 to 2.5), SE = 0.04	–0.0 (–0.1 to 0.1), SE = 0.03; $p = 0.7299$
Communication	2.1 (2.0 to 2.2), SE = 0.04	2.2 (2.1 to 2.3), SE = 0.04	–0.0 (–0.1 to 0.0), SE = 0.03; $p = 0.1282$	2.1 (2.0 to 2.2), SE = 0.04	2.1 (2.1 to 2.2), SE = 0.04	–0.0 (–0.1 to 0.0), SE = 0.04; $p = 0.2939$
Problem-solving	2.0 (2.0 to 2.1), SE = 0.04	2.1 (2.0 to 2.2), SE = 0.04	–0.0 (–0.1 to 0.0), SE = 0.03; $p = 0.2628$	2.0 (1.9 to 2.1), SE = 0.04	2.0 (1.9 to 2.1), SE = 0.04	–0.0 (–0.1 to 0.1), SE = 0.04; $p = 0.7037$

a GHQ-12: higher scores are indicative of greater psychological distress.
b Family Questionnaire: higher scores indicate greater levels of expressed emotion.
c SDQ: higher scores in all but the prosocial score represent greater issues in that category. For the prosocial score, lower scores represent greater issues.
d McMaster FAD: higher scores are indicative of poorer family functioning.
e Difference: FT – TAU.

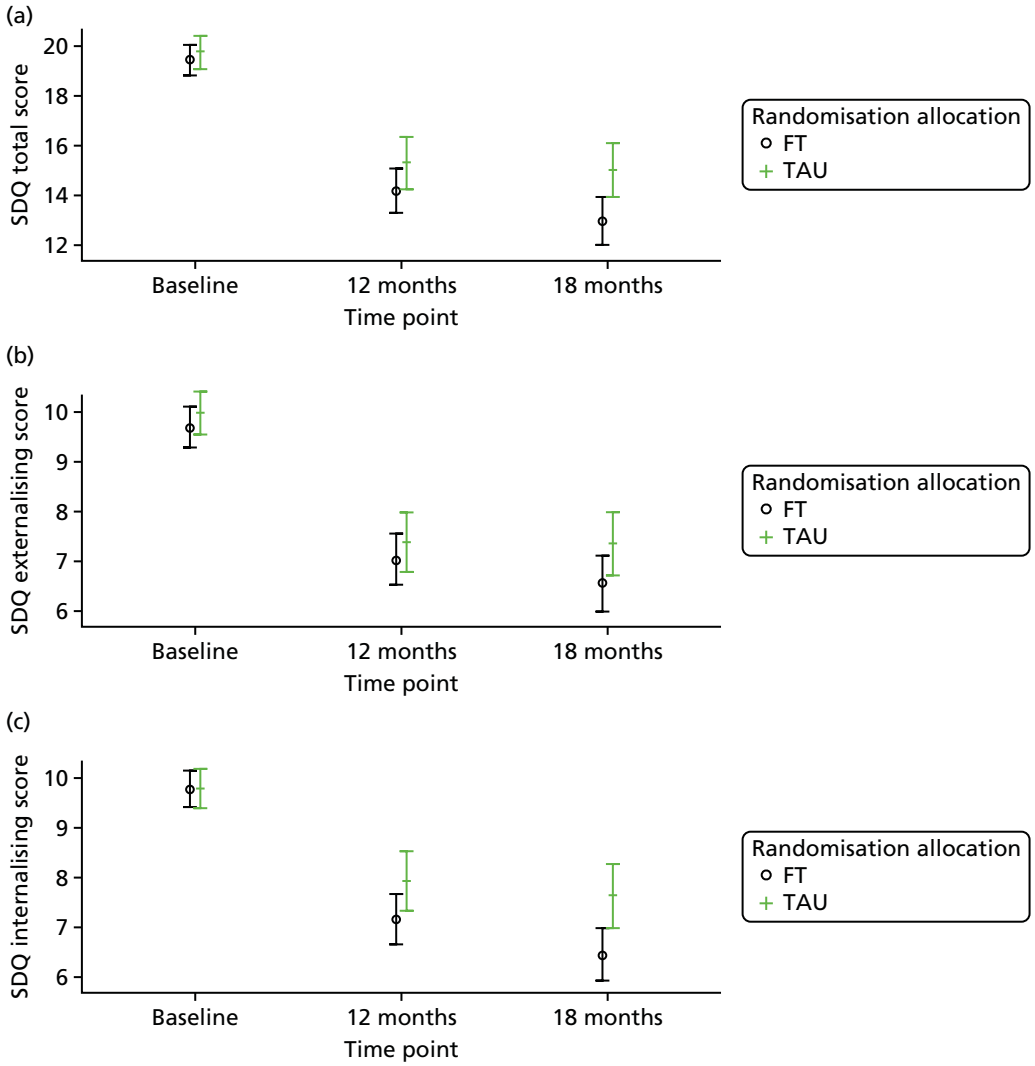


FIGURE 26 Unadjusted 95% CIs for caregiver SDQ (a) total difficulties; (b) externalising; and (c) internalising.

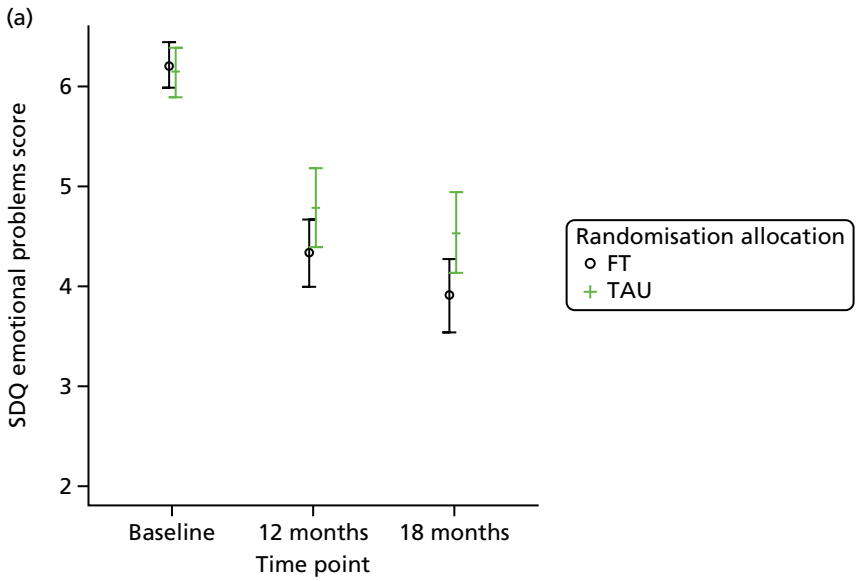


FIGURE 27 Unadjusted 95% CIs for caregiver SDQ (a) emotional problems; (b) conduct problems; (c) peer problems; and (d) impact problems. (*continued*)

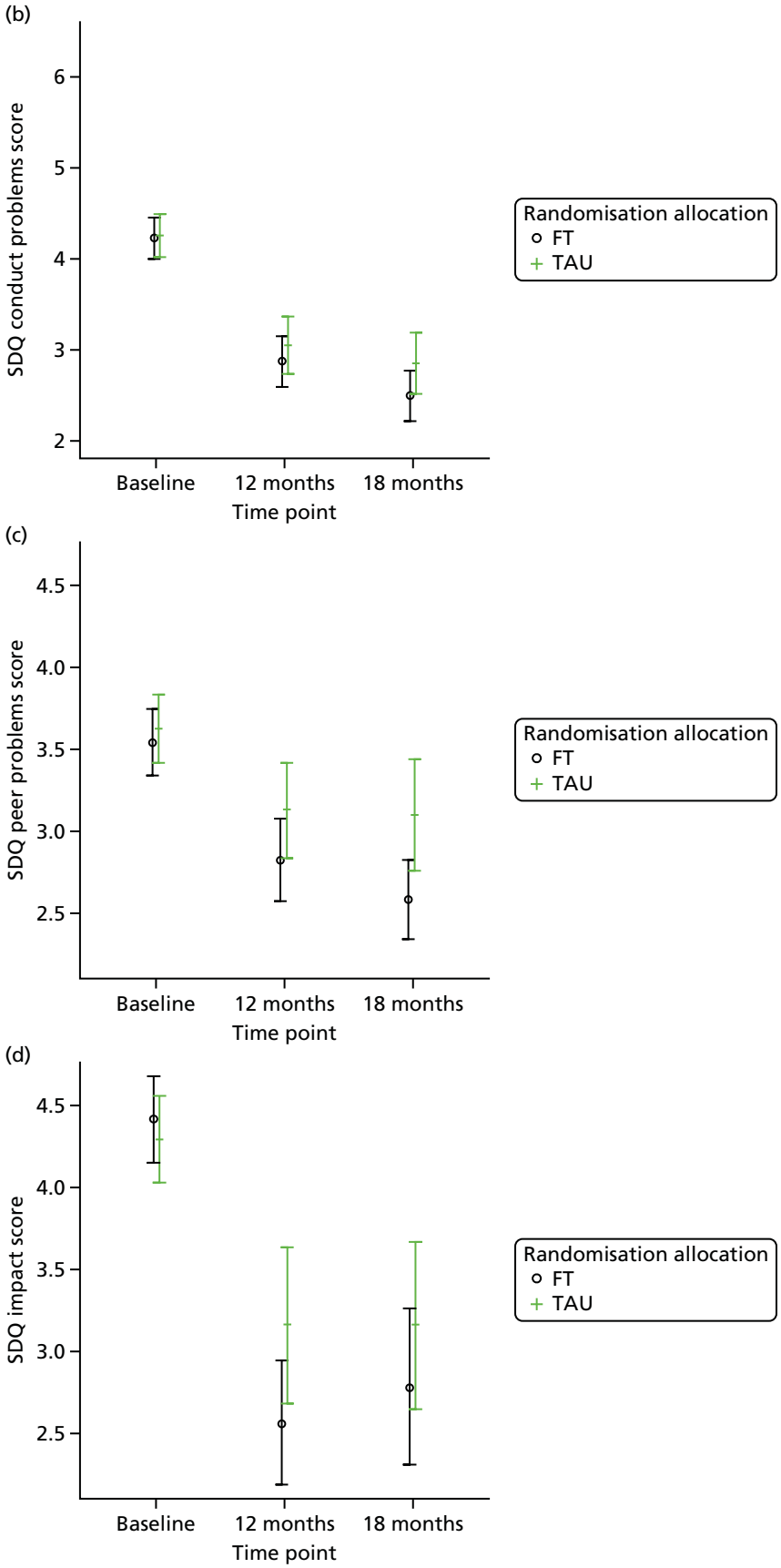


FIGURE 27 Unadjusted 95% CIs for caregiver SDQ (a) emotional problems; (b) conduct problems; (c) peer problems; and (d) impact problems.

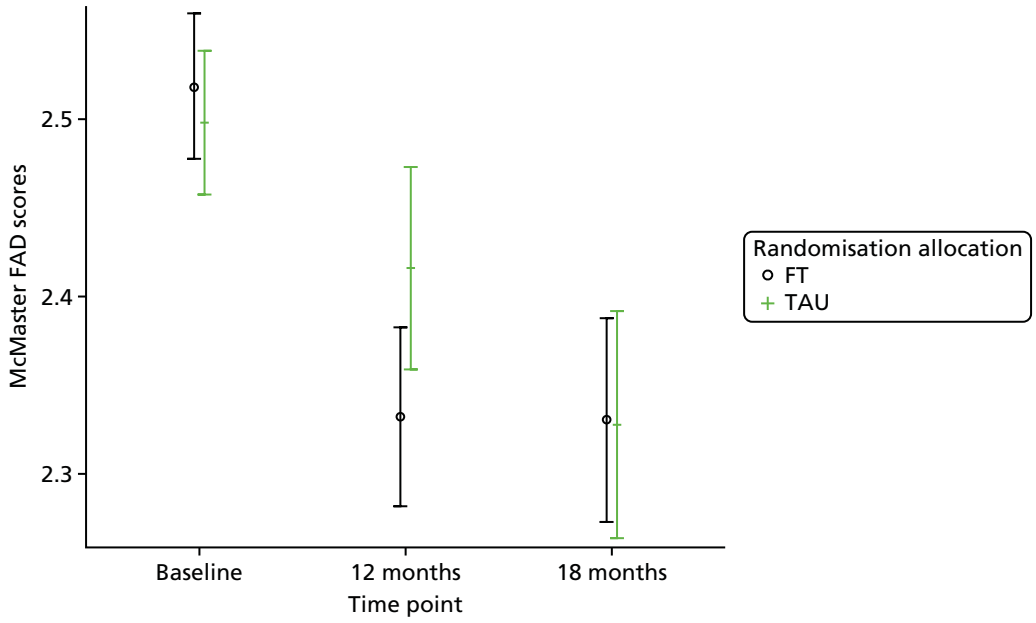


FIGURE 28 Unadjusted 95% CIs for caregiver McMaster FAD roles score.

Responder characteristics

Tables 109–112 and *Figure 66* in *Appendix 6* present the characteristics of participants with and without missing data (based on young person questionnaire return rates) in terms of responders and non-responders (followed up or lost to follow-up) by trial arm to explore if differential patterns exist that could impact on the conclusions for the secondary outcomes. Two logistic models were fitted to explore response status. The first included main effects for treatment arm, the participant characteristic and their interaction, to evaluate whether or not the response profile differed by the characteristic overall and further differentially by treatment arm. As tests for interaction do not provide a very powerful test for moderation, we also modelled response status by participant characteristics separately for each treatment arm, providing separate tests of the main characteristic effect in each arm to examine any differential effects.

Overall, participants lost to follow-up had significantly ($p < 0.1$) higher caregiver scores, representing particularly concerning participant baseline characteristics on the SDQ conduct problems, impact and externalising subscale; the FAD total score, roles and affective involvement subscales; the Family Questionnaire total score; the GHQ-12; and the ICU total score. Participants lost to follow-up at both 12 and 18 months were also more likely to have been referred into CAMHS from hospital. Those lost to follow-up at 12 months also had particularly concerning characteristics based on the young person FAD total score and at 18 months on the caregiver SDQ total difficulties score.

Examination of differential loss to follow-up by treatment arm found some indication ($p < 0.1$) that more males were lost to follow-up in the TAU arm than in the FT arm, whereas more females were lost to follow-up in the FT arm than in the TAU arm and that participants lost to follow-up at 18 months were slightly older in the TAU arm than in the FT arm (see *Figure 66a* and *b*, *Appendix 6*).

There was some indication that participants with particularly concerning initial presentations were lost to follow-up in the TAU arm and/or participants with particularly concerning initial presentations were followed up in the FT arm at 12 and/or 18 months based on participants' baseline method of index self-harm (self-poisoning and combined methods) and baseline scores for the young person and caregiver FAD total score; caregiver SDQ total, conduct, externalising and impact scores; caregiver FAD affective involvement score; caregiver ICU total score; and the caregiver Family Questionnaire score. Furthermore, the number of participants who self-harmed and who were followed up at 12 months was higher in the FT arm than in the TAU arm. These characteristics are described in *Figure 66 f–p*, *Appendix 6*.

In contrast, there was an indication that more participants with particularly concerning initial presentations were lost to follow-up in the FT arm and/or followed up in the TAU arm at 12 and/or 18 months, as indicated through hospital CAMHS referrals and based on baseline scores for the caregiver FAD roles scores and the caregiver GHQ-12 scores. These characteristics are described in *Figure 66 c–e*, *Appendix 6*.

In summary, participants lost to follow-up tended to have particularly concerning baseline characteristics; a greater proportion of participants were lost to follow-up in TAU than in FT and there was some indication of differential characteristics across treatment allocation with the majority of differences suggesting that participants with particulary concerning baseline characteristics lost to follow-up in TAU and/or participants with particularly concerning initial presentations followed up in FT. Based on these results, despite the high loss to follow-up for secondary participant-reported outcomes, examination of missing data patterns suggests that our estimates of treatment effect are conservative, as the potential effect of responses from those with missing data would likely increase the size of the treatment effect detected and therefore the conclusions for these outcomes remain valid.

Predictive and process measures

Moderator variables

Moderator analysis was used to explore whether or not the treatment effect in the Cox proportional hazards model from the primary analysis depended upon participants' baseline characteristics.

Despite the indication of moderation by referral source in summary statistics (see *Table 50*), a statistically significant interaction with treatment was not detected (see *Table 74*). The magnitude of moderation can be seen in *Figure 29*, which shows a HR very close to 1 (no treatment difference) and an increased hazard in the FT arm compared with the TAU arm for participants referred from the community. The results therefore provide some evidence, albeit not statistically significant, of a moderating effect, with an increased risk of self-harm among young people in the FT arm referred to CAMHS via hospital.

Significant interactions with treatment (at the 5% level), indicating moderation, were detected for the unemotional subscale on the young person-reported ICU ($p = 0.0104$) and the affective involvement subscale on the caregiver-reported McMaster FAD for both the overall score ($p = 0.0338$) and score categorised by healthy versus unhealthy families ($p = 0.0444$) (*Tables 74–76*).

Further potential interactions with treatment were detected, but only at the 10% level: the behaviour control subscale on the caregiver-reported McMaster FAD for both the score ($p = 0.571$) and healthy versus unhealthy families ($p = 0.0702$), the Beck screening for suicidal ideation ($p = 0.0881$) and also for

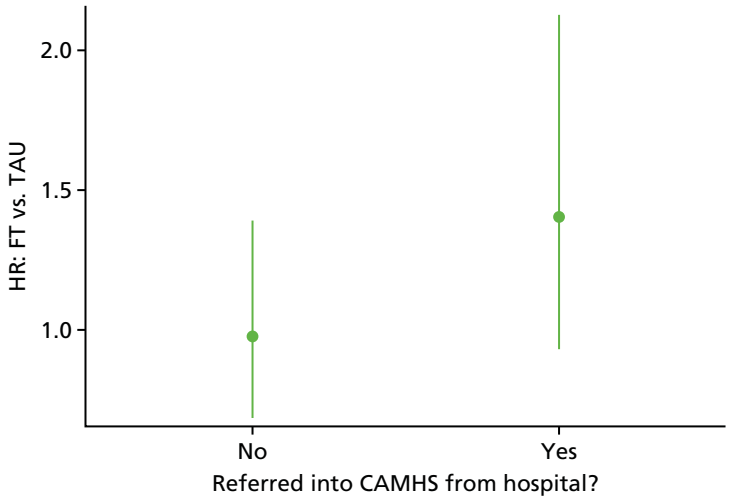


FIGURE 29 Hazard ratio for FT vs. TAU by referral source.

TABLE 74 Type 3 tests for moderation based on treatment × moderator interaction in the primary outcome model for baseline covariates

Baseline covariates: treatment × moderator			
Potential moderator	**df**	**Wald chi-squared test**	***p*-value**
Age	1	0.4730	0.4916
Centre	14	9.6983	0.7839
Referred from hospital[a]	1	1.7130	0.1906
Baseline number of self-harm episodes	1	0.1549	0.6939
Type of index self-harm episode	2	2.3900	0.3027
Sex	1	1.5219	0.2173

df, degrees of freedom.
a Source of referral was also investigated in the model excluding centre owing to known confounding effects. The interaction with treatment remained non-significant ($p = 0.1334$).

TABLE 75 Type 3 tests for moderation based on treatment × moderator interaction in the primary outcome model for categorised young person and caregiver baseline questionnaire responses

Potential moderator	**df**	**Wald chi-squared test**	***p*-value**
Young person questionnaires: treatment × moderator			
BSS	1	2.9095	0.0881
CDRS-R	1	0.0609	0.8051
McMaster FAD			
Affective involvement	1	1.0241	0.3116
Affective responsiveness	**1**	**3.1896**	**0.0741**
Behaviour control	1	0.0525	0.8188
Communication	1	0.0479	0.8268
General functioning	1	0.5484	0.4590
Problem-solving	1	0.0017	0.9671
Roles	1	0.5958	0.4402
Caregiver questionnaires: treatment × moderator			
McMaster FAD			
Affective involvement	1	4.0424	0.0444
Affective responsiveness	1	0.0508	0.8217
Behaviour control	1	3.2787	0.0702
Communication	1	1.1161	0.2908
General functioning	1	0.6174	0.4320
Problem-solving	1	0.0048	0.9445
Roles	1	0.9611	0.3269

df, degrees of freedom.
Note
Shading indicates significant findings.

TABLE 76 Type 3 tests for moderation based on treatment × moderator interaction in the primary outcome model for continuous young person and caregiver baseline questionnaire responses

Potential moderator	df	Wald chi-squared test	*p*-value
Young person questionnaires: treatment × moderator			
BSS	1	0.1620	0.6873
CDRS-R	1	0.1108	0.7392
PQ-LES-Q	1	0.0008	0.9774
Hopelessness	1	0.0000	0.9967
McMaster FAD			
Affective involvement	1	0.9190	0.3377
Affective responsiveness	1	1.0668	0.3017
Behaviour control	1	0.1908	0.6623
Communication	1	0.3647	0.5459
General functioning	1	0.7798	0.3772
Problem-solving	1	0.5967	0.4399
Roles	1	0.0509	0.8216
Total score	1	0.4413	0.5065
ICU			
Callousness	1	0.0876	0.7672
Total score	1	0.6874	0.4071
Uncaring	1	0.0988	0.7533
Unemotional	1	6.5713	0.0104
SDQ			
Conduct problems	1	0.0489	0.8250
Emotional problems	1	0.5393	0.4627
Externalising	1	0.0021	0.9638
Hyperactivity	1	0.2173	0.6411
Impact	1	1.1025	0.2937
Internalising	1	0.3037	0.5815
Peer problems	1	0.0008	0.9776
Prosocial	1	0.2672	0.6052
Total difficulties	1	0.2181	0.6405
Caregiver questionnaires: treatment × moderator			
Family Questionnaire			
Criticism	1	0.0098	0.9212
Emotional overinvolvement	1	0.4756	0.4904
GHQ-12 (Likert)	1	0.0342	0.8532

TABLE 76 Type 3 tests for moderation based on treatment × moderator interaction in the primary outcome model for continuous young person and caregiver baseline questionnaire responses (*continued*)

Potential moderator	df	Wald chi-squared test	*p*-value
McMaster FAD			
Affective involvement	1	4.5066	0.0338
Affective responsiveness	1	0.1042	0.7468
Behaviour control	1	3.6183	0.0571
Communication	1	0.9454	0.3309
General functioning	1	0.0019	0.9656
Problem-solving	1	0.0122	0.9120
Roles	1	1.3937	0.2378
Total score	1	0.9956	0.3184
ICU			
Callousness	1	0.7379	0.3903
Total score	1	0.5244	0.4690
Uncaring	1	0.1057	0.7451
Unemotional	1	0.1604	0.6888
SDQ			
Conduct problems	1	0.9628	0.3265
Emotional problems	1	0.9752	0.3234
Externalising	1	0.0561	0.8128
Hyperactivity	1	0.1353	0.7130
Impact	1	0.0486	0.8255
Internalising	1	0.5917	0.4418
Peer problems	1	0.0677	0.7947
Prosocial	1	0.4686	0.4936
Total difficulties	1	0.5405	0.4622

df, degrees of freedom.
Note
Shading indicates significant findings.

healthy versus unhealthy families according to the young person-reported affective responsiveness subscale of the McMaster FAD ($p = 0.0741$); however, this result was not supported by the uncategorised score ($p = 0.3017$).

No moderation was detected for other questionnaire responses or covariates.

Moderation varied in direction across the moderators (*Table 77* and *Figures 30–34*).

The risk of self-harm decreased in the TAU arm for young people with more unemotional traits (HR 0.932, 95% CI 0.875 to 0.992, for a 1-point increase), whereas the risk of self-harm increased in the FT arm for young people with more unemotional traits (HR 1.05, 95% CI 0.982 to 1.122, for a 1-point increase).

TABLE 77 Hazard ratios and overall type 3 *p*-values for moderation based on treatment × moderator interaction in the primary outcome model for baseline questionnaire responses

Moderator	HR (95% CI)	*p*-value (treatment × moderator)
Young person ICU unemotional score		***0.0104***
FT vs. TAU: at mean score = 9.01	1.127 (0.860 to 1.475)	
A 1-point score increase: in FT	1.050 (0.982 to 1.122)	
A 1-point score increase: in TAU	0.932 (0.875 to 0.992)	
Caregiver FAD affective involvement score		***0.0338***
FT vs. TAU: at mean score = 2.24	1.146 (0.875 to 1.501)	
A 1-point score increase: in FT	0.885 (0.585 to 1.339)	
A 1-point score increase: in TAU	1.627 (1.107 to 2.390)	
Caregiver FAD affective involvement category		***0.0444***
FT vs. TAU: in healthy	1.695 (1.047 to 2.743)	
FT vs. TAU: in unhealthy	0.930 (0.671 to 1.289)	
Healthy vs. unhealthy: in FT	1.165 (0.799 to 1.699)	
Healthy vs. unhealthy: in TAU	0.639 (0.410 to 0.997)	
Caregiver FAD behaviour control score		***0.0571***
FT vs. TAU: at mean score = 1.84	1.129 (0.862 to 1.477)	
A 1-point score increase: in FT	0.772 (0.487 to 1.224)	
A 1-point score increase: in TAU	1.466 (0.910 to 2.363)	
Caregiver FAD behaviour control category		***0.0702***
FT vs. TAU: in healthy	1.470 (0.990 to 2.180)	
FT vs. TAU: in unhealthy	0.887 (0.610 to 1.289)	
Healthy vs. unhealthy: in FT	1.152 (0.794 to 1.672)	
Healthy vs. unhealthy: in TAU	0.695 (0.465 to 1.039)	
Young person FAD affective responsiveness category		***0.0741***
FT vs. TAU: in healthy	0.584 (0.276 to 1.234)	
FT vs. TAU: in unhealthy	1.215 (0.909 to 1.624)	
Healthy vs. unhealthy: in FT	0.496 (0.264 to 0.930)	
Healthy vs. unhealthy: in TAU	1.033 (0.621 to 1.716)	
Young person BSS category		***0.0881***
FT vs. TAU: no SI	0.772 (0.464 to 1.284)	
FT vs. TAU: yes SI	1.303 (0.948 to 1.791)	
SI vs. no SI: in FT	1.814 (1.169 to 2.815)	
SI vs. no SI: in TAU	0.931 (0.611 to 1.417)	

SI, suicidal ideation.

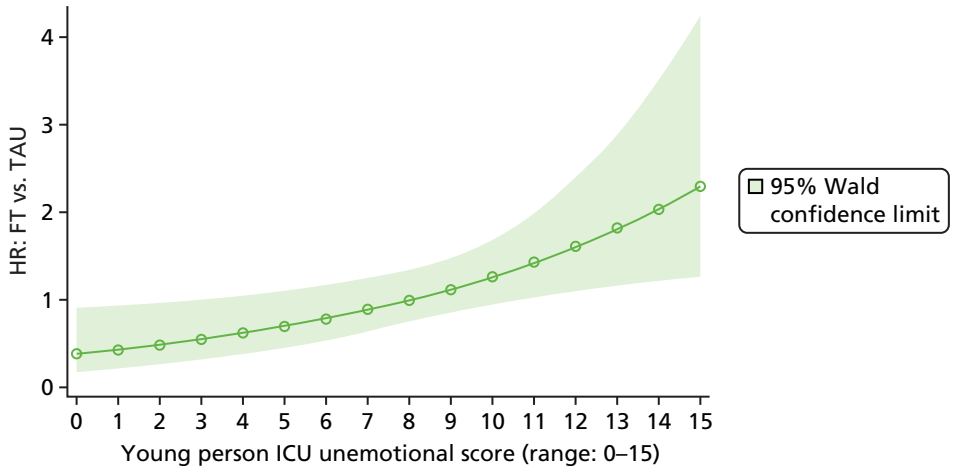


FIGURE 30 Hazard ratio for FT vs. TAU by the baseline young person ICU unemotional score. Higher scores represent higher unemotional traits. Adapted from Cottrell *et al.*[70] © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

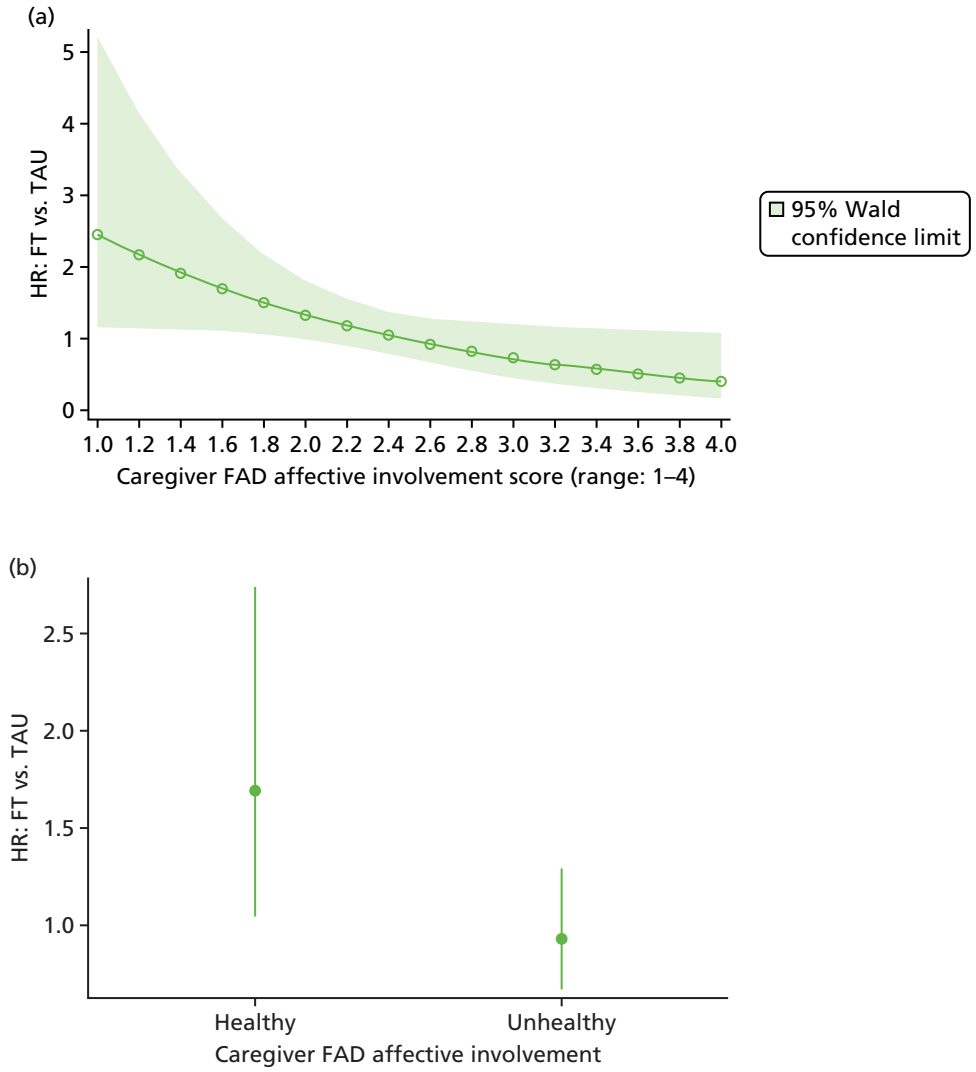


FIGURE 31 Hazard ratio for FT vs. TAU by the baseline caregiver McMaster FAD affective involvement score and category. Higher scores are indicative of poorer family functioning. (a) Continuous score; and (b) families classified as healthy vs. unhealthy. Adapted from Cottrell *et al.*[70] © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

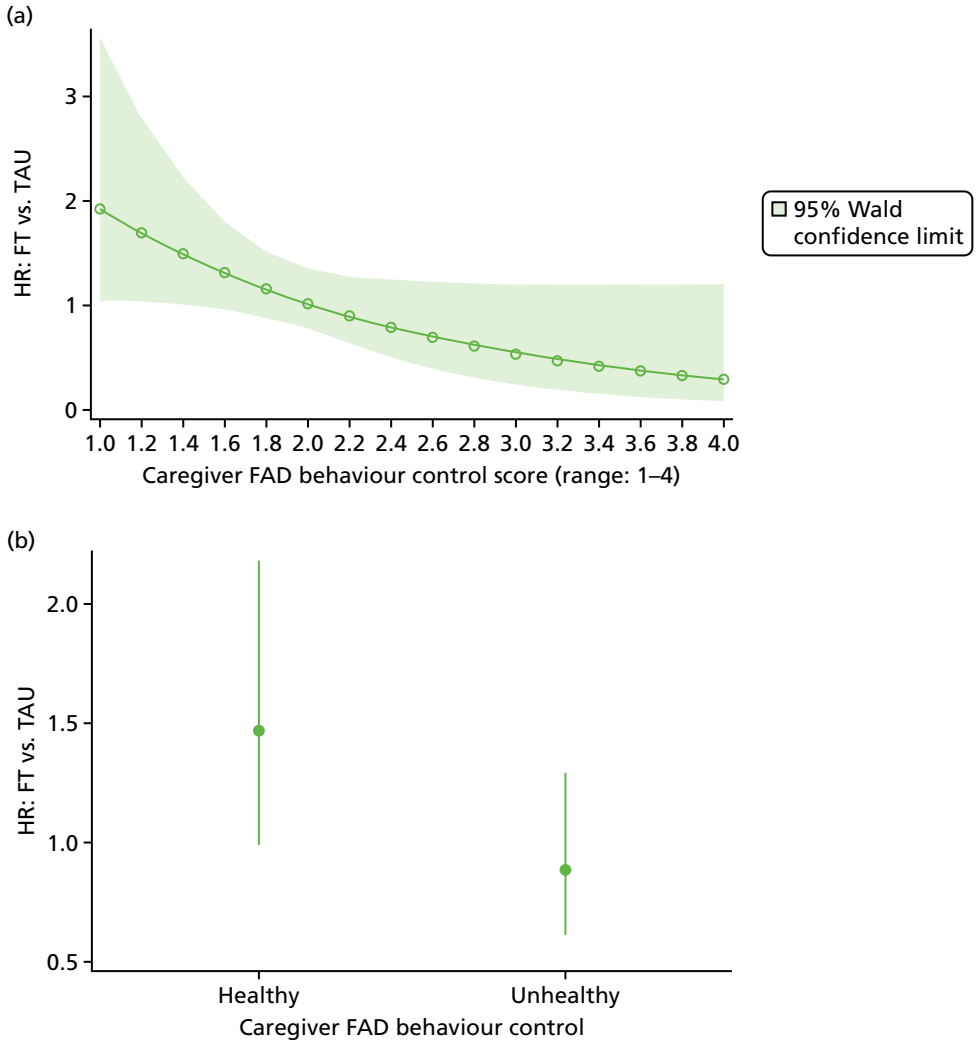


FIGURE 32 Hazard ratio for FT vs. TAU by the baseline caregiver McMaster FAD behaviour control score and category. Higher scores are indicative of poorer family functioning. (a) Continuous score; and (b) families classified as healthy vs. unhealthy.

There was only limited evidence ($p = 0.0881$) of a higher risk of self-harm in the FT arm for young people with suicidal ideation than for those without, as identified through screening on the BSS (HR 1.814, 95% CI 1.169 to 2.815). Young people's responses on the affective responsiveness subscale also provided limited evidence ($p = 0.0741$) that unhealthy families (as defined by the FAD) had an increased risk of self-harm in the FT arm compared with the TAU arm (HR 1.215, 95% CI 0.909 to 1.624); however, moderation was not supported by the continuous score.

On the caregiver-reported affective involvement and behaviour control subscales, moderation differed in direction compared to moderation by young person-reported unemotional traits, suicide ideation and affective responsiveness. The risk of self-harm was higher in the TAU arm for participants with higher caregiver scores indicating poorer family functioning (HR 1.627, 95% CI 1.107 to 2.390, for a 1-point increase in affective involvement; and HR 1.466, 95% CI 0.910 to 2.363, for behaviour control), whereas the risk of self-harm was lower in the FT arm for participants with poorer family functioning (HR 0.885, 95% CI 0.585 to 1.339, for a 1-point increase in affective involvement; and HR 0.772, 95% CI 0.487 to 1.224, for behaviour control). This moderation was supported by the categorisation of families as healthy versus unhealthy based on their scores, with healthy families in the FT arm having on both scales a higher risk of self-harm than those in the TAU arm (HR 1.695, 95% CI 1.047 to 2.743, for affective involvement; and HR 1.470, 95% CI 0.990 to 2.180, for behaviour control).

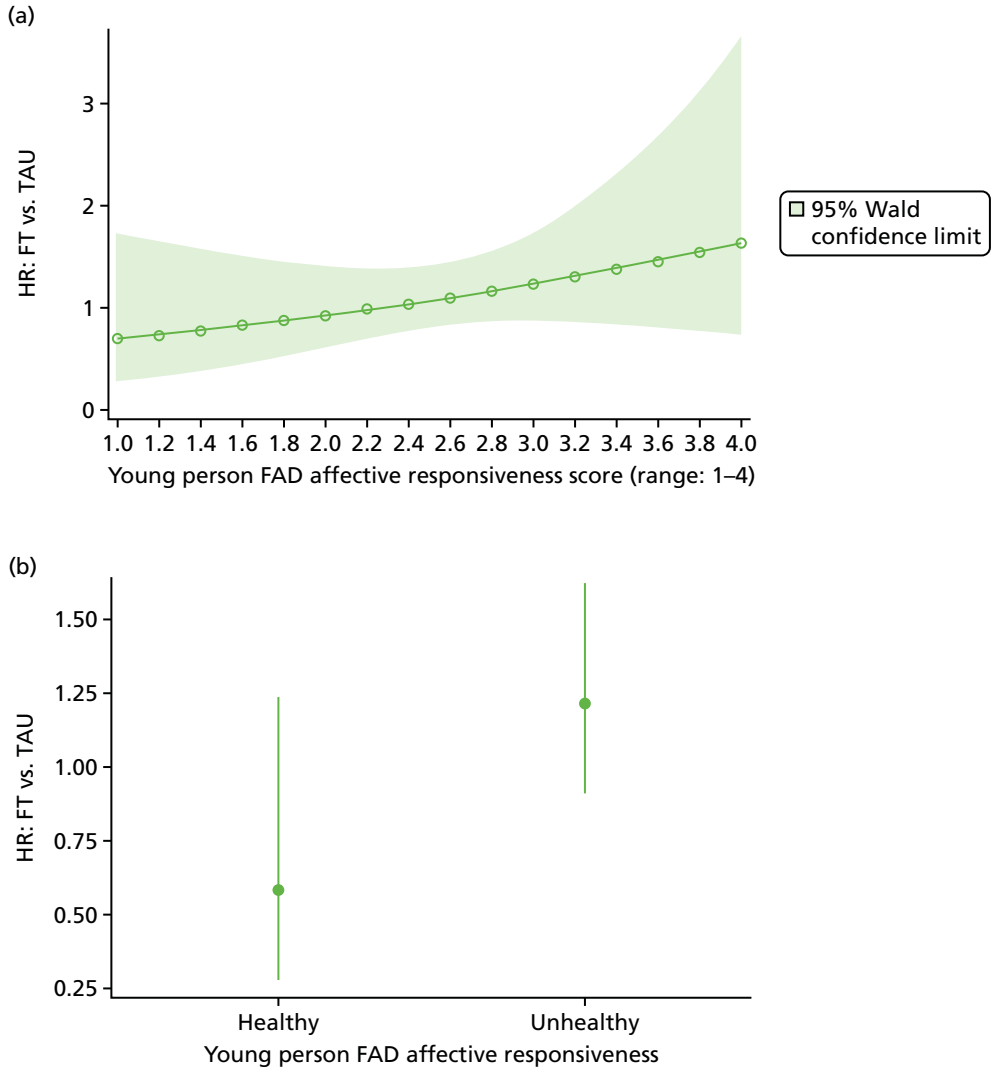


FIGURE 33 Hazard ratio for FT vs. TAU by the baseline young person McMaster FAD affective responsiveness score and category. Higher scores are indicative of poorer family functioning.

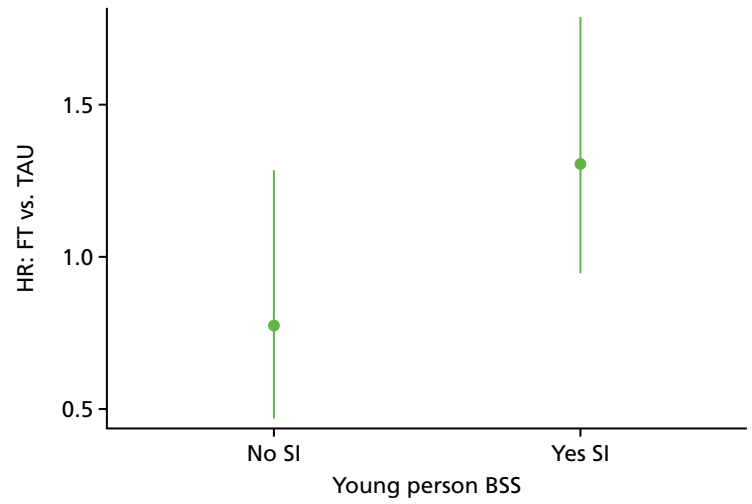


FIGURE 34 Hazard ratio for FT vs. TAU by baseline suicide ideation screening on the young person BSS. SI, suicidal ideation. (a) Continuous score; and (b) families classified as healthy vs. unhealthy.

Mediator variables

Complier average causal effect analysis

A CACE analysis was conducted to model the causal effect of FT receipt (as opposed to randomisation) on the primary outcome.

Summary statistics showing the proportion of participants with an event in each arm and by receipt of FT or formal systemic FT as part of usual care are shown in *Table 78* and *Figures 35* and *36*. The highest overall self-harm rate was exhibited by TAU participants with missing treatment data, with a primary outcome event reported in 15 (33.3%) participants in this group. The lowest repetition rate was reported in three (14.3%) participants allocated to receive FT who attended no FT sessions; however, they also constituted the smallest group. The repetition rate in participants allocated to and who attended SHIFT FT was 29.2%, while for those in the TAU arm not attending FT the repetition rate was lower, at 24.6%. Furthermore, in participants allocated to TAU who attended FT, the repetition rate was lower again, at 20.7%. Considering attendance at FT sessions irrespective of randomisation, the repetition rate was 27.7% in those who received at least one FT session and 23.9% in those who did not.

The CACE analysis shows a very similar effect of the intervention among those who received FT [0.12 (SE 0.13); $p = 0.3424$] compared with the standard 'ITT' estimate of the 'allocation' of FT [0.11 (SE 0.10); $p = 0.2395$] and the 'as treated' estimate [0.10 (SE 0.10); $p = 0.3136$], with no significant differences detected between trial arms, or receipt of FT (*Table 79*). Furthermore, the correlation parameter estimate from the CACE model indicates that treatment selection bias on the outcome is not a problem [0.03 (SE 0.11); $p = 0.7869$].

Finally, *Figure 37* further displays the time to first event for participants allocated to FT who received FT and those allocated to TAU who received a form of FT as part of their usual care. The curves appear to diverge, with an increased rate of self-harm in those allocated to FT; however, 95% CIs around the two curves overlap at each time point and the log-rank estimate of a difference between trial arms finds no statistically significant difference ($p = 0.1087$).

Mediation

Summary statistics showing the proportion of participants with a primary outcome event in each arm by variables considered to be potential process mediators are shown in *Table 80*. Young people on a psychotropic medication during follow-up were more likely to self-harm than those who were not prescribed such drugs: 41 out of 104 (39.4%) young people on a medication engaged in self-harm, compared with 164 out of 670 (24.5%) not on medication. In the FT arm, rates of self-harm were higher in young people whose lead therapist had been working in CAMHS for ≥ 4 years than in those whose lead therapist had been working for < 4, years with repetition observed for 77 (27.8%) and nine (20.9%) participants, respectively. Conversely, in the TAU arm, rates of self-harm were lower among participants seen by more experienced therapists than among those seen by less experienced therapists, with repetition observed for 24 (19.7%) and 12 (26.7%) participants, respectively. In both arms, participants who self-harmed attended more sessions overall [mean 11.2 (SD 12.68) sessions] than those who did not [mean 7.5 (10.03) sessions].

For potential mediators based on 3-, 6- and 12-month questionnaire responses, mediation was explored in relation to the time to event post 3, 6 and 12 months, respectively. *Table 81* presents the proportion of participants with a self-harm event leading to hospital attendance and *Figures 38–40* present Kaplan–Meier plots of time to self-harm for each period.

Following the Baron and Kenny[90] steps, there was no evidence that any of the variables investigated formally mediated the effect of treatment on the time to self-harm, largely because of lack of evidence of a treatment effect (*Table 82*).

TABLE 78 Primary outcome self-harm event observed according to whether or not any formal systemic FT sessions were attended as part of randomisation or TAU

SHIFT FT or formal systemic FT sessions attended?	FT			TAU				Total			
	Yes (*N* = 394)	No (*N* = 21)	Total (*N* = 415)	Yes (*N* = 87)	No (*N* = 285)	Missing (*N* = 45)	Total (*N* = 417)	Yes (*N* = 481)	No (*N* = 306)	Missing (*N* = 45)	Total (*N* = 832)
Primary outcome event?											
Yes, *n* (%)	115 (29.2)	3 (14.3)	118 (28.4)	18 (20.7)	70 (24.6)	15 (33.3)	103 (24.7)	133 (27.7)	73 (23.9)	15 (33.3)	221 (26.6)
No, *n* (%)	279 (70.8)	18 (85.7)	297 (71.6)	69 (79.3)	215 (75.4)	30 (66.7)	314 (75.3)	348 (72.3)	233 (76.1)	30 (66.7)	611 (73.4)

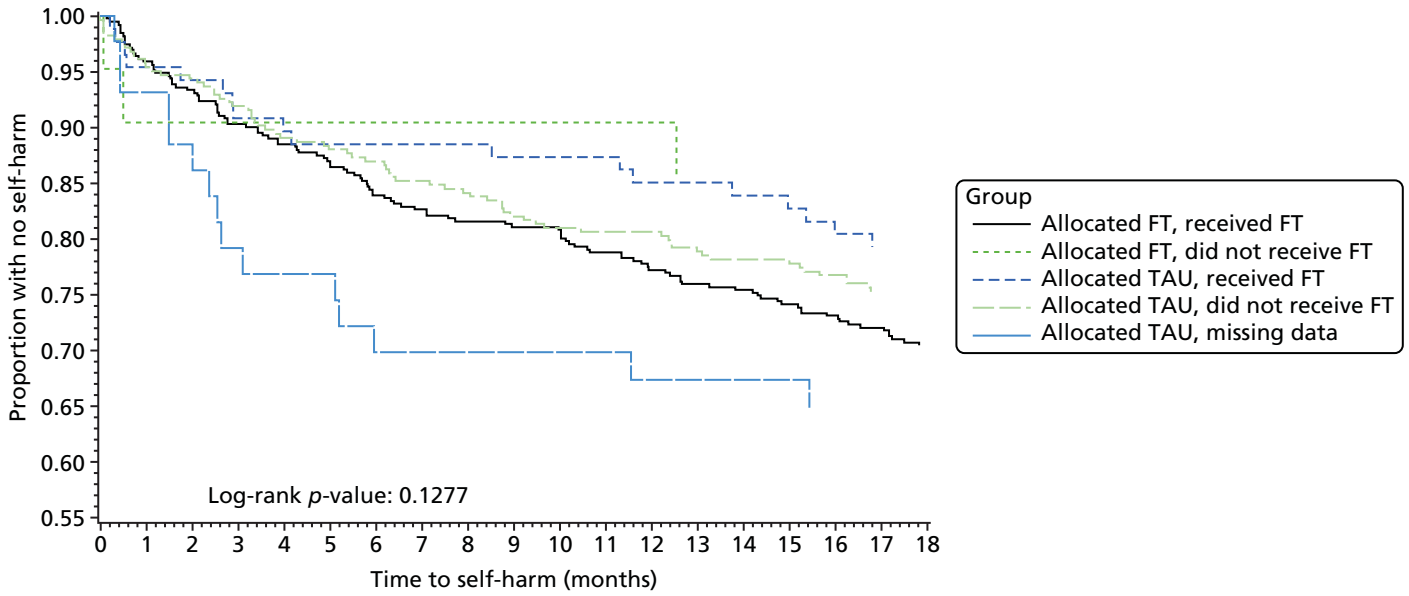


FIGURE 35 Kaplan–Meier plot of time to self-harm by randomised treatment group and receipt of FT. Reproduced from Cottrell *et al.*[70] © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

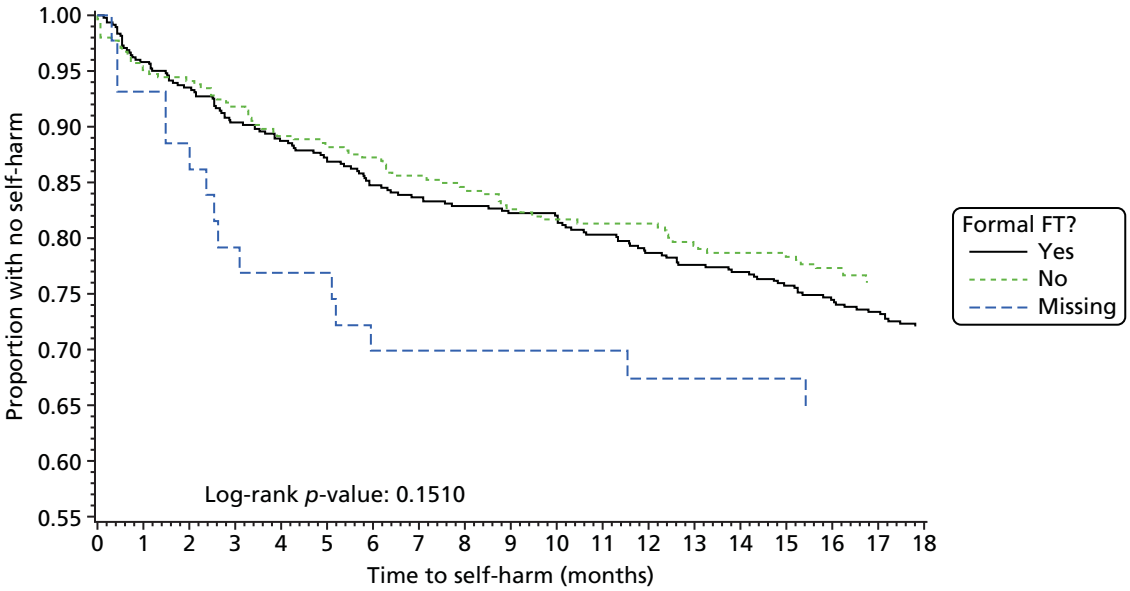


FIGURE 36 Kaplan–Meier plot of time to self-harm by receipt of FT: time to first event for young person who received FT vs. those who did not, regardless of randomisation.

TABLE 79 Adjusted multivariable probit regression estimates for self-harm during follow-up for the ITT, as treated and CACE analysis

	ITT estimate: randomised treatment			As-treated estimate: FT receipt			CACE estimate:[a] stage 2 – FT receipt with randomisation instrumental variable		
Model parameter	Parameter estimate (SE)	df	*p*-value	Parameter estimate (SE)	df	*p*-value	Parameter estimate (SE)	df	*p*-value
Intercept	–0.89 (0.24)	1	0.0002	–0.89 (0.24)	1	0.0003	–0.90 (0.25)	1	0.0003
Randomised treatment: FT (vs. TAU)	0.11 (0.10)	1	0.2395		–	–		–	–
FT received: FT (vs. no FT)		–	–	0.10 (0.10)	1	0.3136	0.12 (0.13)	1	0.3424
Age group (years): 15–17 (vs. 11–14)	–0.24 (0.10)	1	0.0164	–0.23 (0.10)	1	0.0188	–0.23 (0.10)	1	0.0194
Gender: female (vs. male)	0.29 (0.16)	1	0.0668	0.30 (0.16)	1	0.0645	0.30 (0.16)	1	0.0635
Number of previous self-harm episodes: ≥ 3 (vs. 2)	0.12 (0.16)	1	0.4654	0.11 (0.16)	1	0.4817	0.11 (0.16)	1	0.4837
Type of index episode (vs. self-injury)		–	–		–	–		–	–
Combined	0.49 (0.19)	1	0.0112	0.48 (0.19)	1	0.0129	0.48 (0.19)	1	0.0131
Self-poisoning	0.04 (0.14)	1	0.7739	0.04 (0.14)	1	0.7885	0.04 (0.14)	1	0.7882
Referred via hospital: yes (vs. no)	0.16 (0.12)	1	0.1880	0.17 (0.12)	1	0.1751	0.17 (0.12)	1	0.1760
Trust (vs. Y trust 1)		–	–		–	–		–	–
M trust 3	0.09 (0.17)	1	0.6120	0.09 (0.17)	1	0.6074	0.09 (0.17)	1	0.6064
L trust 5	–0.42 (0.28)	1	0.1440	–0.42 (0.28)	1	0.1424	–0.42 (0.28)	1	0.1417
Y trust 3	–0.43 (0.24)	1	0.0718	–0.43 (0.24)	1	0.0747	–0.43 (0.24)	1	0.0750
Y trust 2	–0.23 (0.27)	1	0.3864	–0.22 (0.27)	1	0.4105	–0.22 (0.27)	1	0.4149

TABLE 79 Adjusted multivariable probit regression estimates for self-harm during follow-up for the ITT, as treated and CACE analysis (*continued*)

	ITT estimate: randomised treatment			As-treated estimate: FT receipt			CACE estimate:[a] stage 2 – FT receipt with randomisation instrumental variable		
Model parameter	Parameter estimate (SE)	df	*p*-value	Parameter estimate (SE)	df	*p*-value	Parameter estimate (SE)	df	*p*-value
M trust 1	0.10 (0.22)	1	0.6557	0.10 (0.22)	1	0.6549	0.10 (0.22)	1	0.6562
L trust 6	–0.35 (0.49)	1	0.4671	–0.41 (0.49)	1	0.4054	–0.41 (0.49)	1	0.3998
Y trust 4	–0.75 (0.31)	1	0.0139	–0.76 (0.31)	1	0.0134	–0.75 (0.31)	1	0.0135
L trust 2	–0.50 (0.21)	1	0.0191	–0.52 (0.21)	1	0.0160	–0.52 (0.21)	1	0.0154
M trust 2	–0.17 (0.18)	1	0.3499	–0.17 (0.18)	1	0.3484	–0.17 (0.18)	1	0.3482
L trust 1	–0.28 (0.18)	1	0.1257	–0.28 (0.18)	1	0.1210	–0.29 (0.18)	1	0.1196
Y trust 6	0.17 (0.25)	1	0.4979	0.18 (0.25)	1	0.4834	0.18 (0.26)	1	0.4776
L trust 4	–0.41 (0.61)	1	0.5085	–0.40 (0.61)	1	0.5133	–0.40 (0.61)	1	0.5143
L trust 3	–0.29 (0.37)	1	0.4385	–0.29 (0.37)	1	0.4328	–0.29 (0.37)	1	0.4321
Y trust 5	–0.20 (0.29)	1	0.4797	–0.22 (0.29)	1	0.4406	–0.23 (0.29)	1	0.4305
Rho (selection bias)		–	–		–	–	0.03 (0.11)	1	0.7869

df, degrees of freedom; SLaM, South London and Maudsley.

a Stage 1 of the CACE analysis found strong evidence that randomisation was associated with FT receipt [estimate:–2.63 (SE 0.14), $p < 0.0001$], satisfying the assumption of a non-zero causal effect of the instrumental variable on treatment.

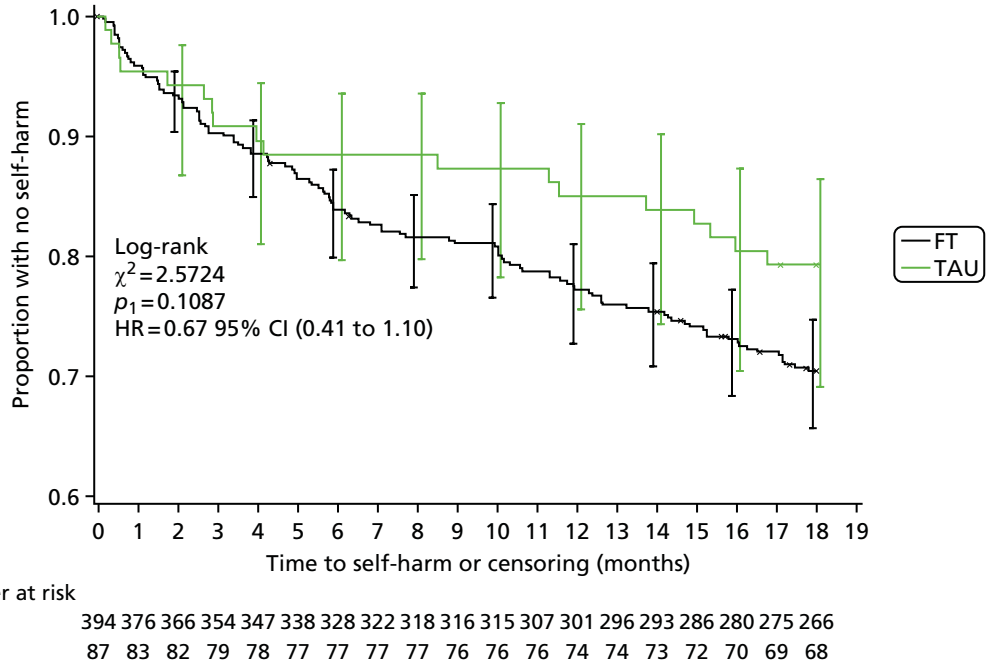


FIGURE 37 Kaplan–Meier plot of time to self-harm by arm for those receiving FT: time to first event for young person who received FT in either arm by allocation.

TABLE 80 Summary of self-harm event by potential process mediator and treatment arm

	Primary outcome event reported					
	FT		TAU		Total	
Potential mediator	Yes (self-harm)	No (no self-harm)	Yes (self-harm)	No (no self-harm)	Yes (self-harm)	No (no self-harm)
Young person on any psychotropic medications, *n* (%)						
Yes (*n* = 104)	18 (40.9)	26 (59.1)	23 (38.3)	37 (61.7)	41 (39.4)	63 (60.6)
No (*n* = 670)	100 (27.2)	267 (72.8)	64 (21.1)	239 (78.9)	164 (24.5)	506 (75.5)
Missing (*n* = 58)	0 (0.0)	4 (100.0)	16 (29.6)	38 (70.4)	16 (27.6)	42 (72.4)
Total (*n* = 832)	118 (28.4)	297 (71.6)	103 (24.7)	314 (75.3)	221 (26.6)	611 (73.4)
Years spent working in CAMHS, *n* (%)						
< 4 years (*n* = 88)	9 (20.9)	34 (79.1)	12 (26.7)	33 (73.3)	21 (23.9)	67 (76.1)
≥ 4 years (*n* = 399)	77 (27.8)	200 (72.2)	24 (19.7)	98 (80.3)	101 (25.3)	298 (74.7)
Missing (*n* = 345)	32 (33.7)	63 (66.3)	67 (26.8)	183 (73.2)	99 (28.7)	246 (71.3)
Total (*n* = 832)	118 (28.4)	297 (71.6)	103 (24.7)	314 (75.3)	221 (26.6)	611 (73.4)
Overall number of sessions attended by anyone						
n	118	297	88	284	206	581
Mean (SD)	9.9 (10.32)	6.8 (5.32)	12.9 (15.18)	8.2 (13.25)	11.2 (12.68)	7.5 (10.03)
Median (range)	7.0 (0–70)	6.0 (0–47)	8.0 (0–90)	4.0 (0–163)	7.0 (0–90)	5.0 (0–163)

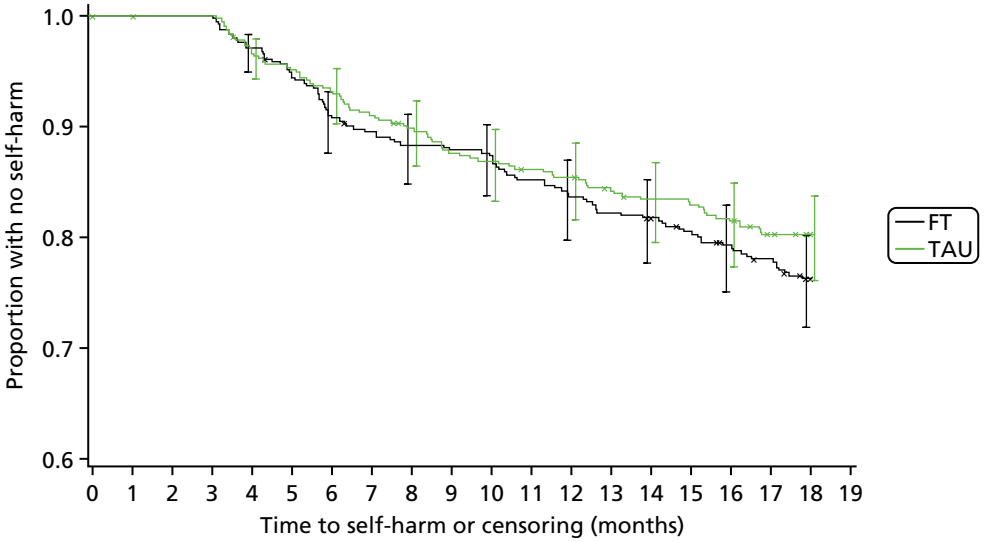
TABLE 81 Summary of self-harm event from 3, 6 and 12 months post randomisation

Self-harm event	FT (*N* = 415), *n* (%)	TAU (*N* = 417), *n* (%)	Total (*N* = 832), *n* (%)
≥ 3 months post randomisation	97 (23.4)	81 (19.4)	178 (21.4)
≥ 6 months post randomisation	74 (17.8)	65 (15.6)	139 (16.7)
≥ 12 months post randomisation	48 (11.6)	33 (7.9)	81 (9.7)

Considering the effect of randomised treatment on the mediator in step 2 and the effect of the mediator of self-harm outcomes in step 3, we further discuss potential mediation for variables where significant associations were detected in both steps (with at least good evidence < 5%).

Process mediators

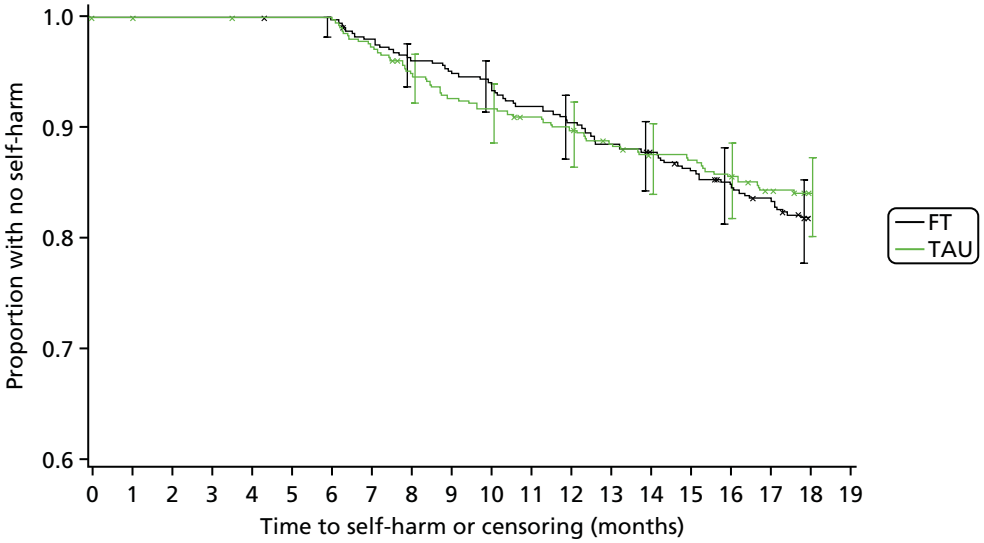
There was good evidence (step 2) of an association between randomised treatment and psychotropic medication use during follow-up (less in FT: OR 0.6; $p = 0.0158$) and strong evidence (step 3) that the use of psychotropic medication is associated with an increased risk of self-harm (HR 2.10; $p < 0.0001$). There was, however, no evidence of a treatment effect on the risk of self-harm (step 1) that could be mediated (step 3) and, furthermore, there is no evidence of a direct treatment effect despite the positive association with FT and the mediator (step 3). It should be noted that these data refer to all young people prescribed any form of psychotropic medication during the treatment period (not specific to selective serotonin reuptake inhibitors; see *Table 37*) and that, furthermore, we do not have reliable data on the timing of medication and therefore its relationship to the timing of primary outcome event.



Number at risk

FT	415 413 413 412 401 389 374 368 363 361 359 350 344 338 334 326 319 313 302
TAU	417 415 414 414 399 393 385 376 369 360 357 353 350 343 339 337 331 323 319

FIGURE 38 Kaplan–Meier plot of time to self-harm from 3 months post randomisation by randomised treatment group with 95% CIs.



Number at risk

FT	415 413 413 413 413 412 411 403 395 391 387 378 372 364 359 350 343 336 325
TAU	417 415 414 414 413 413 413 402 390 381 377 372 367 360 354 352 346 338 333

FIGURE 39 Kaplan–Meier plot of time to self-harm from 6 months post randomisation by randomised treatment group with 95% CIs.

There was strong evidence (step 3) that the number of therapy sessions was associated with the risk of self-harm during follow-up (HR 1.02; $p < 0.0001$), with risk increasing with more sessions. There was only weak evidence (step 2), however, that treatment group was associated with number of sessions (fewer sessions in FT: –1.44; $p = 0.0640$) and no evidence of a treatment effect on the risk of self-harm (step 1) that could be mediated (step 3). When the number of sessions attended is accounted for, there is weak evidence of a direct treatment effect (step 3) with an increased risk of self-harm in FT (HR 1.29; $p = 0.0825$).

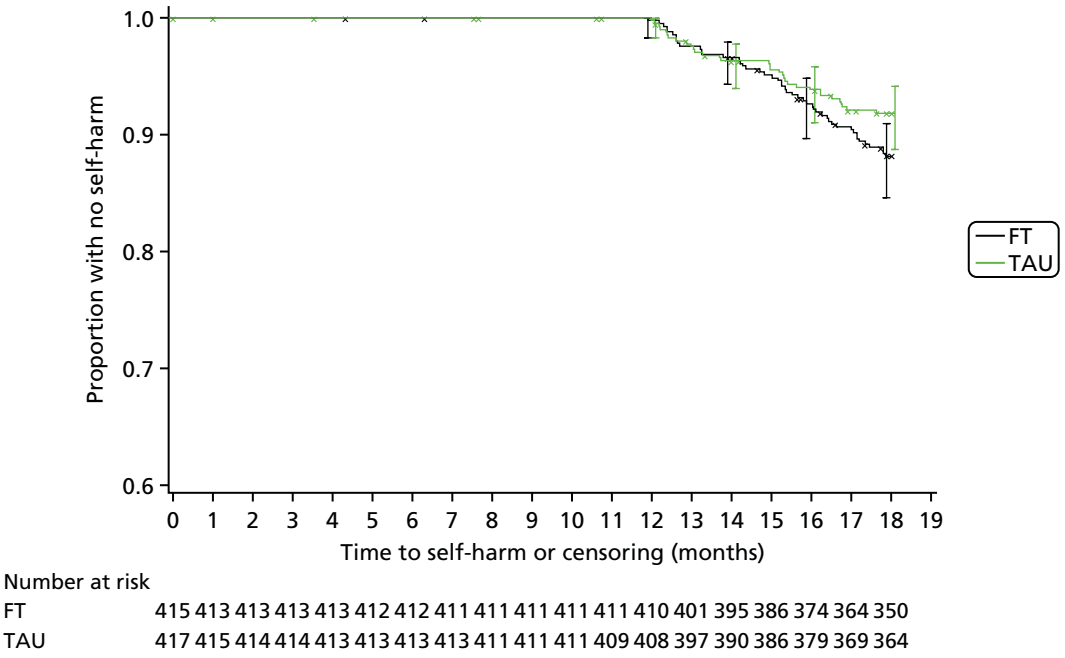


FIGURE 40 Kaplan–Meier plot of time to self-harm from 12 months post randomisation by randomised treatment group with 95% CIs.

There was strong evidence of an association between treatment group and the length of experience of the lead therapist (more experienced in FT: OR 2.38; $p = 0.0003$), but no evidence that length of experience was associated with the risk of self-harm (step 3). Summary statistics, however, do indicate differences between self-harm rates in relation to therapist experience across the two arms that are not accounted for via this mediation model and that warrant further investigation.

For both the number of sessions attended and the use of psychotropic medication, although there is no evidence of a treatment effect in steps 1 or 3, the magnitude of the treatment coefficient is increased and the suggested relationship between treatment and mediator, and between mediator and self-harm outcome, warrants further investigation.

Three-month questionnaires

There was strong evidence of a treatment effect on the risk of self-harm (steps 1 and 3) when investigating responses to the Family Questionnaire as potential mediators. However, this is for a subsample (53%) of responding trial participants and, as there is no evidence that treatment is associated with either mediator (step 2) and there is no change in the treatment effect when allowing for the potential mediating effect (step 3), no mediation effects are suggested.

Twelve-month questionnaires

Significant treatment effects on the mediator (step 2) and significant mediator effects on the risk of self-harm (step 3) were observed for the young person SDQ impact, caregiver SDQ emotional problems and caregiver McMaster FAD roles subscales. Participants in the FT arm had improved outcomes on each mediator at 12 months (step 2), and the risk of self-harm decreased as mediator scores improved. The positive treatment effect on the mediator outcome did not, however, translate through to self-harm outcomes in either step 1 or 2; rather, the estimated treatment effect actually increased in favour of TAU, although not significantly so. Therefore, further investigation is warranted to explore why improved mediator outcomes did not translate to self-harm outcomes.

TABLE 82 Mediators analysis: assessment of the Baron and Kenny mediation steps[a]

Mediator variable	*n*	Step 1 [$Y = h_0(t)\exp(\beta_{10} X)$] Total effect of Rand on outcome – c (β_{10})			Step 2 ($Me = \beta_{20} + \beta_{21} X$) Effect of Rand on mediator – a (β_{21})			Step 3 [$Y = h_0(t)\exp(\beta_{31} X + \beta_{32} Me)$] Effect of mediator on outcome – b (β_{32})			Direct effect of Rand on outcome – c′ (β_{31})			Suggestion of potential for mediation
		Coefficient (SE)	HR (95% CI)	*p*-value	Coefficient (SE)	(95% CI)	*p*-value	Coefficient (SE)	HR (95% CI)	*p*-value	Coefficient (SE)	HR (95% CI)	*p*-value	
Process mediators														
Overall young person on any psychotropic medications[b]	774	0.17 (0.14)	1.19 (0.90 to 1.58)	0.2209	–0.26 (0.11)	OR 0.60 (0.39 to 0.91)	0.0158	0.74 (0.18)	2.10 (1.47 to 3.00)	< 0.0001	0.23 (0.14)	1.26 (0.95 to 1.67)	0.1096	X
Years spent working in CAMHS by lead therapist[b]	487	0.26 (0.20)	1.30 (0.88 to 1.93)	0.1933	0.43 (0.12)	OR 2.38 (1.49 to 3.82)	0.0003	–0.24 (0.27)	0.78 (0.46 to 1.33)	0.3650	0.30 (0.21)	1.35 (0.90 to 2.02)	0.1505	X
Number of sessions attended	787	0.18 (0.14)	1.20 (0.91 to 1.59)	0.1970	–1.44 (0.77)	(–2.95 to 0.08)	0.0640	0.02 (0.00)	1.02 (1.01 to 1.03)	< 0.0001	0.25 (0.14)	1.29 (0.97 to 1.71)	0.0825	X
3-month caregiver questionnaires														
Family Q criticism	439	0.73 (0.24)	2.08 (1.31 to 3.33)	0.0021	0.33 (0.46)	(–0.56 to 1.23)	0.4624	0.03 (0.02)	1.03 (0.98 to 1.08)	0.2250	0.73 (0.24)	2.07 (1.29 to 3.31)	0.0024	
Family Q emotional involvement	439	0.73 (0.24)	2.08 (1.30 to 3.32)	0.0022	0.51 (0.40)	(–0.28 to 1.29)	0.2061	0.11 (0.03)	1.11 (1.05, 1.18)	0.0005	0.74 (0.24)	2.09 (1.30 to 3.34)	0.0022	
6-month caregiver questionnaires														
Family Q criticism	360	0.51 (0.29)	1.66 (0.94 to 2.93)	0.0802	–0.86 (0.57)	(–1.98 to 0.26)	0.1302	0.05 (0.03)	1.05 (1.00, 1.10)	0.0564	0.56 (0.29)	1.74 (0.98 to 3.09)	0.0567	
Family Q emotional involvement	360	0.50 (0.29)	1.66 (0.94 to 2.92)	0.0815	–0.64 (0.53)	(–1.68 to 0.40)	0.2264	0.11 (0.03)	1.11 (1.05, 1.18)	0.0005	0.57 (0.29)	1.76 (1.00 to 3.11)	0.0517	
12-month young person questionnaires														
Beck (continuous score)	449	0.23 (0.32)	1.26 (0.67 to 2.36)	0.4801	–1.00 (0.69)	(–2.36 to 0.35)	0.1467	0.08 (0.02)	1.08 (1.04, 1.12)	< 0.0001	0.36 (0.33)	1.43 (0.75 to 2.73)	0.2806	
Beck (SI screen)[b]	449	0.23 (0.32)	1.26 (0.67 to 2.36)	0.4801	–0.17 (0.10)	OR 0.71 (0.47 to 1.06)	0.0968	1.04 (0.35)	2.83 (1.43, 5.57)	0.0027	0.30 (0.33)	1.35 (0.71 to 2.55)	0.3639	
CDRS-R	430	0.29 (0.32)	1.33 (0.71 to 2.50)	0.3676	–0.40 (1.30)	(–2.95 to 2.15)	0.7566	0.04 (0.01)	1.05 (1.02, 1.07)	0.0002	0.28 (0.32)	1.32 (0.70 to 2.48)	0.3902	
PQ-LES-Q	453	0.21 (0.32)	1.24 (0.66 to 2.32)	0.5122	1.11 (0.91)	(–0.68 to 2.89)	0.2234	–0.08 (0.02)	0.92 (0.89, 0.96)	< 0.0001	0.32 (0.33)	1.37 (0.72 to 2.63)	0.3417	
SDQ total difficulties	463	0.27 (0.32)	1.31 (0.70 to 2.47)	0.3946	–0.70 (0.55)	(–1.78 to 0.39)	0.2074	0.12 (0.03)	1.13 (1.06, 1.20)	0.0001	0.38 (0.32)	1.47 (0.78 to 2.78)	0.2357	

continued

TABLE 82 Mediators analysis: assessment of the Baron and Kenny mediation steps[a] (*continued*)

Mediator variable	*n*	Step 1 [$Y = h_0(t)\exp(\beta_{10} X)$] Total effect of Rand on outcome – c (β_{10})			Step 2 ($Me = \beta_{20} + \beta_{21} X$) Effect of Rand on mediator – a (β_{21})			Step 3 [$Y = h_0(t)\exp(\beta_{31} X + \beta_{32} Me)$] Effect of mediator on outcome – b (β_{32})			Direct effect of Rand on outcome – c′ (β_{31})			Suggestion of potential for mediation
		Coefficient (SE)	HR (95% CI)	*p*-value	Coefficient (SE)	(95% CI)	*p*-value	Coefficient (SE)	HR (95% CI)	*p*-value	Coefficient (SE)	HR (95% CI)	*p*-value	
SDQ prosocial	463	0.28 (0.32)	1.33 (0.71 to 2.49)	0.3777	0.41 (0.16)	(0.10 to 0.72)	0.0098	–0.18 (0.09)	0.84 (0.70, 1.00)	0.0550	0.34 (0.32)	1.40 (0.74 to 2.64)	0.2967	
SDQ impact	325	0.12 (0.39)	1.13 (0.52 to 2.44)	0.7593	–0.56 (0.25)	(–1.05 to –0.07)	0.0253	0.24 (0.09)	1.28 (1.08, 1.51)	0.0047	0.32 (0.41)	1.37 (0.62 to 3.07)	0.4376	X
McMaster FAD overall	450	0.27 (0.33)	1.31 (0.69 to 2.49)	0.4139	–0.02 (0.03)	(–0.08 to 0.04)	0.4760	1.99 (0.50)	7.33 (2.75, 19.58)	0.0001	0.35 (0.33)	1.41 (0.74 to 2.71)	0.2985	
12-month caregiver questionnaires														
GHQ-12	445	0.30 (0.34)	1.35 (0.70 to 2.62)	0.3689	–0.71 (0.59)	(–1.86 to 0.45)	0.2296	0.04 (0.02)	1.04 (1.00, 1.09)	0.0758	0.34 (0.34)	1.41 (0.72 to 2.74)	0.3127	
SDQ total difficulties	447	0.16 (0.32)	1.18 (0.62 to 2.22)	0.6195	–1.35 (0.59)	(–2.52 to –0.18)	0.0238	0.05 (0.03)	1.05 (1.00, 1.11)	0.0635	0.22 (0.33)	1.24 (0.65 to 2.37)	0.5050	
SDQ emotional problems	447	0.15 (0.33)	1.17 (0.62 to 2.21)	0.6344	–0.52 (0.24)	(–0.99 to –0.04)	0.0332	0.14 (0.07)	1.15 (1.01, 1.31)	0.0390	0.24 (0.33)	1.27 (0.67 to 2.42)	0.4666	X
SDQ peer problems	447	0.17 (0.32)	1.18 (0.63 to 2.24)	0.6035	–0.34 (0.17)	(–0.68 to –0.00)	0.0471	0.08 (0.09)	1.08 (0.91, 1.29)	0.3815	0.20 (0.33)	1.22 (0.64 to 2.32)	0.5403	
SDQ impact	336	0.27 (0.37)	1.31 (0.63 to 2.69)	0.4675	–0.52 (0.30)	(–1.10 to 0.07)	0.0838	0.12 (0.07)	1.13 (0.99, 1.28)	0.0687	0.32 (0.37)	1.37 (0.66 to 2.85)	0.3924	
SDQ internalising	447	0.16 (0.32)	1.18 (0.62 to 2.22)	0.6191	–0.88 (0.36)	(–1.58 to –0.17)	0.0149	0.08 (0.04)	1.08 (0.99, 1.18)	0.0776	0.25 (0.33)	1.28 (0.67 to 2.44)	0.4546	
McMaster FAD overall	441	0.10 (0.33)	1.11 (0.58 to 2.12)	0.7541	–0.05 (0.03)	(–0.10 to 0.01)	0.0796	2.32 (0.61)	10.13 (3.06, 33.53)	0.0002	0.21 (0.34)	1.23 (0.63 to 2.39)	0.5398	
McMaster FAD roles	448	0.16 (0.32)	1.17 (0.62 to 2.22)	0.6234	–0.09 (0.03)	(–0.15 to –0.04)	0.0018	2.70 (0.55)	14.85 (5.03, 43.89)	< 0.0001	0.37 (0.34)	1.44 (0.75 to 2.80)	0.2755	X
McMaster FAD affective involvement	445	0.07 (0.33)	1.07 (0.56 to 2.04)	0.8426	–0.07 (0.04)	(–0.14 to 0.01)	0.0778	0.82 (0.40)	2.27 (1.04, 4.95)	0.0391	0.10 (0.33)	1.10 (0.57 to 2.12)	0.7701	

SI, suicidal ideation; Q, questionnaire.

a The effect of Rand is in reference to FT compared with TAU (TAU is the reference category). For the binary mediators psychotropic medications and Beck screening for suicide ideation the reference category is 'No', hence the effect of the mediator is in relation to yes; young person was on psychotropic medications during follow-up and suicide ideation indicated at 12 months, respectively; and for years spent working in CAMHS the reference category is < 4 years, hence the effect of the mediator is in relation to young person with a lead therapist working in CAMHS for ≥ 4 years.

b Binary mediators: step 2 uses a logistic regression model (excluding centre covariate).

Therapist alliance

Therapist alliance, reported in the FT arm, was further explored by participants' primary outcome. *Figures 41* and *42* present the total and subscale scores for SOFTA alliance as reported by the young person, caregiver and therapist according to whether or not the young person has a primary outcome event. The greatest differences for those with and without a primary outcome event, although not statistically significant, can be seen based on the young person's perception of purpose, engagement and safety and the caregiver's perception of safety, with lower scores (weaker alliance) for those with self-harm during follow-up. Note that the rate of self-harm was similar in participants who completed the SOFTA and those who did not, albeit slightly higher in the former (29.6% vs. 26.2% repetition rate, respectively).

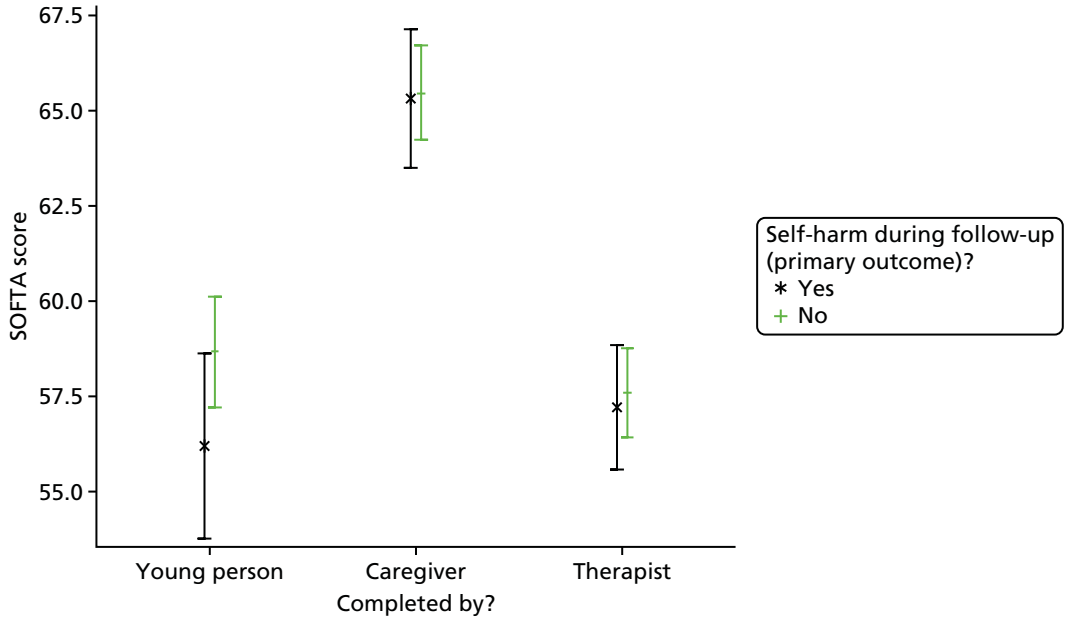


FIGURE 41 System for Observing Family Therapy Alliances total score by completer and primary outcome.

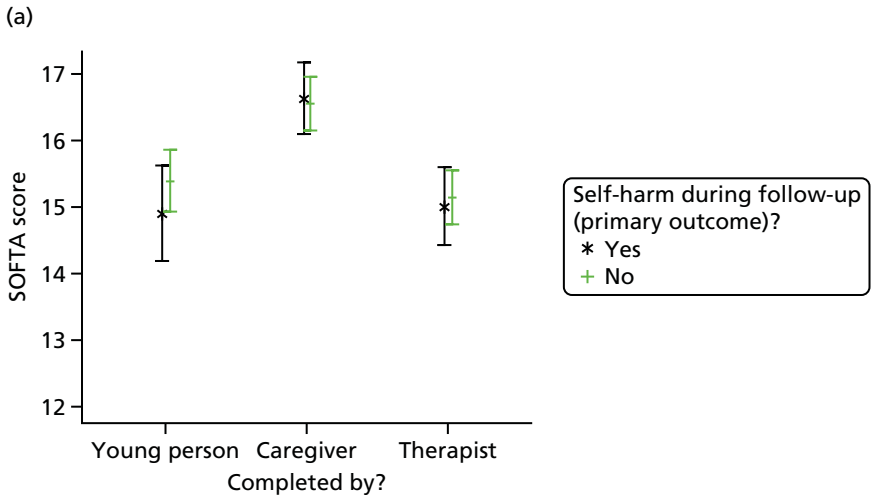


FIGURE 42 System for Observing Family Therapy Alliances subscale scores by completer and primary outcome. (a) Purpose subscale; (b) engagement subscale; (c) emotional subscale; and (d) safety subscale. (*continued*)

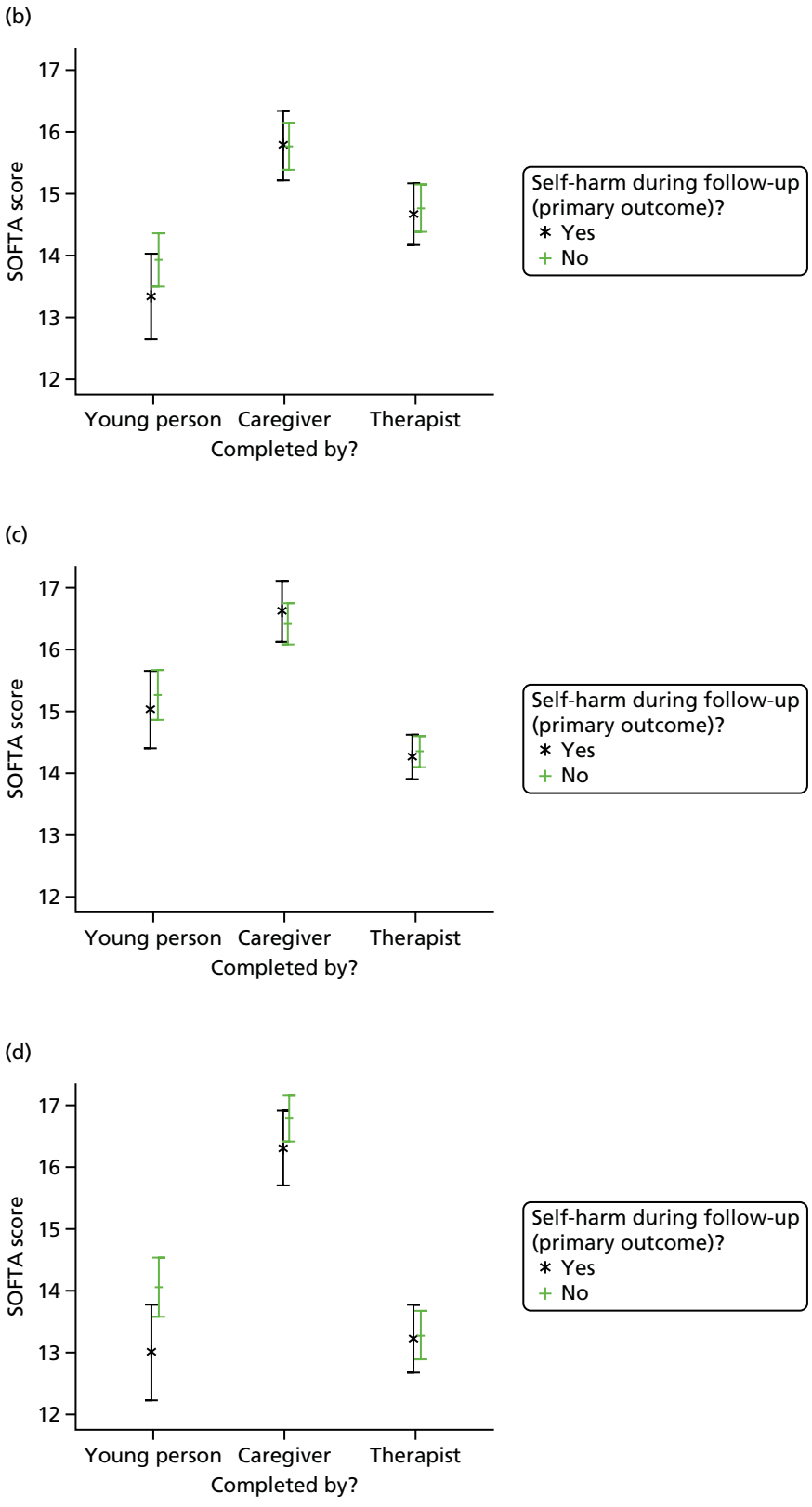


FIGURE 42 System for Observing Family Therapy Alliances subscale scores by completer and primary outcome. (a) Purpose subscale; (b) engagement subscale; (c) emotional subscale; and (d) safety subscale.

Chapter 4 Health economics results

Introduction

The health economics analysis was designed to provide an economic evaluation of the management of self-harm in adolescents using FT compared with TAU. The aim was to assess the cost-effectiveness of FT versus TAU in the management of self-harm in adolescents from the NHS perspective. A societal perspective for costs was not included in the sensitivity analysis owing to the very limited private expenses, productivity costs and out-of-pocket expenses reported in the trial.

Unit cost data

Unit cost of resource use

Individual-level resource use was combined with unit costs to calculate the total health services cost for each participant. Participants' resource use was self-reported by the young person as well as being reported separately by the caregiver. When missing, young people's resource use was replaced by the caregiver reports. Visits to A&E and inpatient visits were available from NHS Digital records; they included the type of event that led to a hospital visit, namely whether it was self-harm related or not self-harm related, and the admission and discharge dates. A&E visits were combined with the national average cost by type of event. The costs for hospital inpatient stays were calculated using the average length of stay and the type of event. Specifically, if the event that led to hospital admission was related to self-harm, hospital stays were costed at £430.72 for a day case, £413.79 for a non-elective overnight admission and £1373.43 for longer stays; equally, for events not related to self-harm, a day case was costed at £697.55, a non-elective overnight was costed at £602.52 and longer stays were costed at £2837.31. Excess bed-days were used when a young person experienced a length of stay beyond the average length of stay. For instance, if the event was related to self-harm, extra bed-day costs of £254.09 were added to the long stay costs of £1373.43. Information on unit costs was taken from national databases such as the PSSRU *Unit Costs of Health and Social Care 2014*[98] and the Department of Health's National Schedule of Reference Costs.[97] *Table 83* presents the summary of unit costs.

TABLE 83 Participant-reported health-care use and associated unit costs

Item	Unit cost (£)	Source
Health services		
GP surgery visit	46.00	PSSRU (2014),[98] p. 195, including direct care staff costs with qualification, per participant contact lasting 11.7 minutes
GP home visit	117.00[a]	PSSRU (2014),[98] p. 191, including direct care staff costs with qualification, per participant contact lasting 23.4 minutes
GP telephone/e-mail	28.00	PSSRU (2014),[98] p. 195, including direct care staff costs with qualification, per participant contact lasting 7.1 minutes
Practice or district nurse	59.50	PSSRU (2014),[98] p. 187 and pp.192 (average), per hour, including qualifications
Physiotherapist	36.00	PSSRU (2014),[98] p. 179, per hour, including qualifications
Occupational therapist	36.00	PSSRU (2014),[98] p. 180, per hour, including qualifications
Drug and alcohol worker	56.00	PSSRU (2014),[98] p. 67, per clinic consultation, including qualifications

continued

TABLE 83 Participant-reported health-care use and associated unit costs (*continued*)

Item	Unit cost (£)	Source
Family planning service	50.00	PSSRU (2014),[98] p. 212, per hour
CAMHS	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
Other (non-hospital-based health service)	57.00	*National Schedule of Reference Costs Year 2013–2014*,[97] community health services, national average
Hospital services		
Self-harm related		
Non-elective short stay	413.79	*National Schedule of Reference Costs Year 2013–2014*,[97] non-elective inpatients – short stay, code: WA11B
Non-elective long stay	1373.43	*National Schedule of Reference Costs Year 2013–2014*,[97] non-elective inpatients – long stay, code: WA11B
Excess day, non-elective	254.09	*National Schedule of Reference Costs Year 2013–2014*,[97] non-elective inpatients – excess bed-days, code: WA11B
Day case	430.72	*National Schedule of Reference Costs Year 2013–2014*,[97] day case, code: WA11B
Not self-harm related		
Non-elective short stay	602.52	*National Schedule of Reference Costs Year 2013–2014*,[97] no-elective inpatients – Short stay, national average
Non-elective long stay	2837.31	*National Schedule of Reference Costs Year 2013–2014*,[97] non-elective inpatients – long stay, national average
Excess day, non-elective	275.05	*National Schedule of Reference Costs Year 2013–2014*,[97] non-elective inpatients – excess bed-days, national average
Day case	697.55	*National Schedule of Reference Costs Year 2013–2014*,[97] day case, national average
Self-harm related/not self-harm related		
Hospital outpatient clinic	128.31	*National Schedule of Reference Costs Year 2013–2014*,[97] outpatients – consultant led, national average
Hospital A&E department	123.67	*National Schedule of Reference Costs Year 2013–2014*,[97] emergency medicine, national average

CCEMG-EPPI, Campbell & Cochrane Economics Methods Group – Evidence for Policy and Practice Information and Coordinating Centre Cost Converter.
a These costs have been adjusted to 2014 prices using the CCEMG-EPPI Centre Cost Converter (http://eppi.ioe.ac.uk/costconversion/).

Unit cost of medications

Unit costs for any psychotropic medication were obtained from the electronic market information tool (eMIT) [2014; URL: www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit (accessed 1 December 2016)] and the BNF.[99] Medication details including the dosage, date started and date stopped were collected directly from the participants and/or from CAMHS notes by the study researchers at baseline and at 12 and 18 months. Using the aforementioned dates, the number of days on which the medication was taken was calculated. To account for cases in which the dosage was not reported, the recommended dose for adolescents was referenced. In cases where the reported medication was not licensed for use in those aged < 18 years, the lowest recommended dose was used, as suggested by a clinical expert on the TMG. *Table 84* shows all the unit costs and standard dosage details for the psychotropic drugs that were reported in the trial. The unit costs, package and dosage details as well as the number of days the medication was taken were combined to calculate the cost of psychotropic medication for each participant.

TABLE 84 Young person-reported medication use, dosage and associated unit costs

Item	Standard dosage	Unit cost (£)	Source
Atomoxetine	40 mg (once per day)	15.38[a]	BNF 2015,[99] package of 7
Methylphenidate	5 mg (twice per day)	2.49	eMIT 2014, code: DDD026, package of 30
Propranolol	80 mg (twice per day)	2.55	eMIT 2014, code: DBD56, package of 56
Olanzapine	5 mg (once per day)	0.68	eMIT 2014, code: DDE000, package of 28
Risperidone	0.5 mg (once per day)	1.32[b]	eMIT 2014, code: DDB098, package of 60
Citalopram	20 mg (once per day)	0.28	eMIT 2014, code: DDC135, package of 28
Fluoxetine	10 mg (once per day)	0.26[b]	eMIT 2014, code: DDI014, package of 30
Sertraline	50 mg (once per day)	0.45	eMIT 2014, code: DDC039, package of 28
Venlafaxine	12.5 mg (once per day)	1.79[b]	eMIT 2014, code: DDI015, package of 56
Mirtazapine	15 mg (once per day)	1.08	eMIT 2014, code: DDC177, package of 28
Benzodiazepine – diazepam	2 mg (twice per day)	0.16	eMIT 2014, code: DDA062, package of 28
Benzodiazepine – temazepam	10 mg (once per day)	3.86	eMIT 2014, code: DDA054, package of 28
Melatonin	2–3 mg (once per day)	15.18[a]	BNF 2015, package of 30
Promethazine	10 mg (once per day)	2.75	eMIT 2014, code: DCD094, package of 56
Zopiclone	3.75 mg (once per day)	0.47	eMIT 2014, code: DDA045, package of 28

CCEMG-EPPI, Campbell & Cochrane Economic Methods Group – Evidence for Policy and Practice Information and Coordinating Centre Cost Converter.

a These costs have been adjusted to 2014 prices using the CCEMG-EPPI Centre Cost Converter [http://eppi.ioe.ac.uk/costconversion/ (accessed 2 December 2016)].

b For risperidone 1 mg, fluoxetine 20 mg and venlafaxine 37.5 mg tablets.

Unit cost of delivering the intervention

The cost of the intervention was calculated separately for the FT arm and the TAU arm. Participants allocated to TAU received care by local CAMHS teams, and clinicians (or researchers) recorded any treatment details (including duration, number of therapists involved in the session, type, attendance and telephone contact with the family between sessions). In addition, the frequency and duration of any supervision meetings were recorded.

Participants allocated to FT had sessions with qualified family therapists and details of these sessions were recorded in a similar way as in the TAU arm. Family therapists also received supervision, but it was reported differently from the supervision in the TAU arm; precisely, supervision in the FT arm was done in group sessions with the number of patients being discussed in each session, and frequency and duration being recorded. The unit costs for TAU and FT are presented in *Table 85*. It is important to note that PSSRU provides costs for CBT delivered in CAMHS, but not for any other modalities of treatment. Generic multidisciplinary CAMHS costs were thus used for all non-CBT in both arms; as tier three CAMHS teams often include family therapists and often provide FT, it was deemed a reasonable assumption to make in the absence of any other reported costs. The unit costs were combined with the duration, number of sessions reported and the number of therapists involved in the sessions in the trial data to provide the pragmatic intervention cost for the FT and TAU groups, which are referred to in the analysis as 'actual intervention costs'.

To test the sensitivity of the results, and as discussed with clinical experts, two other costing scenarios were explored for FT as an intervention. In the first scenario, it was assumed that only one therapist was involved in each treatment session by local CAMHS in the FT arm (intervention costs from scenario 1). In the second scenario, the average number of therapists involved in each CAMHS session in the FT arm was used (intervention costs from scenario 2). For the TAU arm, the number of therapists involved in each session, as recorded in the trial data, was used.

TABLE 85 Unit costs of delivering the interventions

Item	Unit cost (£)	Source
FT		
Qualified family therapist sessions	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
TAU (by local CAMHS)		
CBT	93.00	PSSRU (2014),[98] p. 94, per session
Formal systemic FT	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
Family work	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
Psychoeducation	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
Interpersonal therapy	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
Psychodynamic psychotherapy	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
DBT	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
Supporting therapy/ counselling	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
Problem-solving therapy	115.00	PSSRU (2014),[98] p. 223, per hour per team member for face-to-face contact
FT/TAU		
Telephone contact	91.00	PSSRU (2014),[98] p. 223, cost per hour per team member for participant-related activities
Supervision	56.00	PSSRU (2014),[98] p. 223, cost per hour per team member

The average number of CAMHS sessions per young person was 10 in the TAU arm, compared with seven in the FT arm (*Table 86*). In addition, these sessions on average were shorter and involved more therapists per session in the FT arm than in the TAU arm. The average number of telephone contacts with the family between treatment sessions differed only very slightly between trial arms, but calls lasted longer in the TAU arm. It was not possible to obtain an average number of supervision sessions per young person in the TAU arm as only the total amount of supervision per young person was recorded. Any conclusions on the duration of supervision should be treated with caution given the different ways in which supervision was reported in the two arms.

TABLE 86 Session details of delivering the intervention

	Average number of sessions per participant[a]		Average duration per participant (minutes)		Average number of therapists involved in each session[a]	
Session type	TAU	FT	TAU	FT	TAU	FT
CAMHS sessions/qualified family therapist sessions[b]	9.80 (13.73)	6.51 (3.90)	588.7	481.8	1.23 (0.50)	2.35 (0.45)
Telephone contact with family sessions	2.79 (2.82)	2.83 (2.26)	90.2	77.4	N/A	N/A
Supervision sessions	N/A[c]	2.11 (1.46)	88.4	77.5	N/A	N/A

N/A, not applicable.
a SD in parentheses.
b CAMHS sessions in the TAU arm and qualified family therapist sessions in the FT arm.
c It is not applicable for TAU because the frequency of supervision was reported rather than the actual number of sessions.

Missing data

A total of 832 young people were recruited to take part in the trial (417 were allocated to TAU and 415 to FT). The complete case refers to all those participants for whom complete EQ-5D and cost data are available at any time point and number 125 young people (44 allocated to TAU and 81 allocated to FT). Details of the quality of life and cost data available at each time point are shown in *Table 87*. Missing utility scores and/or total health services and hospital services costs at 6, 12 and 18 months were imputed using multiple imputation via chained equations assuming missingness at random at each time point in order to use as much as possible of the available information for each variable. For consistency with the statistical analysis, but also to ensure best fit of the imputed results, we used as predictor variables in the imputation process: treatment allocation, gender, age, centre, total number of self-harm episodes, type of index episode, source of referral and derived scores at baseline for a number of assessment instruments (BSS, CDRS-R, Hopelessness Scale for Children and PQ-LES-Q). It would be very speculative to impute missing utility at baseline; therefore, 37 participants (4.4%) are excluded from the health economics study sample. Additionally, imputation was not possible for a further 13 participants (i.e. it was not possible to impute any missing utility or cost data for these participants at any of the follow-up periods) and this was because of missing scores in at least one of the predictors. The base case sample (i.e. after imputation) was 782 participants (388 allocated to TAU and 394 allocated to FT); this is the sample used for the cost-effectiveness analyses.

Cost-effectiveness analysis

Descriptive statistics of participant characteristics and the additional predictors used in multiple imputations are presented in *Table 88*. More than 50% of the participants in the TAU or FT arm were between 11 and 14 years old. More than two-thirds of the participants in both arms were young women in whom more than three self-harm episodes occurred over the duration of the trial. Self-harm was primarily caused by self-injury and did not vary across trial arms. Differences in any of the assessment instruments were marginal across arms and more than 25% of the participants presented with mild or moderate depression at baseline. These results are consistent across samples (base case and complete case) and with the main statistical analysis in *Chapter 3*.

TABLE 87 Availability of quality-of-life and cost data (before imputation)

Quality of life and cost data	Baseline, *n* (%)	6 months, *n* (%)	12 months, *n* (%)	18 months, *n* (%)	Overall, *n* (%)
Quality of life					
EQ-5D[a]	795 (95.6)	341 (41.0)	455 (54.7)	390 (46.9)	**190 (22.8)**
Costs					
Health and social services costs[a]	N/A	405 (48.7)[b]	453 (54.4)	395 (47.5)	**221 (26.6)**
Hospital outpatient visits costs[a]	N/A	420 (50.5)[b]	426 (51.2)	363 (43.6)	**196 (23.6)**
Hospital inpatient stays and A&E visits costs[c]	N/A	832 (100.0)	832 (100.0)	832 (100.0)	**832 (100.0)**
Quality of life and costs	–	**250 (30.0)**	**410 (49.3)**	**355 (42.7)**	**125 (15.0)**

N/A, not applicable.
a Self-reported information.
b Non-missing costs at 3 months were multiplied by 2 to replace *only* those missing costs at 6 months.
c Information collected from participants' records.

TABLE 88 Participants' baseline characteristics by trial arm

	Base case		Complete case	
Baseline characteristic	TAU ($N = 388$)	FT ($N = 394$)	TAU ($N = 44$)	FT ($N = 81$)
Gender, *n* (%)				
Male	47 (12)	45 (11)	6 (14)	14 (17)
Female	341 (88)	349 (89)	38 (86)	67 (83)
Age (years), *n* (%)				
11–14	204 (53)	204 (52)	29 (66)	43 (53)
15–17	184 (47)	190 (48)	15 (34)	38 (47)
Centre, *n* (%)				
Yorkshire	142 (37)	141 (36)	15 (34)	39 (48)
Manchester	135 (35)	136 (35)	16 (36)	23 (28)
London	111 (29)	117 (30)	13 (30)	19 (23)
Total number of self-harm episodes, *n* (%)				
2	42 (11)	42 (11)	2 (5)	8 (10)
≥ 3	346 (89)	352 (89)	42 (95)	73 (90)
Type of index episode, *n* (%)				
Self-poisoning	87 (22)	89 (23)	7 (16)	21 (26)
Self-injury	274 (71)	280 (71)	34 (77)	56 (69)
Combined	27 (7)	25 (6)	3 (7)	4 (5)
Source of referral (from hospital), *n* (%)				
Yes	139 (36)	149 (38)	14 (32)	26 (32)
No	249 (64)	245 (62)	30 (68)	55 (68)
BSS				
Mean (SD)	10.07 (9.35)[a]	10.86 (8.89)[a]	11.23 (9.71)[a]	11.84 (8.82)
Min.	0	0	0	0
Max.	35	38	29	28
CDRS-R,[b] *n* (%)				
Not depressed (< 30)	23 (6)	41 (10)	5 (11)	11 (14)
Mild depression (30–42)	103 (27)	102 (26)	15 (34)	25 (31)
Moderate depression (43–57)	152 (39)	147 (37)	15 (34)	25 (31)
Severe depression (58–72)	96 (25)	88 (22)	9 (20)	19 (23)
Very severe depression (> 72)	13 (3)	16 (4)	–	1 (1)
Hopelessness Scale for Children				
Mean (SD)	7.18 (4.22)[c]	7.64 (4.27)[c]	8.14 (4.19)[c]	8.01 (4.38)[c]
Min.	0	0	1	1
Max.	17	17	15	16

TABLE 88 Participants' baseline characteristics by trial arm (*continued*)

	Base case		Complete case	
Baseline characteristic	TAU (*N* = 388)	FT (*N* = 394)	TAU (*N* = 44)	FT (*N* = 81)
PQ-LES-Q				
Mean (SD)	41.18 (9.53)	41.25 (9.28)	41.93 (9.31)	42.52 (9.72)
Min.	17	16	22	22
Max.	70	67	61	62

Max., maximum; min., minimum.
a The score was not available for one participant in the TAU arm (either base or complete case) and one participant in the FT arm (base case).
b The scale was not available for one participant in the TAU arm (base case).
c The score was not available for one participant in each trial arm (either base or complete case).

Intervention costs

Table 89 shows the average cost of each of the different components that were used to calculate the intervention costs. It is not possible to make any comparison between arms in terms of costs for the CAMHS or qualified family therapists, as part of the intervention, because these services were not offered to both arms. The average cost of telephone contacts with the family between sessions is slightly higher in the FT arm and the average cost of supervision is about four times higher in the FT arm than in the TAU arm.

TABLE 89 Average intervention costs by trial arm

	Cost (£)	
Type of service	TAU (*n* = 388)	FT (*n* = 394)
CAMHS		
Mean (SD)	800.73 (1412.65)	N/A
Min.	0	N/A
Max.	18,103.49	N/A
Qualified family therapists		
Mean (SD)	N/A	2075.51 (1506.28)
Min.	N/A	0
Max.	N/A	9265.98
Telephone contact		
Mean (SD)	56.05 (219.69)	59.58 (255.93)
Min.	0	0
Max.	3640	4095
Therapist's supervision		
Mean (SD)	18.50 (46.87)	48.08 (60.50)
Min.	0	0
Max.	373.35	438.70

Max., maximum; min., minimum; N/A, not applicable.
Reproduced from Cottrell *et al.*[70] © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Health-care resource use

Table 90 shows the average resource use per participant in each trial arm observed in the data for the complete cases only. Participants in the TAU arm were more likely to use the majority of health services than those in the FT arm, independently of the follow-up period. Similarly, participants in the TAU arm seemed to have used more social or hospital services than those in the FT arm. Any other conclusions should be tempered given that very little was reported in the trial follow-up questionnaires.

TABLE 90 Average resource use per participant by trial arm (complete case)

	0–6 months		6–12 months		12–18 months	
Health services	TAU ($n = 44$)	FT ($n = 81$)	TAU ($n = 44$)	FT ($n = 81$)	TAU ($n = 44$)	FT ($n = 81$)
GP (family doctor), surgery visit						
Mean (SD)	0.84 (1.07)	0.70 (1.36)	1.05 (1.41)	0.96 (1.29)	1.18 (1.81)	1.16 (1.58)
Min.	0	0	0	0	0	0
Max.	4	9	6	6	10	8
GP (family doctor), home visit						
Mean (SD)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Min.	0	0	0	0	0	0
Max.	0	0	0	0	0	0
GP (family doctor), telephone/e-mail						
Mean (SD)	0 (0)	0 (0)	0.09 (0.47)	0.02 (0.16)	0.02 (0.15)	0 (0)
Min.	0	0	0	0	0	0
Max.	0	0	3	1	1	0
Practice or district nurse						
Mean (SD)	0.18 (0.66)	0.14 (0.54)	0.11 (0.39)	0.21 (0.52)	0.09 (0.42)	0.19 (0.53)
Min.	0	0	0	0	0	0
Max.	4	3	2	3	2	3
Physiotherapist						
Mean (SD)	0 (0)	0 (0)	0.07 (0.33)	0.05 (0.27)	0.02 (0.15)	0.21 (1.21)
Min.	0	0	0	0	0	0
Max.	0	0	2	2	1	9
Occupational therapist						
Mean (SD)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.01 (0.11)
Min.	0	0	0	0	0	0
Max.	0	0	0	0	0	1
Drug and alcohol worker						
Mean (SD)	0 (0)	0 (0)	0 (0)	0.14 (1.22)	0 (0)	0.01 (0.11)
Min.	0	0	0	0	0	0
Max.	0	0	0	11	0	1
Family planning service						
Mean (SD)	0.07 (0.33)	0.07 (0.31)	0.05 (0.21)	0.04 (0.19)	0.09 (0.36)	0.05 (0.31)
Min.	0	0	0	0	0	0
Max.	2	2	1	1	2	2

TABLE 90 Average resource use per participant by trial arm (complete case) (*continued*)

	0–6 months		6–12 months		12–18 months	
Health services	**TAU ($n = 44$)**	**FT ($n = 81$)**	**TAU ($n = 44$)**	**FT ($n = 81$)**	**TAU ($n = 44$)**	**FT ($n = 81$)**
CAMHS						
Mean (SD)	2.36 (4.00)	2.55 (10.41)	2.18 (4.13)	1.21 (2.48)	2.30 (4.55)	1.05 (2.51)
Min.	0	0	0	0	0	0
Max.	17	90	15	12	20	12
Any other non-hospital-based health service (e.g. NHS Direct)						
Mean (SD)	0.07 (0.33)	0.11 (0.61)	0.09 (0.36)	0.05 (0.22)	0 (0)	0.09 (0.57)
Min.	0	0	0	0	0	0
Max.	2	5	2	1	0	5
Social services						
Social worker						
Mean (SD)	0.16 (0.68)	0.28 (1.63)	0.07 (0.33)	0.10 (0.44)	0.14 (0.90)	0.02 (0.22)
Min.	0	0	0	0	0	0
Max.	4	14	2	3	6	2
Helpline (Childline, Samaritans)						
Mean (SD)	0 (0)	0 (0)	0.02 (0.15)	0.15 (0.87)	0.02 (0.15)	0.02 (0.22)
Min.	0	0	0	0	0	0
Max.	0	0	1	6	1	2
Family or young person support or self-help groups						
Mean (SD)	0.84 (2.98)	0.22 (1.11)	0.73 (2.73)	0.27 (1.41)	0.27 (1.52)	0.02 (0.16)
Min.	0	0	0	0	0	0
Max.	12	8	12	8	10	1
Any other social services						
Mean (SD)	0 (0)	0 (0)	0 (0)	0.04 (0.33)	0 (0)	0.01 (0.11)
Min.	0	0	0	0	0	0
Max.	0	0	0	3	0	1
Hospital services						
Hospital inpatient stay (staying in hospital overnight)						
Mean (SD)	0.12 (0.63)	0.40 (1.36)	0.49 (1.26)	0.35 (0.78)	0.58 (1.78)	0.42 (1.02)
Min.	0	0	0	0	0	0
Max.	4	10	7	4	10	7
Hospital outpatient clinic (doctor visits, scans, etc.)						
Mean (SD)	0.30 (0.67)	0.33 (0.85)	0.30 (0.79)	0.26 (0.74)	0.18 (0.58)	0.12 (0.43)
Min.	0	0	0	0	0	0
Max.	3	4	4	5	3	2
Hospital A&E department						
Mean (SD)	0.09 (0.60)	0.16 (0.43)	0.32 (0.77)	0.19 (0.57)	0.41 (1.41)	0.30 (0.66)
Min.	0	0	0	0	0	0
Max.	4	2	3	4	9	3

Max., maximum; min., minimum.

Mean health-care provider costs broken down by type of service, trial arm and for each time point are presented in *Table 91*. With or without imputing for missing data, it is difficult to see any patterns between the different follow-up periods, which is consistent with the discussion earlier on the participant-reported resource use.

TABLE 91 Average health-care provider costs by trial arm

Health-care provider costs (£)	0–6 months		6–12 months		12–18 months[a]	
Before imputation	***TAU (n = 44)***	***FT (n = 81)***	***TAU (n = 44)***	***FT (n = 81)***	***TAU (n = 44)***	***FT (n = 81)***
Health and social services costs (excluding intervention cost)						
Mean (SD)	789.06 (942.62)	703.20 (1312.58)	726.77 (1125.69)	491.85 (678.42)	698.61 (1045.22)	427.11 (679.17)
Min.	0	0	0	0	0	0
Max.	4571	10,994	4894	2852	4611.98	3375.44
Hospital services costs from the NHS Digital records (inpatient stays and A&E visits)[b]						
Mean	294.98 (1036.52)	367.69 (992.05)	353.18 (857.64)	259.32 (633.66)	479.53 (1455.01)	313.7 (737.59)
Min.	0	0	0	0	0	0
Max.	6572.4	6784.06	4480.41	3286.2	7996.47	4480.41
Reported hospital outpatient visits costs						
Mean (SD)	75.82 (144.63)	83.96 (179.36)	75.82 (203.95)	66.53 (189.35)	46.66 (149.25)	31.68 (110.25)
Min.	0	0	0	0	0	0
Max.	513.24	898.17	1026.48	1283.1	769.86	513.24
Medication costs[b]						
Mean (SD)	0.07 (0.31)	0.01 (0.08)	0.06 (0.30)	0 (0)	0.03 (0.22)	0 (0)
Min.	0	0	0	0	0	0
Max.	1.82	0.64	1.83	0	1.47	0
After imputation[c]	***TAU (n = 388)***	***FT (n = 394)***	***TAU (n = 388)***	***FT (n = 394)***	***TAU (n = 388)***	***FT (n = 394)***
Health and social services costs						
Mean (SD)	639.30 (846.01)	492.29 (834.68)	359.52 (677.17)	411.65 (714.27)	405 (829.79)	355.19 (605.82)
Min.	0	0	0	0	0	0
Max.	9648	10,994	4894	5520	7195.17	3829.56
Hospital services costs from the NHS Digital records (inpatient stays and A&E visits)[b]						
Mean (SD)	484.37 (1067.78)	490.66 (1033.11)	421.4 (885.74)	441.84 (1121.65)	430.08 (1024.74)	480.34 (1133.52)
Min.	0	0	0	0	0	0
Max.	8522.76	8693.99	4931.94	11,747.48	8960.82	10,785.98

TABLE 91 Average health-care provider costs by trial arm (*continued*)

Health-care provider costs (£)	0–6 months		6–12 months		12–18 months[a]	
Reported hospital outpatient visits costs						
Mean (SD)	57.82 (184.9)	64.67 (354.82)	29.75 (120.87)	46.05 (161.86)	30.11 (213.02)	25.93 (135.14)
Min.	0	0	0	0	0	0
Max.	1411.41	6415.5	1026.48	1283.1	3849.3	1539.72
Medication costs[b]						
Mean (SD)	0.38 (5.88)	1.11 (20.20)	0.29 (4.77)	1.11 (20.31)	0.05 (0.58)	1.09 (19.96)
Min.	0	0	0	0	0	0
Max.	114.25	399.85	93.51	402.05	8.81	395.19

Max., maximum; min., minimum.
a These are the discounted estimates.
b Hospital services costs from IC records and medication costs were not imputed.
c Include estimated costs after imputation – these figures were used in the cost–utility analyses.

The total costs to the NHS providers are shown in *Table 92*. They comprise the total health and social services costs, total hospital costs (separately by data source), total intervention costs and medication costs, as well as any costs of appointments that occurred after randomisation, but before the first treatment appointment (only in the FT arm). Overall, they substantially vary from one participant to the other, as illustrated by the large SDs.

In detail, the average total costs for the use of health and social services (before imputation) were £2214.44 (SD £2334.32) for the TAU arm compared with £1622.16 (SD £1866.72) for the FT arm. Furthermore, the average total costs for the use of any hospital services (before imputation) were higher for the participants in the TAU arm: the average cost of inpatient stays and A&E visits was £1127.7 (SD £2782.9) in the TAU arm, compared with £940.71 (SD £1563.95) in the FT arm. The corresponding costs of outpatient visits were £198.3 (SD £405.24) in the TAU arm and £182.17 (SD £299.45) in the FT arm. The mean total costs of any psychotropic medications were also higher for the TAU group. However, the average total intervention costs observed within the trial were much higher in the FT arm [£2928.74 (SD £1541.14)] than in the TAU arm [£1005.41 (SD £1034.46)]. The intervention costs and the cost of appointments that took place post randomisation but before the first treatment appointment in the FT arm are likely to drive the average total costs to the NHS. Pragmatically, the mean total costs to the NHS (before imputation) were £5588.21 (SD £3399.39) for the FT arm and £4461.22 (SD £4280.41) for the TAU arm. Most patterns remain even after imputing for the missing cost data; an exception is the average total costs of reported hospital outpatient visits, which were higher for the FT arm. Mann–Whitney tests show the average total NHS costs to be marginally different (at the 5% significance level) between the trial arms, but significantly different (at the 1% significance level) after the imputations.

Quality of life

Table 93 shows the mean EQ-5D scores at each time point between the two arms of the trial when scores were not imputed (complete case) and when scores were imputed (base case). In both arms, there was an increase in EQ-5D from baseline to 18 months. The fluctuations between 6 months and 12 months, and between 12 months and 18 months, were small for both arms.

TABLE 92 Total costs of NHS resources used by trial arm

Before imputation	TAU ($n = 44$)	FT ($n = 81$)	p-value (Mann–Whitney test)
Total NHS costs[a] (including actual intervention cost) (£)			
Mean (SD)	4461.22 (4280.41)	5588.21 (3399.39)	0.012
Min.	164.01	1012.4	
Max.	17,889.53	17,546.38	
Total actual intervention costs[b] (£)			
Mean (SD)	1005.41 (1034.46)	2928.74 (1541.14)	< 0.001
Min.	0	121.39	
Max.	4329.91	9349.98	
Total health and social services costs (£)			
Mean (SD)	2214.44 (2334.32)	1622.16 (1866.72)	0.426
Min.	0	0	
Max.	9712.70	12,451.16	
Total hospital services costs from IC records (inpatient stays and A&E visits)[b] (£)			
Mean (SD)	1127.7 (2782.9)	940.71 (1563.95)	0.438
Min.	0	0	
Max.	15,390.42	7766.61	
Total reported hospital outpatient visits costs (£)			
Mean (SD)	198.3 (405.24)	182.17 (299.45)	0.464
Min.	0	0	
Max.	2052.96	1539.72	
Total costs of appointments post randomisation but before first treatment appointment[b] (£)			
Mean (SD)	N/A	4.26 (28.43)	N/A
Min.	N/A	0	
Max.	N/A	230	
Medication costs[b] (£)			
Mean (SD)	0.17 (0.81)	0.01 (0.08)	0.410
Min.	0	0	
Max.	5.12	0.64	
After imputation[c]	**TAU ($n = 388$)**	**FT ($n = 394$)**	**p-value (Mann–Whitney test)**
Total NHS costs[a] (including actual intervention costs) (£)			
Mean (SD)	3725.49 (3785.98)	4991.72 (3766.88)	< 0.001
Min.	164.01	403.37	
Max.	29,215.88	32,085.23	
Total actual intervention costs[b] (£)			
Mean (SD)	875.28 (1471.21)	2183.17 (1558.77)	< 0.001
Min.	0	0	
Max.	18,683.46	9349.98	
Total health and social services costs (£)			
Mean (SD)	1403.82 (1760.43)	1259.13 (1462.82)	0.121
Min.	0	0	
Max.	15,650.02	12,451.16	

TABLE 92 Total costs of NHS resources used by trial arm (*continued*)

Before imputation	TAU ($n = 44$)	FT ($n = 81$)	p-value (Mann–Whitney test)
Total hospital services costs from NHS Digital records (inpatient stays and A&E visits)[b] (£)			
Mean (SD)	1335.85 (2217.43)	1412.85 (2550.78)	0.285
Min.	0	0	
Max.	15,390.42	24,190.24	
Total reported hospital outpatient visits costs (£)			
Mean (SD)	117.67 (325.83)	136.65 (444.62)	0.783
Min.	0	0	
Max.	3852.82	7194.35	
Total costs of appointments post randomisation but before first treatment appointment[b] (£)			
Mean (SD)	N/A	5.25 (39.84)	N/A
Min.	N/A	0	
Max.	N/A	575	
Medication costs[b] (£)			
Mean (SD)	0.72 (11.04)	3.31 (60.47)	0.709
Min.	0	0	
Max.	214.8	1197.1	

Max., maximum; min., minimum; N/A, not applicable.
a The cost of appointments that took place after randomisation but before the first treatment appointment in the FT arm is also included.
b These costs were not imputed.
c Includes estimated costs after imputation – these figures were used in the cost–utility analyses.
Reproduced from Cottrell *et al.*[70] © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

TABLE 93 Mean EQ-5D scores (unadjusted) by trial arm

	Complete case		Base case (imputed)	
Time point	TAU ($n = 44$), mean (SD)	FT ($n = 81$), mean (SD)	TAU ($n = 388$), mean (SD)	FT ($n = 394$), mean (SD)
Baseline	0.663 (0.260)	0.663 (0.263)	0.682 (0.264)	0.674 (0.277)
6 months	0.798 (0.203)	0.814 (0.201)	0.760 (0.161)	0.799 (0.178)
12 months	0.841 (0.176)	0.825 (0.229)	0.780 (0.181)	0.811 (0.193)
18 months	0.865 (0.184)	0.852 (0.204)	0.804 (0.159)	0.813 (0.189)

The change in mean EQ-5D between the two arms of the trial at any time point is presented in *Table 94*. The observed changes in either the complete or the base case were marginal. The largest changes were observed at 6 months and 12 months, showing mostly that participants in the FT arm had higher utility scores than participants in the TAU arm; differences were not significant in the complete case but were significant at 5% significance level in the base case at 6 and 12 months.

TABLE 94 Changes in mean EQ-5D scores (unadjusted) between the two arms of the trial

Time point	Difference[a]	Complete case	Base case (imputed)
Baseline	FT vs. TAU (*p*-value)	–0.001 (0.988)	0.008 (0.680)
6 months	FT vs. TAU (*p*-value)	–0.016 (0.669)	–0.039 (0.002)
12 months	FT vs. TAU (*p*-value)	0.016 (0.688)	–0.031 (0.020)
18 months	FT vs. TAU (*p*-value)	0.013 (0.731)	–0.009 (0.493)

a *p*-value of *t*-test of the difference.

Cost-effectiveness results

Primary analysis

Table 95 shows the costs and QALYs for the TAU and FT arm for the primary analysis. In addition, it provides the incremental cost and incremental effectiveness expressed as QALY gains. On average, FT participants incurred £1266.23 (95% CI £736.04 to £1796.43) higher costs and gained 0.034 (95% CI –0.004 to 0.065) extra QALYs than TAU participants, which is equivalent to an extra 12.4 days of perfect health. The ICER equalled £36,811.80 per QALY, which is above the recommended threshold range currently specified for NICE decision-making in England and Wales (£20,000–30,000 per QALY gain), indicating that FT was unlikely to be cost-effective.[93]

Figure 43 shows the cost-effectiveness plane for FT compared with TAU based on 10,000 bootstrapped estimates of costs and QALYs. The diagonal line represents a willingness to pay per QALY threshold of £20,000. The average costs from the bootstrapped estimates were £3725.03 (SD £191.31) and £4979.54 (SD £190.01) for the TAU and FT arms, respectively. The corresponding mean QALYs were 1.122 (SD 0.010) for the TAU arm and 1.156 (SD 0.011) for the FT arm. The simulation estimates were all above the x-axis, showing that FT was always more costly than TAU. Most of the estimates were spread in the north-east quadrant, suggesting that FT was likely to lead to better health outcomes, although some estimates were also in the north-west quadrant, where FT is likely to result in a lower number of QALYs than TAU.

TABLE 95 Base-case cost-effectiveness results (outcome measure: QALY, NHS perspective including actual intervention costs)

Treatment arm	Cost-effectiveness
TAU	
Costs (£), mean (SD)	3725.49 (3785.98)
QALY, mean (SD)	1.122 (0.203)
FT	
Costs (£), mean (SD)	4991.72 (3766.88)
QALY, mean (SD)	1.157 (0.226)
FT vs. TAU	
Incremental cost (95% CI)	1266.23 (736.04 to 1796.43)
Incremental QALY (95% CI)	0.034 (–0.004 to 0.065)
ICER (£/QALY)	36,811.80

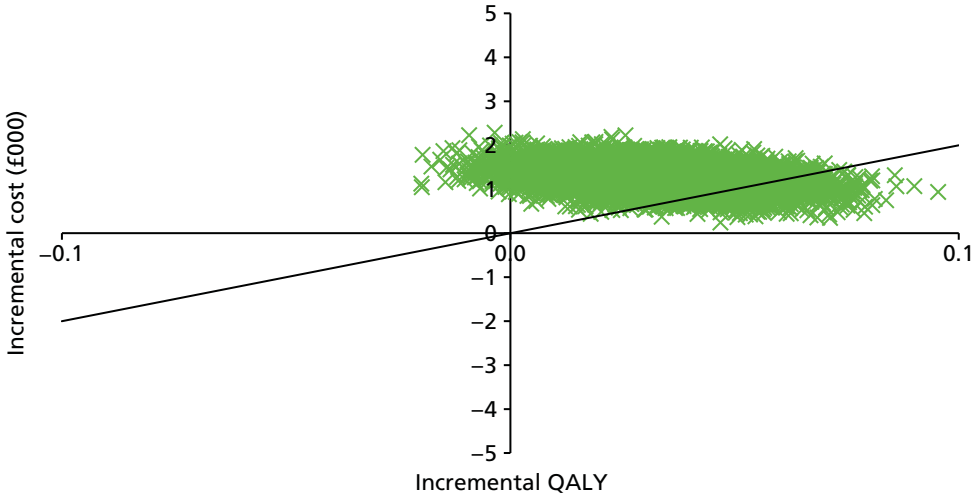


FIGURE 43 Base-case cost-effectiveness plane of FT compared with TAU (outcome measure: QALY, NHS perspective).

A cost-effectiveness acceptability curve of FT compared with TAU is presented in *Figure 44*. At a threshold of £20,000, FT had a 12% chance of being cost-effective, and this percentage increased to 36% when the threshold was £30,000. There is a positive relationship between the cost-effectiveness threshold and the chance of FT being cost-effective, and this is because FT was, on average, more effective (in terms of QALY gains) than TAU.

To account for uncertainty in the incremental costs and QALYs estimation, a number of sensitivity analyses were conducted, including a non-parametric bootstrapping (*Table 96*). The mean incremental cost and QALY estimates from the bootstrapping were along the lines of the deterministic base-case scenario, yielding an ICER of £36,705.79 per QALY gain. Other sensitivity analyses included making different assumptions about the number of therapists involved in each of the treatment sessions in the FT arm, that is, using either intervention costs from scenario 1 or intervention costs from scenario 2 as defined earlier in the analysis; in any of these scenarios, FT remained unlikely to be cost-effective. When adjusting for EQ-5D differences at baseline, participants in the FT arm gained an incremental QALY of 0.039 (equivalent to an extra 14.2 days at perfect health) compared with those in the TAU arm and the ICER was £32,852.22 per QALY gain. In any of these analyses, the ICER lay above the recommended NICE threshold (£20,000–30,000) and it was concluded that FT was unlikely to be cost-effective. Considering the complete cases only (i.e. no missing costs or quality of life data at any time point), FT was dominated by TAU (i.e. FT was more costly and less effective). Finally, information was exploited on caregivers' quality of life and cumulated QALY gains for both young people and

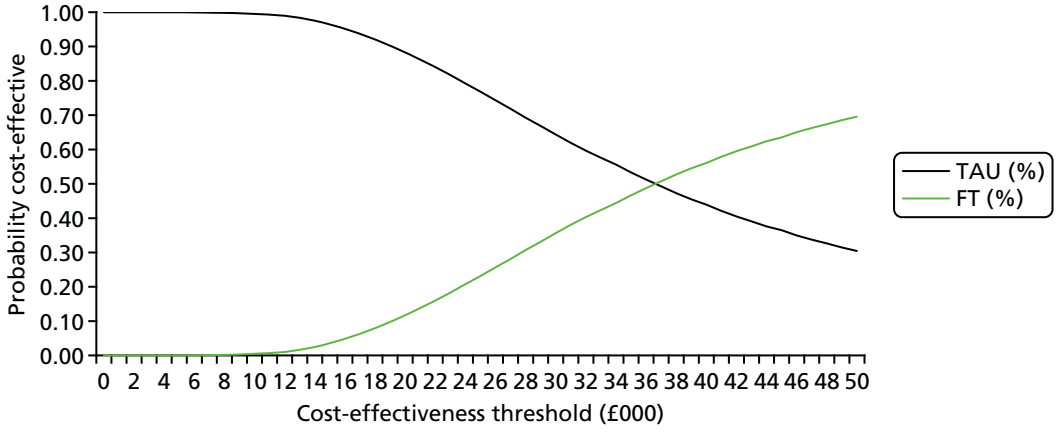


FIGURE 44 Base-case cost-effectiveness acceptability curve of FT compared with TAU (outcome measure: QALY, NHS perspective).

TABLE 96 Sensitivity analyses (outcome measure: QALY, NHS perspective)

FT vs. TAU	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER (£/QALY)
Bootstrapped average (10,000 replications)	1254.51 (1149.23 to 1259.80)	0.034 (0.0338 to 0.0344)	36,705.79
Assume intervention costs from scenario 1	1380.36 (747.98 to 2012.75)	0.034 (–0.004 to 0.065)	40,129.79
Assume intervention costs from scenario 2	1546.38 (909.59 to £2183.17)	0.034 (–0.004 to 0.064)	44,956.26
Adjusting for baseline EQ-5D differences	1266.23 (736.04 to 1796.43)	0.039 (0.035 to 0.042)	32,852.22
Complete case	1135.13 (267.39 to 2537.64)	–0.003 (–0.086 to 0.080)	FT dominated
Including caregivers' QALYs	1207.16[a] (662.33 to 1752.00)	0.058 (0.002 to 0.114)	20,808.21

a The incremental cost is different from that presented in *Table 95* because the sample size in this analysis is different; 364 participants and their caregivers in the TAU arm and 379 participants and their caregivers in the FT arm were considered.

caregivers. An aggregate QALY was created by simply summing the young people's and caregivers' QALYs.[103] The associated cost-effectiveness ratio was then £20,808.21 per QALY gain, demonstrating a potential for FT to bring a total of 21.2 extra days at full health for both the young person and caregiver and hence to be cost-effective, as it was below the recommended NICE threshold (£20,000–30,000). The cost-effectiveness plane and the cost-effectiveness acceptability curve of FT compared with TAU when young people's and caregivers' quality of life are taken into account can be found in *Appendix 7*.

Secondary analysis

Table 97 shows the costs and the number of self-harm events for each of the trial arms to the end of trial follow-up (18 months). It also shows the incremental cost and incremental number of self-harm events avoided because of FT. More self-harm events were avoided in the TAU arm than in the FT arm over the full trial period and, together with the higher cost to the NHS providers of FT, suggests that FT was dominated by TAU (i.e. FT was both more costly and less effective).

TABLE 97 Cost-effectiveness results (outcome measure: self-harm events avoided, NHS perspective)

Treatment arm	Cost-effectiveness result
TAU	
Costs (£),[a] mean (SD)	3733.85 (3788.80)
Number of self-harm events, mean (SD)	0.468 (1.268)
FT	
Costs (£),[a] mean (SD)	4986.41 (3758.29)
Number of self-harm events, mean (SD)	0.501 (1.063)
FT vs. TAU	
Incremental cost (£), (95% CI)	1252.57 (724.80 to 1780.33)
Incremental number of self-harm events (95% CI)	0.033 (–0.130 to 0.197)
ICER (£/self-harm event avoided)	FT dominated

a The costs are different from those presented in *Table 95* because the sample size in the secondary analysis is different; 391 participants in the TAU arm and 397 participants in the FT arm were considered.

Decision model analysis

As the difference in QALY gains between FT and TAU was marginal at 18 months in the within-trial analysis and did not exist in the complete-case analysis, a decision-analysis model would not have been required. However, as planned in the original protocol, a discrete-time state-transition (modified Markov) model was developed in order to estimate the cost-effectiveness of FT compared with TAU over a horizon of 5 years. In line with the within-trial analysis, the base-case model adopted a NHS perspective and future costs and QALYs were discounted at an annual rate of 3.5% following the NICE guidelines.[93] The model was built and analysed using Microsoft Excel® version 2013 (Microsoft Corporation, Redmond, WA, USA).

Model parameters

The model structure was presented in *Figure 3*. The full list of the model parameters and distributions applied in the model is given in *Table 98*, with the chosen distributions being based on the observed variance data.

TABLE 98 Markov input parameters

Parameter	Name	Mean	Distribution	SE	Source
Global parameters	Discount rate	0.035	Fixed		NICE guidance
Health state costs in TAU arm (6 months)	SH	£1182	Log-normal	£1493	SHIFT trial data
Health state costs in TAU arm (12 months)	SH	£1698	Log-normal	£1628	
	noSH	£709	Log-normal	£1116	
	Death	0	Fixed	–	
Health state costs in TAU arm (18 months)	SH	£1510	Log-normal	£1022	
	noSH	£817	Log-normal	£1432	
	Death	0	Fixed	–	
Health state costs in FT arm (6 months)	SH	£1049	Log-normal	£1482	
Health state costs in FT arm (12 months)	SH	£2186	Log-normal	£2198	
	noSH	£763	Log-normal	£1228	
	Death	0	Fixed	–	
Health state costs in FT arm (18 months)	SH	£2530	Log-normal	£2282	
	noSH	£649	Log-normal	£1054	
	Death	0	Fixed	–	
Health state utilities in TAU arm (6 months)	SH	0.760	Beta	0.161	
Health state utilities in TAU arm (12 months)	SH	0.751	Beta	0.187	
	noSH	0.784	Beta	0.180	
	Death	0	Fixed	–	
Health state utilities in TAU arm (18 months)	SH	0.754	Beta	0.033	
	noSH	0.808	Beta	0.157	
	Death	0	Fixed	–	
Health state utilities in FT arm (6 months)	SH	0.799	Beta	0.178	
	SH	0.793	Beta	0.184	

continued

TABLE 98 Markov input parameters (*continued*)

Parameter	Name	Mean	Distribution	SE	Source
Health state utilities in FT arm (12 months)	noSH	0.813	Beta	0.194	
	Death	0	Fixed	–	
Health state utilities in FT arm (18 months)	SH	0.732	Beta	0.239	
	noSH	0.823	Beta	0.179	
	Death	0	Fixed	–	
Transition probabilities (at 6 months)	Proportion of young persons stopping SH (noSH) from SH in the TAU arm	0.858	Beta	0.0003	
	Proportion of young persons stopping SH (noSH) from SH in the FT arm	0.845	Beta	0.0003	
Transition probabilities (at 12 months)	Proportion of young persons stopping SH (noSH) from SH in the TAU arm	0.716	Beta	0.004	
	Proportion of young persons SH from noSH in the TAU arm	0.066	Beta	0.0002	
	Proportion of young persons stopping SH (noSH) from SH in the FT arm	0.803	Beta	0.003	
	Proportion of young persons SH from noSH in the FT arm	0.078	Beta	0.0002	
Transition probabilities (at 18 months)	Proportion of young persons stopping SH (noSH) from SH in the TAU arm	0.775	Beta	0.004	
	Proportion of young persons SH from noSH in the TAU arm	0.063	Beta	0.0002	
	Proportion of young persons stopping SH (noSH) from SH in the FT arm	0.684	Beta	0.006	
	Proportion of young persons SH from noSH in the FT arm	0.095	Beta	0.0002	

noSH, no self-harm; SH, self-harm.

Base-case model results

The results of the base-case analyses are presented in *Tables 99* and *100*. Taking into account uncertainty around the model parameters (i.e. looking at the probabilistic results), FT was associated with an extra cost of £1262.13 and a 5-year QALY gain of 0.065 (equivalent to 23.7 extra days of full health) compared with TAU. The ICER indicated that 1 QALY would be gained for every £19,486.97 spent by adopting FT; the ICER was below the NICE-recommended threshold (£20,000–30,000) and, therefore, FT was expected to be cost-effective in the longer term.

The results of the probabilistic analysis are presented in *Figure 45*. Each point on the graph represents the result of one probabilistic simulation of the model and indicates a potential incremental cost and incremental QALY for FT compared with TAU. The diagonal line represents the NICE willingness-to-pay threshold of £20,000 per QALY. The points were widely distributed around the origin, in both cost-effective and non-cost-effective regions, which indicated a high level of uncertainty around the cost-effectiveness of FT compared with TAU over a 5-year horizon.

TABLE 99 Deterministic 5-year model cost-effectiveness results

Treatment arm	Cost-effectiveness result
TAU	
Total costs (£)	10,338.77
Total QALYs	4.201
FT	
Total costs (£)	11,650.18
Total QALYs	4.272
FT vs. TAU	
Incremental cost (£)	1311.41
Incremental QALY	0.070
ICER (£/QALY)	18,659.74

TABLE 100 Probabilistic 5-year model cost-effectiveness results

Treatment arm	Cost-effectiveness result
TAU	
Costs (£), mean (SD)	11,0301.58 (11,092.43)
QALY, mean (SD)	4.187 (0.623)
FT	
Costs (£), mean (SD)	11,563.70 (8,111.12)
QALY, mean (SD)	4.251 (0.698)
FT vs. TAU	
Incremental cost (£), (95% CI)	1262.13 (1106.62 to 1417.62)
Incremental QALY (95% CI)	0.065 (0.053 to 0.075)
ICER (£/QALY)	19,486.97

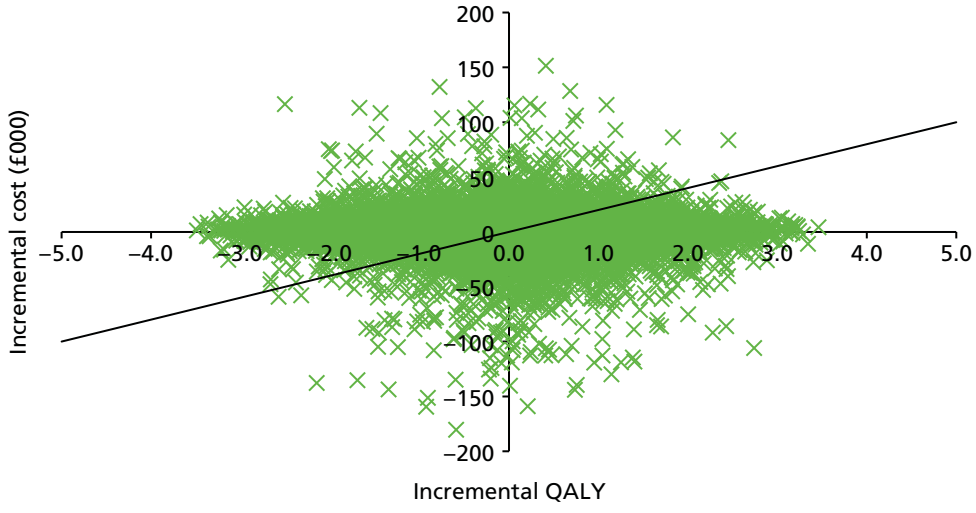


FIGURE 45 Long-term cost-effectiveness plane of FT compared with TAU (outcome measure: QALY, NHS perspective).

The cost-effectiveness acceptability curve presented in *Figure 46* shows the proportion of model simulation points being under different cost-effectiveness threshold values and indicated the probability that each treatment was cost-effective at given willingness-to-pay values. At low cost per additional QALY thresholds, FT was associated with a low probability of being cost-effective. As the threshold value increased, the probability of FT being cost-effective increased marginally. In particular, at a threshold of £20,000 per QALY, FT had a 50% chance of being cost-effective, and this increased to 52% when the threshold was £30,000.

Sensitivity analyses

A number of deterministic one-way sensitivity analyses were undertaken, the results of which are presented in *Table 101*. They included making different assumptions about the number of therapists involved in each of the treatment sessions in the FT arm, similarly to the within-trial analysis; we considered intervention costs from scenario 1 and then intervention costs from scenario 2, as defined earlier in the analysis. In both cases, participants in the FT arm incurred higher costs than those in the TAU arm, but with an incremental QALY of 0.070, giving ICERs between £1000 and £4500 per QALY, implying that FT was likely to be cost-effective in the longer term. Another analysis considered using the costs and utilities from the trial data at 6 months but without using the observed costs and utilities at 12 or 18 months. In this case, participants in the FT arm incurred costs that were £1041.55 higher than those incurred in the TAU arm and achieved a 5-year QALY gain of 0.077 (28.1 days at full health). The final sensitivity analysis included the derivation of QALYs after adjusting for EQ-5D differences at baseline and a set of baseline characteristics. Again, FT was found to be more costly and more effective than TAU, with an ICER of £16,533.03 per QALY, which was below the recommended NICE threshold of £20,000.

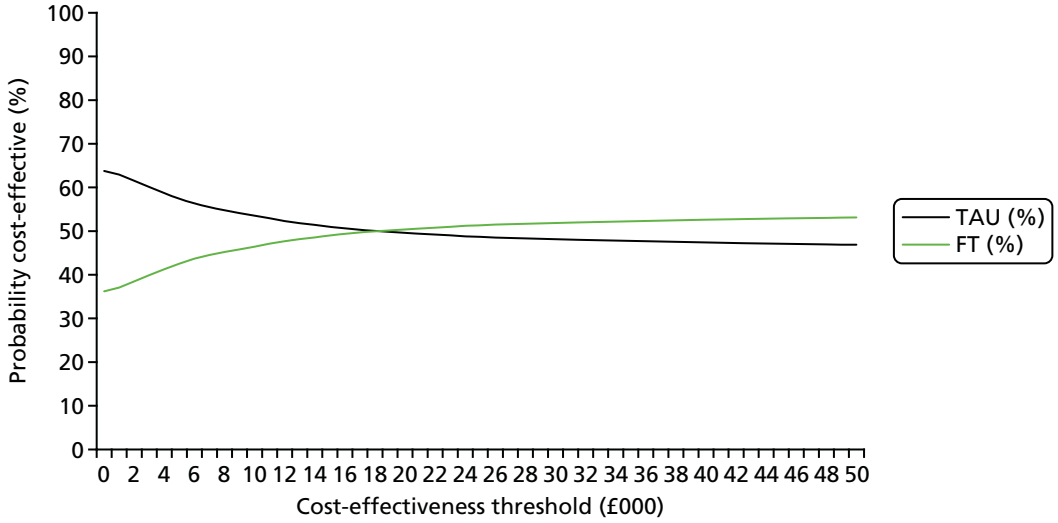


FIGURE 46 Long-term cost-effectiveness acceptability curve of FT compared with TAU (outcome measure: QALY, NHS perspective).

TABLE 101 Deterministic sensitivity analyses for 5-year model

FT vs. TAU	Incremental cost (£)	Incremental QALY	ICER (£/QALY)
Assume intervention costs from scenario 1	117.59	0.070	1673.21
Assume intervention costs from scenario 2	282.89	0.070	4025.21
Costs and utilities from trial data at 6 months and after 6 months	1041.55	0.077	13,458.40
Adjusting for baseline EQ-5D differences	1311.41	0.079	16,533.03

Chapter 5 Discussion

Summary of results

Primary outcome: self-harm

In this study, 221 (26.6%) participants inflicted self-harm leading to hospital attendance within 18 months of randomisation. Older participants (aged 15–17 years) had a lower rate of repeat self-harm than younger ones (aged 11–14 years). Participants recruited via referral direct from hospital to CAMHS had higher rates of repeat self-harm, as did females and participants whose index episode combined self-injury and poisoning as opposed to self-injury alone. The self-harm method used for the primary outcome event was commonly not the same as that used at the index event, with just over half of those whose index event involved self-injury 'switching' to self-poisoning.

In this study, in young people aged 11–17 years who had self-harmed and had at least one previous episode of self-harm, there was no evidence of a difference in subsequent repetition of self-harm between those receiving TAU and FT within 18 months.

Secondary outcomes: self-harm

There was no evidence of a difference in subsequent repetition of self-harm within 12 months between those receiving TAU and those receiving FT, nor in all further repeat episodes of self-harm leading to hospital attendance within 18 months.

Researchers rated young people's self-report of their index episode of self-harm. The behaviour of 313 (37.6%) participants was classified as a suicide attempt. Remaining episodes were largely classed as NSSI. It was considered that just under half the participants (412, 49.5%) exhibited some intent to die, and in over three-quarters there was a low medical risk of death. Nevertheless, medical treatment was sought or received for 659 (79.2%) participants.

Although loss to follow-up means we have data (or partial data) on self-reported self-harm at 18 months from only 478 participants, it is clear that in our sample there were many self-harm events that did not lead to hospital attendance (the primary outcome), with 349 out of 478 (73.0%) participants reporting self-harm. An unadjusted analysis of time to first self-reported self-harm event does not contradict the primary outcome analysis.

Secondary outcomes: health economics

Both trial arms showed an increase in the mean EQ-5D over the 18 months' trial follow-up. The largest differences in EQ-5D scores between the two arms were at 6 and 12 months, with participants in the FT arm exhibiting significantly higher scores than those in the TAU arm, whereas there were no significant differences in quality of life between the two study arms at 18 months.

A cost-effectiveness analysis indicated that FT was £1266.23 more expensive and slightly more effective in terms of QALY gains (equivalent to 12.4 extra days of perfect health) than TAU, leading to an ICER of £36,811.80 per QALY gained. Hence, FT was found unlikely to be cost-effective at 18 months compared with TAU.

Secondary outcomes: other clinical outcomes

We also found no significant treatment differences for young person questionnaire outcomes on the CDRS-R, PQ-LES-Q, Hopelessness scale or McMaster FAD. However, young people in the FT arm reported significantly better outcomes on the prosocial scale of the SDQ at 12 and 18 months and on the impact of their problems scale at 12 months, but not at 18 months. They also reported significantly lower rates of suicidal ideation on the BSS at 12 months, but not at 18 months.

There were no significant differences between treatment groups in participants' caregiver questionnaire outcomes on the GHQ-12 or Family Questionnaire. However, caregivers in the FT arm reported a range of significantly better outcomes than those in TAU arm: on the total difficulties score of the SDQ; on the emotional problems, peer problems and internalising scales of the SDQ at 12 and 18 months; on the conduct problems and externalising scales of the SDQ at 18 months, but not at 12 months; and on the impact scale of the SDQ at 12 months, but not at 18 months. Caregivers also reported significantly better outcomes in FT on the roles scale of the McMaster FAD at 12 months but not at 18 months.

The number of participants with other 'administrative' outcomes, such as referrals to other services, including to inpatient units, and safety outcomes including re-referrals to CAMHS, A&E attendances and hospital admissions for any reason, were similar across the arms.

Moderator analyses

The risk of self-harm among young people whose scores on the unemotional subscale at baseline suggested that they had difficulty in talking about feelings was higher in the FT arm than in the TAU arm, but among those whose scores indicated that they found talking about feelings easier was lower in the FT arm than in the TAU arm.

Among young people whose caregivers reported healthier affective involvement scores (degree to which family members are involved and interested in one another) on the McMaster FAD, risk of self-harm was higher in the FT am than in the TAU arm, while among those with poorer affective involvement scores risk of self-harm was lower in the FT arm than in the TAU arm.

Additional health economic analyses

In a secondary analysis in which the number of self-harm events avoided over 18 months was the outcome measure, FT was found to be less effective than TAU, with 0.033 more self-harm events on average, and more expensive by £1252.57, indicating that TAU dominated FT (i.e. was more effective and less costly) in the management of repeated self-harm in adolescents.

When combining young people's and caregivers' QALY gains, FT was associated with higher costs and better health outcomes than TAU, with an ICER of £20,808.21 per QALY gain (within the NICE cost-effectiveness range of £20,000–30,000 per QALY) and with a probability of being cost-effective of 41% at a threshold of £20,000 per QALY and of 64% at a threshold of £30,000 per QALY. Hence, when health benefits were considered beyond the young person's own health benefits, FT would be considered a cost-effective use of NHS money.

Mediator analysis

Complier average causal effect analyses were conducted to model the causal effect of FT receipt (as opposed to randomisation) on the primary outcome, finding a very similar effect as in the primary analysis. Exploratory analyses of the time to first event for participants who were allocated to and received FT, and those allocated to TAU who received a form of FT as part of their usual care, show that the curves appear to diverge, with an increased rate of self-harm in those allocated to FT, but no statistically significant difference was found.

There was no evidence that any of the variables investigated formally mediated the effect of treatment on the time to self-harm, largely because of lack of evidence of a treatment effect. However, further analysis using alternative approaches to explore mediation in the absence of a treatment effect could be undertaken for a number of potential mediators identified (therapist experience, young person SDQ impact, caregiver SDQ emotional problems and caregiver McMaster FAD roles subscales).

Interpretation of results

The nature of the sample and generalisability

The proportion of females in the sample (88.3%) is not dissimilar to that seen in hospital samples[13,104] and is also similar to that seen in a well-conducted community survey of 15- to 16-year-olds in England,[5] and highly similar to a number of randomised controlled trials among this population[47,105–109] in the UK and in countries with similar health-care arrangements, including Australia.[110,111]

Baseline data on the sample as a whole suggest that the recruited group has experienced significant difficulties and significant mental disorders. A total of 26.2% reported a health or disability problem, 29.3% had been involved with CAMHS in the past, 21.4% reported marked physical abuse, 16.6% reported sexual abuse and 16.0% reported experience of being bullied. On the total difficulties score of the SDQ, 66.2% of participants scored in the high/very high range, with their caregivers reporting that 69.6% scored in this range. On the general functioning subscale of the FAD, 84.7% of participants scored their families as 'unhealthy', with the equivalent figure from caregivers being 75.8%. On the CDRS-R, 65.7% of participants scored themselves as being in the moderate, severely or very severely depressed category.

To be eligible for the trial, young people needed to have self-harmed at least twice; in the event, most (88.8%) had self-harmed at least three times prior to randomisation. The index events leading to enrolment in the trial was cutting in 63.3% of participants (71.4% if other forms of self-injury are included, e.g. biting and burning), self-poisoning in 22.7% and a combination of cutting and self-poisoning in 6.5%. Ratings on the SASII suggested that 62.0% of index events would be classified as NSSI, but 49.5% of participants were rated as having some intent to die. This suggests that the decision to use UK definitions of self-harm 'regardless of motive' is justified because of the ambiguity about and difficulty of interpretation of statements about intent.

The self-harm method used by the young people in the sample is much more slanted towards self-injury than seen in hospital-referred cases – around 71.4% in the study, but only 17% in a monitoring study in three English cities of 5200 consecutive episodes of hospital attendance caused by self-harm in those aged under 18 years.[13] This is likely to be related to the fact that, although the primary outcome was self-harm leading to hospital attendance, the eligibility criteria for enrolment in the study included adolescents who had recently self-harmed and, as a consequence, been referred to CAMHS – irrespective of whether or not they had attended hospital. In fact, 63.5% of the sample were community referrals and were not referred directly by hospital services to CAMHS following self-harm. Some of the community referrals to SHIFT had been discharged from hospital following self-harm without a CAMHS referral and had then been referred to the trial via community services. Others had never presented to hospital in the first place but were referred from primary care following a self-harm episode. Interestingly, although the largest proportion of those screened were hospital referrals, many of these were ineligible, mostly because this was their first episode of self-harm. Although a smaller proportion of those screened were community referrals, more were eligible for the trial.

The method and location of recruitment are comparable with those of several more recent studies in the UK, using a mixture of hospital and CAMHS-presenting populations. Several older studies in the UK have focused on CAMHS participants[106,107,109] with more recent studies diversifying into A&E plus CAMHS[47] or A&E plus GP referrals.[105] A study in Australia recruited from both A&E and CAMHS.[111] Interestingly, none of these combined sample-source studies has noted any effect or diversity in response to the interventions by source of referral.

The primary outcome

Interpretation of any results related to self-harm is complicated by debates about terminology – in particular, differences in usage between clinicians and researchers in the USA and those in the UK and Europe. In this trial, which was a pragmatic trial taking place in a UK setting, self-harm was defined 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 stepping into traffic), regardless of motivation or the degree of intention to die. This was in line with the commissioning brief, but also with clinical practice and UK policy, where this definition is standard practice.[16] This definition includes what in the USA would be described as NSSI as well as suicidal behaviour.

In the USA, self-harm is often seen as synonymous with self-injury and implies a lack of suicidal intent, whereas in the UK the term encompasses self-injury and self-poisoning and is associated with concerns about making judgements about suicidal intent from the methods used to self-harm (see, for example, Wilkinson *et al.*,[108] Kapur *et al.*,[112] Butler and Malone,[113] Klonsky *et al.*[114] and Edmondson *et al.*[115]). There is increasing evidence, supported by data from this trial, that a significant proportion of people who repeatedly harm themselves will change methods in subsequent episodes[116] and that 'non-suicidal self-injury' should be taken more seriously.[117]

Overall, 26.6% of participants attended hospital as a result of repeated self-harm within 18 months of randomisation. Repetition of self-harm among the young people in the English multicentre monitoring study was around 28% in an average of more than 4 years of follow-up.[13] Repetition of around the same frequency in just 18 months in the present trial is not surprising, because participants were all young people who had already self-harmed on more than one occasion, while the monitoring study was a mix of first-time and repeat episodes: the monitoring study found a HR of 1.8 for repetition among those with a history of earlier episodes compared with those with no such history.[13] Similarly, the HR for repetition where the method was cutting or other injury was 1.5 and the present trial had a very high proportion of self-injury among the methods used at the most recent self-harm episode.

Age-related repetition in this study showed a pattern that does not reflect the picture seen in large-scale monitoring of hospital attendance because of self-harm. We found that that younger people in the trial (aged 11–14 years) were significantly more likely to show repeated self-harm than older adolescents (aged 15–17 years), but that was not the case in the three cities monitoring study, in which the 1-year repetition rate was exactly the same among those aged 10–14 years as those aged 15–18 years (18%).[16]

In the monitoring study,[13] over a median 6 years of follow-up after an index episode, there was a suicide (or probable suicide) rate of 1.0%. On the basis of that rate, from a sample with some comparability, and in one-quarter of that follow-up time, the present study might have expected something in the region of two suicides; plainly zero suicides during the SHIFT trial follow-up period is therefore unsurprising.

Recent reviews have set out mixed findings in relation to interventions to prevent subsequent self-harm,[48,49] noting that the studies reviewed have relatively small sample sizes, with an average follow-up of 6 months, with most not powered for ascertaining risk of repetition, which is statistically a relatively rare event. This trial therefore adds to the existing literature that suggests that there is no clear evidence of any one intervention being more effective than others in reducing repeat self-harm in adolescents. It is, however, a significantly more robust study than others. It should be noted that the primary outcome in this study was self-harm leading to hospital attendance. Many of the other studies reviewed by Ougrin *et al.*[49] and Brent *et al.*[48] are based on self-reports of self-harm, again making comparisons with the existing literature problematic.

The pragmatic design of this study conferred a potential advantage on TAU as experienced CAMHS clinicians in the TAU arm had greater flexibility in selecting treatments (or combinations of treatments) based on the results of their clinical assessment and discussions with the young person and family about their preferred treatment strategies.

Secondary outcomes

There were marked improvements on nearly all measures from baseline to 18 months, but only for those summarised below were there significant between-group differences. The absence of a between-group difference on scores for depression and hopelessness is somewhat unexpected, given that depression is

most frequently identified as the strongest psychiatric risk factor associated with a range of self-harming behaviours.[14,21] It has been found in previous studies of both depression itself (see, for example, Hazell *et al.*[118] and Pineda and Dadds[119]) and suicidal behaviour to have an impact on responsiveness to treatment[120] or to frequently respond to treatments for adolescent self-harm.[46,121] Similarly, hopelessness has been identified in a number of epidemiological and clinical population studies as being associated with risk of repetition[108] and having an impact on treatment responsiveness.[118,122]

There was good evidence of reduced odds of suicidal ideation in the FT arm at 12 months but not at 18 months. However, this was because suicidal ideation had decreased further in the TAU group between 12 and 18 months, while suicidal ideation decreased to a lesser extent in the FT arm. It might be argued that if FT reduces suicidal ideation sooner, this is a potentially important clinical benefit. This is consistent with research using attachment-based FT,[123] a slightly different form of FT to that used here, but also addressing specific issues related to self-harm. The similarity of the findings strengthens our conclusions.

The significantly better outcomes on different elements of the SDQ (a widely used measure of general emotional and behavioural difficulties) for the FT arm reported by young people and their caregivers suggest that FT did have a significant, positive impact on general mental health, even if this did not translate into reduced repetition of self-harm. Caregivers reported a wide range of benefits across the total difficulties score and the subscales of the SDQ.

It has long been known that self-report questionnaire findings from caregivers and their children will differ. In the field of self-harm, in Pineda and Dadds' 2013[111] evaluation of an interactive psychoeducation programme, parents reported greater improvements in family functioning (FAD general scale) and adolescents reported lower levels of suicidal behaviour (including ideation, plans, self-harm or attempts) in the intervention group. Similarly, in a much cited study by Huey *et al.*,[124] young people, but not parents, reported lower repetition in the experimental group.

Moderator analyses

It might seem to make intuitive sense that families who report poorer family functioning in relation to affective involvement obtain benefits from an intervention that focuses specifically on family function. However, other studies have suggested that it is families with good functioning that benefit most from FT.[125] The unemotional subscale of the ICU[71] focuses on an absence of emotional expression with items related to the ability to express feelings openly and to show feeling to others. That adolescents who find this difficult may do less well in an intervention like FT, where such expression is encouraged, is more in line with the literature, but the two findings together suggest a more complex relationship between adolescent and family functioning and the ability to talk about feelings than has hitherto been reported. From a clinical perspective it may be that, where adolescents themselves have difficulty in expressing feelings after self-harm, a different approach from FT or a modification of FT practice may be indicated. In eating disorders, for example, it has been shown that families with high levels of criticism tend to drop out of FT and do worse in conjoint FT than when parents and adolescents are seen in parallel.[126,127]

Summary statistics appear to suggest a differential treatment effect according to source of referral, with an increased rate of self-harm in participants referred through hospital in FT. However formal testing for moderation did not detect a statistically significant difference, but it is recognised that interaction testing for moderation is not particularly powerful and the possibility that referral source might have an influence should be explored further in future research.

Health economics

The SHIFT approach required family therapists to work in teams and, therefore, more therapists were involved with each participant. It is interesting that, despite this, FT did not cost substantially more than TAU.

As the difference in QALY gains between FT and TAU was marginal at 18 months and did not exist in the complete-case analyses, the cost-effectiveness within-trial results would not have required a decision-analysis model. However, as the model was planned in the original protocol, the within-trial analysis was completed with a decision-analysis model using a 5-year horizon. When extrapolating the analysis beyond 18 months, it was found that participants in the FT incurred arm higher costs (by £1262.13) and experienced a 5-year QALY gain of 0.065 compared with those in the TAU arm. The ICER was below the £20,000 recommended threshold, indicating that FT was likely to be cost-effective in the longer term, although the level of uncertainty around this estimate was high and the cost-effectiveness acceptability curve exhibited a probability to be cost-effective of 50%. All the sensitivity analyses undertaken confirmed that FT could be cost-effective over a 5-year follow-up but with a very high level of uncertainty. The decision-analysis model showed that the cost-effectiveness of FT as an intervention in the management of self-harm in young people is highly uncertain in the longer term.

The finding that FT could be considered cost-effective when considering benefits beyond just those for the participants relies on the strong assumption that QALYs can be aggregated across individuals as a simple aggregation; while this has been done in prior studies on child health,[128] this is not yet part of the NICE reference case. It is, however, consistent with other health economic research showing benefits to other family members.[129] Such considerations require a theoretical justification and a discussion on the interdependence between the utility function of the adolescent and the parent that has not been pursued in this report, as it was not part of the trial's objective. This sensitivity analysis result suggests that caregivers' QALYs affected the conclusions of the economic evaluation, probably because the incremental participant QALYs were small and the base-case ICER was near the threshold value.[103]

Study strengths

The SHIFT team believe this to be among the largest (by participant number) studies in the world evaluating a psychological intervention in CAMHS. Many other studies in adolescent self-harm have much smaller sample sizes. The combined total of randomised participants in the 19 studies in Ougrin *et al.*'s[49] systematic review of self-harm interventions was 2432, with the smallest study randomising just 39 participants and the largest 448. By selecting a primary outcome that was measurable without contact with participants it was possible to ascertain the primary outcome in a very high proportion of participants, 795 (95.6%) and partially (< 18 months) for a further 33 (4.0%). This is thus an adequately powered study with a very high success rate in obtaining the primary outcome and an ITT analysis suggesting that the findings are robust. The follow-up period of 18 months is longer than many self-harm intervention studies, which typically have a 6-month follow-up.[48]

The piloting and subsequent use of routinely collected HES data obtained from NHS Digital (supplemented by researcher visits when clarification was needed for the reason for attendance or method of self-harm) was novel and effective and should be considered in other trials.

The presence of a full cost-effectiveness analysis is also unusual in trials aimed at reducing repeat self-harm and is a further strength.

This trial recruited 'high-risk' participants, as young people had to have self-harmed at least twice to be eligible; in fact, nearly all had self-harmed at least three times. This complicates comparison with other studies, as the number of previous attempts is not always reported consistently. The exclusion criteria largely related to conditions that would necessitate alternative, more specialist treatments – for example, anyone presenting with self-harm and an eating disorder would have been excluded and referred to a specialist eating disorder clinic. This, combined with sample baseline descriptive data described above, means that the sample is broadly typical of young people presenting to CAMHS and the findings are generalisable to those services. The small number of participants in foster care means that caution should be exercised in generalising to that group of young people. Although the sample comprised a mix of referrals to CAMHS

direct from hospital and from community sources, this is also typical of CAMHS referral, albeit different from the pattern often seen in adult services. TAU was also broadly comparable to CAMHS practice in the UK with a mixture of supportive/individual counselling, CBT, family work and more formal FT (although see *Study weaknesses* for commentary on this latter intervention).

A further strength of the SHIFT study is that there is detailed self-report information regarding the range of self-harm behaviours an adolescent may engage in, with these data available in a small number of recent randomised controlled trials but not generally for older studies in this area.[48]

This was also an evaluation of FT as it is practised in the UK. Therapists were relatively senior and had formal qualifications. They worked in teams and had regular supervision. Analysis of session recordings suggests that they did adhere to the manualised version of therapy designed for the trial.

Study weaknesses

There was significant loss to follow-up and therefore there are many missing data for secondary participant-reported outcomes and for the health economic analyses. Participants lost to follow-up tended to have particularly concerning baseline characteristics, the proportion of participants were lost to follow-up was higher in the TAU arm than in the FT arm and there was some indication of differential characteristics for those lost to follow-up across treatment allocation. Clinically, young people who self-harm do not readily engage with treatment and our choice of a primary outcome that could be obtained without contact with the young person was for this reason. Nevertheless, despite the relatively large numbers involved, interpretation of secondary outcome data has to be made with some caution because of the number of missing data.

The primary outcome used also has a built-in disadvantage in that it misses those who do not attend hospital following a self-harm episode. If the two treatments differ in changing the level of awareness of self-harm in the family and the willingness to act on this, then there is potential confounding. The large number of missing self-report data makes it difficult to test this hypothesis.

It is known that many incidents of self-harm never lead to hospital attendance,[5] and, indeed, in this study, nearly three-quarters of participants self-reported further episodes of self-harm within 18 months, in addition to those leading to hospital attendance, with only one in seven episodes resulting in medical attention being sought. The extent of missing self-report data in this sample at the 12- and 18-month follow-up means that we are unable to comment on between-arm differences in self-reported self-harm.

When recruiting centres for the trial, there was the expectation that they would be compliant with UK guidance on management of self-harm and not be already routinely offering formal FT to those who had self-harmed. It was therefore surprising that so many participants in the TAU arm received formal FT (between 10% and 20%, although sometimes only for one session). Strictly speaking, this was not a protocol violation. This was a pragmatic trial and TAU was not restricted in any way. As described in the methods section, clear steps were taken to avoid those delivering FT within TAU becoming aware of the specific manualised form of FT that was being delivered in the FT arm. There had always been the expectation that participants in the TAU arm would receive some form of 'family work', as in the UK occasional family meetings are a routine part of practice, but we did not expect so many TAU participants to receive FT. As noted earlier, analysis exploring the primary outcome for those in TAU who received a version of FT suggest that this group did not do significantly better than those in the FT arm or those in the TAU group who received other interventions.

In addition, the high degree of variability in TAU could also have also impacted on and limited interpretation of study findings. If there was a lot of geographical variation in standard care, then it is difficult to evaluate

what FT was ultimately being compared against. It remains possible that it may have been more effective (and cost-effective) depending on the actual alternative package of care in specific cases.

We used the Baron and Kenny approach[90] to provide a preliminary analysis exploring the mediating effect of process variables and self-reported participant outcomes on the primary outcome; however, because of the lack of an overall treatment effect, interpretation of these results is limited. Further methods exist in order to explore mediation in the absence of a treatment effect and also hidden confounding (selection effects), such as structural equation modelling, instrumental variable, principle stratification[130] and counterfactual methods.[131] However, such current methodology largely focuses on linear outcomes and mediators, with limited guidance for survival outcomes[130] in particular, although, Valeri and VanderWeele[132] have recently extended their work in the counterfactual framework to account for survival data. Further exploration of potential mediators currently identified could therefore be beneficial, if methods allow, to further understand the impact of therapist experience across arms and self-reported outcomes on the young person SDQ impact, caregiver SDQ emotional problems and caregiver McMaster FAD roles subscales.

Dissemination and patient and public involvement

In addition to presentations at appropriate academic and clinical conferences and academic publications, the research team convened a SHIFT conference in Leeds in the autumn of 2016 in order to share the results of the trial with all of the clinical teams that participated and other interested CAMHS clinicians. Results are posted on the SHIFT website (URL: www.medhealth.leeds.ac.uk/info/615/research/315/shift).

The research team have had fruitful consultation with the NIHR-funded Young People's Mental Health Advisory Group in order to plan how best to present the findings of the research and will continue to liaise with this group.

In order to ensure that the research findings are available to service users and carers, the research team will publicise trial findings with the help of organisations such as Young Minds and will also liaise with the clinical teams that took part in the trial to ensure that findings are disseminated to their local PPI groups.

Chapter 6 Conclusions

Clinical implications

- This trial did not demonstrate that SHIFT manualised FT following repeated self-harm reduced subsequent hospital attendances for self-harm compared with TAU.
- The trial does not demonstrate that brief psychological interventions generally, or some form of family intervention (received by nearly one-third of the TAU group), are ineffective. It demonstrates that this particular version of FT did not seem to confer extra benefits in terms of reducing the risk of further self-harm in an unselected group of adolescents referred after repeat self-harm when compared with TAU.
- An important caveat is that this was a higher-risk group and therefore conclusions cannot be drawn about those presenting for the first time following self-harm.
- It should also be noted that in the TAU arm clinicians had considerable flexibility in tailoring treatments offered to individual needs.
- The high proportion of participants whose index episode of self-harm involved self-injury and the fact that two-thirds of the sample was not recruited directly following admission to hospital mean that, although the sample is representative of all self-harm referrals to CAMHS, the findings may not be generalisable to the smaller subset of adolescents who present to hospital following a first episode of self-harm. Generalisation to those in foster care may also be inappropriate given the small number of such participants in the sample.
- There is good evidence that FT has a positive impact on general emotional and behavioural problems at 12 and 18 months and may reduce suicidal ideation faster than TAU.
- There is some evidence to support the increased effectiveness of FT compared with TAU in reducing self-harm where caregivers report poor family functioning or young people report ease in discussing emotions.
- Conversely, when young people reported difficulty in expressing emotion or families reported healthy functioning, other interventions or modifications of FT may be indicated.
- Although there was no evidence of the cost-effectiveness of FT in the primary analyses focused on health benefits to young people, there is a suggestion that FT may be cost-effective if health benefits to the caregivers are additionally taken into account.
- This trial adds further to the evidence that young people are likely to switch methods in subsequent episodes of self-harm and that the method of self-harm itself may not be a useful indicator of suicidal intent or risk.
- There are indications that participants randomised to FT exhibited higher levels of engagement, with fewer participants who 'did not attend at all' and less 'drop-out with no negotiation' and higher rates of participants 'completing all required sessions'. This aligns with other trials of FT in which higher levels of engagement have been noted in the management of self-harm and in evaluation of FT for other conditions.[21,133]

Research implications

Young people who self-harm form a varied and heterogeneous group. Self-harm may be the final common pathway for a wide range of interpersonal and mental health predicaments. This was a commissioned call that specified a family intervention, and that is what we evaluated. It is possible that interventions based on other theoretical models may be more effective, although the lack of evidence of effectiveness of a wide range of approaches from earlier studies suggests that this may not be the case. Future research needs to evaluate interventions targeted at the characteristics of specific subgroups within the self-harming population. It may also be the case that family-oriented interventions are effective for only some of these

subgroups. Obvious candidate groups arising from this research would be families that self-report poorer family functioning and young people who are more unemotional and those referred to CAMHS directly from hospital rather than the community. The match in perceptions of family functioning between young people and their caregivers may also be an area worth exploring.

Further research into the characteristics of these two groups is also indicated: what is the exact nature of the family dysfunction that some groups report and how might psychological interventions be targeted at this group? Are unemotional traits shared by other family members and is it possible that these two findings are aspects of the same underlying issue?

Collecting follow-up data from participants was clearly problematic. Collecting data about therapeutic alliance also proved difficult in the TAU arm and yet is of potential benefit in understanding complex treatment effects. Further consideration needs to be given as to how this might be achieved in future studies; exploring the potential benefits of routine data and/or new technologies in enhancing such data collection would be valuable.

At 18 months, this trial already has a longer follow-up period than many other published reports, but 18 months is still a relatively short time, given that self-harming behaviour continues into adult life. Longer follow-up studies are needed. For SHIFT, the NIHR Health Technology Assessment (HTA) programme has agreed funding to extend follow-up by 18 months. This means that the research team will be able to use routinely collected data to obtain the primary outcome (and hospital attendances for other reasons) for all consenting participants at 3 years and for some at 6–7 years post randomisation, and will also collect adult mental health records of participants who are now aged > 18 years. Given the availability of routine data, follow-up could and should continue beyond this and ideally should explore, with appropriate consent, linkage to other relevant national databases.

The possibility that FT may have benefits beyond that of the identified participant requires further exploration as to how health economic benefits might be aggregated for family members.

The significant differences observed in self-reported episodes of self-harm and episodes requiring hospital attendance and the very different patterns of self-harm recorded suggest that further work is needed to clarify the most appropriate outcome measures in self-harm research and how they might best be measured.

Acknowledgements

Trial Steering Committee

Professor Simon Gowers (Independent Chairperson), Dr Sarah Byford, Ms Barbara Riddell (PPI representative), Professor Chris Roberts and Professor Arlene Vetere.

Data Monitoring and Ethics Committee

Professor Simon Gilbody (Independent Chairperson), Professor Mike Campbell and Professor Alan Carr.

Leeds Institute for Clinical Trials Research

Ms Emma Batman, Mr Cris Bunker, Mrs Kayleigh Burton, Mrs Claire Caffrey, Mr Andrew Carter, Mr Andrew Davies, Ms Alison Fergusson, Ms Marie Fletcher, Ms Andrea Franklin, Ms Madeline Harms, Mrs Amanda Lilley-Kelly, Ms Vicki McLellan, Professor Jane Nixon (co-applicant), Mrs Caroline Paley, Mr Peter Saunders, Ms Laura Stubbs and Mr John Turgoose.

Academic Unit of Health Economics, Leeds Institute of Health Sciences

Professor Chris McCabe (co-applicant), Dr Yemi Oluboyede, Dr Sophie Guthmuller and Dr Andrew Sutton.

Leeds Institute of Health Sciences

Self-Harm Intervention: Family Therapy trial researchers: Ms Fiona Lambert, Ms Kirsty Blay and Ms Cara Gates.

University of Manchester

Self-Harm Intervention: Family Therapy trial researchers: Dr Justine Rothwell, Dr Clare Harrop, Dr Catherine Kay, Ms Amy Burns, Ms Ami Brooks and Ms Hannah Tobin.

King's College London

Self-Harm Intervention: Family Therapy trial researchers: Ms Jude Gittins, Ms Claire Kambamettu, Ms Lucy Brett-Taylor, Ms Kuljit Heer, Ms Sarah Fife and Ms Elise Campbell.

The Tavistock and Portman NHS Foundation Trust

Dr Reenee Singh and Dr Charlotte Burck.

Patient and public involvement

With thanks to Young Minds (the UK's leading charity committed to improving the emotional well-being and mental health of children and young people) and the NIHR's Young People's Mental Health Advisory Group for their advice and help with the review of participant materials.

Participating sites

With thanks to all clinicians at participating CAMHS (within the NHS trusts detailed below) who identified and treated trial participants and provided trial data:

Leeds Community Healthcare NHS Trust (three CAMHS), Bradford District Care Trust (two CAMHS), South West Yorkshire Partnership NHS Foundation Trust (five CAMHS), Leeds and York Partnership NHS Foundation Trust (one CAMHS), Humber NHS Foundation Trust (three CAMHS), Central Manchester and Manchester Children's University Hospitals Trust (four CAMHS), Trafford Healthcare NHS Trust (one CAMHS), Pennine Care NHS Foundation Trust (five CAMHS), Five Boroughs Partnership NHS Foundation Trust (four CAMHS), The Tavistock and Portman NHS Foundation Trust (one CAMHS), Whittington Health NHS (one CAMHS), Barnet, Enfield and Haringey Mental Health NHS Trust (two CAMHS), South London and Maudsley NHS Foundation Trust (four CAMHS), Oxleas NHS Foundation Trust (three CAMHS) and East London NHS Foundation Trust (one CAMHS).

Family therapists

Thank you to all the dedicated SHIFT trial family therapists and local supervisors:

Annette Allen, Sarah Amoss, Sarah Baker, James Barclay, Roger Barford, Esther Blessitt, Mary Bratley, Charlotte Burck, Steve Cansdale, Iris Corrales, Claire Dempster, Annie Fatima, Sarah Favier, Mike Gibson, Cynthia Hamlin, Lumi Henshaw, Carole Hunt, Simon Jubb, Monica Lynch, Eva Murphy, Kerri Newns, Ann Overton, Judy Rathbone, Martin Richards, Alice Rogers, Lydia Rolley, Petra Schmidt, Paul Stockwell, Annette Wetherell, Jane Wilson and Kath Wissett.

National Institute for Health Research Clinical Research Networks

With thanks to the CSOs from the local NIHR Clinical Research Networks in Yorkshire, London and Greater Manchester for their help with recruitment and with the collection of clinical data from participating CAMHS.

Participants

Thank you to all trial participants for their essential contribution to this trial.

Contributions of authors

The SHIFT trial was conceived by the trial team in response to a call by the NIHR HTA programme for a study investigating the clinical and cost-effectiveness of FT for adolescents who self-harm (HTA no. 07/33). It was designed by Professor **David J Cottrell**, together with Ms **Paula Boston** (Programme Leader for Family Therapy training, University of Leeds), Professor **Ivan Eisler** (Professor of Family Psychology and Family Therapy), Dr **Sarah Fortune** (Clinical Psychologist), Professor **Jonathan Green** (Professor of Child and Adolescent Psychiatry), Professor **Allan O House** (Professor of Liaison Psychiatry), Professor

Michael Kerfoot (Professor of Child and Adolescent Policy and Research), Dr **David W Owens** (Associate Professor in Psychiatry), Dr **Mima Simic** (Consultant Child and Adolescent Psychiatrist) and Professor **Amanda J Farrin** (Professor of Clinical Trials and Evaluation of Complex Interventions and Director of the Complex Interventions Division at Leeds Institute for Clinical Trials Research). Professor David J Cottrell, Professor Ivan Eisler and Professor Jonathan Green were responsible for researchers undertaking recruitment at CAMHS within the Yorkshire, London and Manchester hubs, respectively. Ms Paula Boston and Professor Ivan Eisler trained the trial family therapists and ensured supervision.

Ms **Alex Wright-Hughes** and Mrs **Michelle Collinson** (Senior Medical Statisticians) provided statistical input into the trial design, implementation, statistical analysis plan and undertook the statistical analysis under the supervision of Professor Amanda J Farrin.

Ms **Elizabeth H Graham** (Trial Manager) was delivery lead for the implementation of the trial, acquisition of trial data, trial monitoring, good clinical practice and reporting requirements.

Dr **Eirini-Christina Saloniki** (Research Fellow in Health Economics) performed the analysis.

Dr **Sandy Tubeuf** (Associate Professor in Health Economics) designed the health economics analysis and had oversight of the economic evaluation.

All authors (excepting Professor Michael Kerfoot, deceased) contributed to the writing of the report and had the opportunity to revise prior to submission.

Publications

Oluboyede Y, Tubeuf S, McCabe C. Measuring health outcomes of adolescents: report from a pilot study. *Eur J Health Econ* 2013;**14**:11–19.

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**:501.

Amoss S, Lynch M, Bratley M. Bringing forth stories of blame and shame in dialogues with families affected by adolescent self-harm. *J Fam Ther* 2016;**38**:189–205.

Boston P, Cottrell D. Trials and Tribulations – an RCT comparing manualized family therapy with treatment as usual and reflections on key issues that arose in the implementation. *J Fam Ther* 2016;**38**:172–88.

Fortune S, Cottrell D, Fife S. Family factors associated with adolescent self-harm: a narrative review. *J Fam Ther* 2016;**38**:226–56.

Masterson C, Barker C, Jackson D, Boston P. Constructing a SHIFT adherence measure (SAM): the development of a family therapy integrity measure for the SHIFT trial. *J Fam Ther* 2016;**38**:274–90.

Rogers A, Schmidt P. Emotion talk in the context of young people self-harming: facing the feelings in family therapy. *J Fam Ther* 2016;**38**:206–25.

Cottrell DJ, Wright-Hughes A, Collinson M, Boston P, Eisler I, Fortune S, *et al*. Effectiveness of systemic family therapy versus treatment as usual for young people after self-harm: a pragmatic, phase 3, multicentre, randomised controlled trial. *Lancet Psychiatry* 2018;**5**:203–16.

Wright-Hughes A, Graham E, Cottrell D, Farrin A. Routine hospital data – is it good enough for trials? An example using England's hospital episode statistics in the SHIFT trial of family therapy vs. treatment as usual in adolescents following self-harm. *Clin Trials* 2018; in press.

Data sharing statement

Data requests should be sent to the corresponding author and will be subject to review by a subgroup of the trial team. All data sharing activities would require a data sharing agreement.

References

1. Swannell SV, Martin GE, Page A, Hasking P, St John NJ. Prevalence of nonsuicidal self-injury in nonclinical samples: systematic review, meta-analysis and meta-regression. *Suicide Life Threat Behav* 2014;**44**:273–303. http://dx.doi.org/10.1111/sltb.12070
2. Hawton K, Saunders KE, O'Connor RC. Self-harm and suicide in adolescents. *Lancet* 2012;**379**:2373–82. http://dx.doi.org/10.1016/S0140-6736(12)60322-5
3. Evans E, Hawton K, Rodham K, Deeks J. The prevalence of suicidal phenomena in adolescents: a systematic review of population-based studies. *Suicide Life Threat Behav* 2005;**35**:239–50. https://doi.org/10.1521/suli.2005.35.3.239
4. Brunner R, Kaess M, Parzer P, Fischer G, Resch F, Carli V, *et al.* 3038–Characteristics of non-suicidal self-injury and suicide attempts among adolescents in Europe: results from the European Research Consortium SEYLE. *Eur Psychiatry* 2013;**28**:1. https://doi.org/10.1016/S0924-9338(13)77531-X
5. Hawton K, Rodham K, Evans E, Weatherall R. Deliberate self harm in adolescents: self report survey in schools in England. *BMJ* 2002;**325**:1207–11. https://doi.org/10.1136/bmj.325.7374.1207
6. Hawton K, Bergen H, Casey D, Simkin S, Palmer B, Cooper J, *et al.* Self-harm in England: a tale of three cities. Multicentre study of self-harm. *Soc Psychiatry Psychiatr Epidemiol* 2007;**42**:513–21. http://dx.doi.org/10.1007/s00127-007-0199-7
7. Hawton K, Hall S, Simkin S, Bale L, Bond A, Codd S, Stewart A. Deliberate self-harm in adolescents: a study of characteristics and trends in Oxford, 1990-2000. *J Child Psychol Psychiatry* 2003;**44**:1191–8. https://doi.org/10.1111/1469-7610.00200
8. Carter G, Reith DM, Whyte IM, McPherson M. Repeated self-poisoning: increasing severity of self-harm as a predictor of subsequent suicide. *Br J Psychiatry* 2005;**186**:253–7. https://doi.org/10.1192/bjp.186.3.253
9. Gibb SJ, Beautrais AL, Fergusson DM. Mortality and further suicidal behaviour after an index suicide attempt: a 10-year study. *Aust N Z J Psychiatry* 2005;**39**:95–100. https://doi.org/10.1080/j.1440-1614.2005.01514.x
10. Goldacre M, Hawton K. Repetition of self-poisoning and subsequent death in adolescents who take overdoses. *Br J Psychiatry* 1985;**146**:395–8. https://doi.org/10.1192/bjp.146.4.395
11. Suominen K, Isometsä E, Suokas J, Haukka J, Achte K, Lönnqvist J. Completed suicide after a suicide attempt: a 37-year follow-up study. *Am J Psychiatry* 2004;**161**:562–3. http://dx.doi.org/10.1176/appi.ajp.161.3.562
12. Hawton K, Harriss L. Deliberate self-harm in young people: characteristics and subsequent mortality in a 20-year cohort of patients presenting to hospital. *J Clin Psychiatry* 2007;**68**:1574–83. https://doi.org/10.4088/JCP.v68n1017
13. Hawton K, Bergen H, Kapur N, Cooper J, Steeg S, Ness J, *et al.* Repetition of self-harm and suicide following self-harm in children and adolescents: findings from the Multicentre Study of Self-harm in England. *J Child Psychol Psychiatry* 2012;**53**:1212–19. https://doi.org/10.1111/j.1469-7610.2012.02559.x
14. Bridge JA, Goldstein TR, Brent DA. Adolescent suicide and suicidal behavior. *J Child Psychol Psychiatry* 2006;**47**:372–94. https://doi.org/10.1111/j.1469-7610.2006.01615.x

15. Muehlenkamp JJ, Claes L, Havertape L, Plener PL. International prevalence of adolescent non-suicidal self-injury and deliberate self-harm. *Child Adolesc Psychiatry Ment Health* 2012;**6**:10. http://dx.doi.org/10.1186/1753-2000-6-10

16. Hawton K, Bergen H, Waters K, Ness J, Cooper J, Steeg S, *et al.* Epidemiology and nature of self-harm in children and adolescents: findings from the multicentre study of self-harm in England. *Eur Child Adolesc Psychiatry* 2012;**21**:369–77. https://doi.org/10.1007/s00787-012-0269-6

17. Hawton K, Rodham K, Evans E. *By their Own Young Hand: Deliberate Self-Harm and Suicidal Ideas in Adolescents*. London: Jessica Kingsley Publishers; 2006.

18. Reith DM, Whyte I, Carter G. Repetition risk for adolescent self-poisoning: a multiple event survival analysis. *Aust N Z J Psychiatry* 2003;**37**:212–18. https://doi.org/10.1046/j.1440-1614.2003.01114.x

19. Beautrais AL. Risk factors for suicide and attempted suicide among young people. *Aust N Z J Psychiatry* 2000;**34**:420–36. https://doi.org/10.1080/j.1440-1614.2000.00691.x

20. Evans E, Hawton K, Rodham K. Factors associated with suicidal phenomena in adolescents: a systematic review of population-based studies. *Clin Psychol Rev* 2004;**24**:957–79. https://doi.org/10.1016/j.cpr.2004.04.005

21. Ougrin D, Tranah T, Leigh E, Taylor L, Asarnow JR. Practitioner review: self-harm in adolescents. *J Child Psychol Psychiatry* 2012;**53**:337–50. http://dx.doi.org/10.1111/j.1469-7610.2012.02525.x

22. King CA, Merchant CR. Social and interpersonal factors relating to adolescent suicidality: a review of the literature. *Arch Suicide Res* 2008;**12**:181–96. http://dx.doi.org/10.1080/13811110802101203

23. Michelson D, Bhugra D. Family environment, expressed emotion and adolescent self-harm: a review of conceptual, empirical, cross-cultural and clinical perspectives. *Int Rev Psychiatry* 2012;**24**:106–14. https://doi.org/10.3109/09540261.2012.657613

24. Wagner BM, Silverman MAC, Martin CE. Family factors in youth suicidal behaviors. *Am Behav Sci* 2003;**46**:1171–91. https://doi.org/10.1177/0002764202250661

25. Fergusson DM, Woodward LJ, Horwood LJ. Risk factors and life processes associated with the onset of suicidal behaviour during adolescence and early adulthood. *Psychol Med* 2000;**30**:23–39. https://doi.org/10.1017/S003329179900135X

26. Klaus NM, Mobilio A, King CA. Parent–adolescent agreement concerning adolescents' suicidal thoughts and behaviors. *J Clin Child Adolesc Psychol* 2009;**38**:245–55. https://doi.org/10.1080/15374410802698412

27. Thompson R, Dubowitz H, English DJ, Nooner KB, Wike T, Bangdiwala SI, *et al.* Parents' and teachers' concordance with children's self-ratings of suicidality: findings from a high-risk sample. *Suicide Life Threat Behav* 2006;**36**:167–81. http://dx.doi.org/10.1521/suli.2006.36.2.167

28. Fortune S, Cottrell D, Fife S. Family factors associated with adolescent self-harm: a narrative review. *J Fam Ther* 2016;**38**:226–56. http://dx.doi.org/10.1111/1467-6427.12119

29. Brent DA, Baugher M, Bridge J, Chen T, Chiappetta L. Age- and sex-related risk factors for adolescent suicide. *J Am Acad Child Adolesc Psychiatry* 1999;**38**:1497–505. https://doi.org/10.1097/00004583-199912000-00010

30. Johnson BA, Brent DA, Bridge J, Connolly J. The familial aggregation of adolescent suicide attempts. *Acta Psychiatr Scand* 1998;**97**:18–24. https://doi.org/10.1111/j.1600-0447.1998.tb09957.x

31. Shaffer D, Gould MS, Fisher P, Trautman P, Moreau D, Kleinman M, Flory M. Psychiatric diagnosis in child and adolescent suicide. *Arch Gen Psychiatry* 1996;**53**:339–48. https://doi.org/10.1001/archpsyc.1996.01830040075012

32. Thompson EA, Mazza JJ, Herting JR, Randell BP, Eggert LL. The mediating roles of anxiety depression, and hopelessness on adolescent suicidal behaviors. *Suicide Life Threat Behav* 2005;**35**:14–34. http://dx.doi.org/10.1521/suli.35.1.14.59266

33. Hawton K, Kingsbury S, Steinhardt K, James A, Fagg J. Repetition of deliberate self-harm by adolescents: the role of psychological factors. *J Adolesc* 1999;**22**:369–78. http://dx.doi.org/10.1006/jado.1999.0228

34. Perloe A, Esposito-Smythers C, Curby TW, Renshaw KD. Concurrent trajectories of change in adolescent and maternal depressive symptoms in the TORDIA study. *J Youth Adolesc* 2014;**43**:612–28. http://dx.doi.org/10.1007/s10964-013-9999-0

35. Agerbo E, Nordentoft M, Mortensen PB. Familial, psychiatric, and socioeconomic risk factors for suicide in young people: nested case-control study. *BMJ* 2002;**325**:74. https://doi.org/10.1136/bmj.325.7355.74

36. Brent DA, Perper JA, Moritz G, Liotus L, Schweers J, Balach L, Roth C. Familial risk factors for adolescent suicide: a case–control study. *Acta Psychiatr Scand* 1994;**89**:52–8. https://doi.org/10.1111/j.1600-0447.1994.tb01485.x

37. Gould MS, Fisher P, Parides M, Flory M, Shaffer D. Psychosocial risk factors of child and adolescent completed suicide. *Arch Gen Psychiatry* 1996;**53**:1155–62. https://doi.org/10.1001/archpsyc.1996.01830120095016

38. Brent DA, Oquendo M, Birmaher B, Greenhill L, Kolko D, Stanley B, *et al.* Familial pathways to early-onset suicide attempt: risk for suicidal behavior in offspring of mood-disordered suicide attempters. *Arch Gen Psychiatry* 2002;**59**:801–7. https://doi.org/10.1001/archpsyc.59.9.801

39. Cox LJ, Stanley BH, Melhem NM, Oquendo MA, Birmaher B, Burke A, *et al.* A longitudinal study of nonsuicidal self-injury in offspring at high risk for mood disorder. *J Clin Psychiatry* 2012;**73**:821–8. http://dx.doi.org/10.4088/JCP.11m07250

40. Mann JJ, Bortinger J, Oquendo MA, Currier D, Li S, Brent DA. Family history of suicidal behavior and mood disorders in probands with mood disorders. *Am J Psychiatry* 2005;**162**:1672–9. https://doi.org/10.1176/appi.ajp.162.9.1672

41. Evans E, Hawton K, Rodham K. Suicidal phenomena and abuse in adolescents: a review of epidemiological studies. *Child Abuse Negl* 2005;**29**:45–58. https://doi.org/10.1016/j.chiabu.2004.06.014

42. van Heeringen K. Stress-Diathesis Model of Suicidal Behavior. In Dwivedi Y, editor. *The Neurobiological Basis of Suicide*. Abingdon: Taylor & Francis; 2012. pp. 113–24. https://doi.org/10.1201/b12215-7

43. Hawton K, Townsend E, Arensman E, Gunnell D, Hazell P, House A, van Heeringen K. Psychosocial versus pharmacological treatments for deliberate self harm. *Cochrane Database Syst Rev* 2000;**2**:CD001764.

44. Newton AS, Hamm MP, Bethell J, Rhodes AE, Bryan CJ, Tjosvold L, *et al.* Pediatric suicide-related presentations: a systematic review of mental health care in the emergency department. *Ann Emerg Med* 2010;**56**:649–59. https://doi.org/10.1016/j.annemergmed.2010.02.026

45. Robinson J, Hetrick SE, Martin C. Preventing suicide in young people: systematic review. *Aust N Z J Psychiatry* 2011;**45**:3–26. http://dx.doi.org/10.3109/00048674.2010.511147

46. Mehlum L, Tørmoen AJ, Ramberg M, Haga E, Diep LM, Laberg S, *et al.* Dialectical behavior therapy for adolescents with repeated suicidal and self-harming behavior: a randomized trial. *J Am Acad Child Adolesc Psychiatry* 2014;**53**:1082–91. http://dx.doi.org/10.1016/j.jaac.2014.07.003

47. Rossouw TI, Fonagy P. Mentalization-based treatment for self-harm in adolescents: a randomized controlled trial. *J Am Acad Child Adolesc Psychiatry* 2012;**51**:1304–13. https://doi.org/10.1016/j.jaac.2012.09.018

48. Brent DA, McMakin DL, Kennard BD, Goldstein TR, Mayes TL, Douaihy AB. Protecting adolescents from self-harm: a critical review of intervention studies. *J Am Acad Child Adolesc Psychiatry* 2013;**52**:1260–71. http://dx.doi.org/10.1016/j.jaac.2013.09.009

49. Ougrin D, Tranah T, Stahl D, Moran P, Asarnow JR. Therapeutic interventions for suicide attempts and self-harm in adolescents: systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry* 2015;**54**:97–107. https://doi.org/10.1016/j.jaac.2014.10.009

50. Glenn CR, Franklin JC, Nock MK. Evidence-based psychosocial treatments for self-injurious thoughts and behaviors in youth. *J Clin Child Adolesc Psychol* 2015;**44**:1–29. http://dx.doi.org/10.1080/15374416.2014.945211

51. Tompson MC, Boger KD, Asarnow JR. Enhancing the developmental appropriateness of treatment for depression in youth: integrating the family in treatment. *Child Adolescent Psychiatr Clin N Am* 2012;**21**:345–84. https://doi.org/10.1016/j.chc.2012.01.003

52. Linehan MM, Comtois KA, Brown MZ, Heard HL, Wagner A. Suicide Attempt Self-Injury Interview (SASII): development, reliability, and validity of a scale to assess suicide attempts and intentional self-injury. *Psychol Assess* 2006;**18**:303. https://doi.org/10.1037/1040-3590.18.3.303

53. Beck AT, Steer RA. *Manual for the Beck Scale for Suicide Ideation*. San Antonio, TX: Psychological Corporation; 1991.

54. Endicott J, Nee J, Yang R, Wohlberg C. Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q): reliability and validity. *J Am Acad Child Adolesc Psychiatry* 2006;**45**:401–7. http://dx.doi.org/10.1097/01.chi.0000198590.38325.81

55. Goldberg DP, Williams P. *A User's Guide to the General Health Questionnaire*. Windsor: NFER Nelson; 1988.

56. Poznanski EO, Mokros HB. *Children's Depression Rating Scale, Revised (CDRS-R)*. Los Angeles, CA: Western Psychological Services Los Angeles; 1996.

57. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry* 2001;**40**:1337–45. https://doi.org/10.1097/00004583-200111000-00015

58. Kazdin AE, Rodgers A, Colbus D. The hopelessness scale for children: psychometric characteristics and concurrent validity. *J Consult Clin Psychol* 1986;**54**:241–5. https://doi.org/10.1037/0022-006X.54.2.241

59. Epstein NB, Baldwin LM, Bishop DS. The McMaster Family Assessment Device. *J Marital and Fam Ther* 1983;**9**:171–80. https://doi.org/10.1111/j.1752-0606.1983.tb01497.x

60. 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 Res* 2002;**109**:265–79. https://doi.org/10.1016/S0165-1781(02)00023-9

61. Escudero V, Friedlander ML, Varela N, Abascal A. Observing the therapeutic alliance in family therapy: associations with participants' perceptions and therapeutic outcomes. *J Fam Ther* 2008;**30**:194–214. https://doi.org/10.1111/j.1467-6427.2008.00425.x

62. Pote H, Stratton P, Cottrell D, Shapiro D, Boston P. Systemic family therapy can be manualized: research process and findings. *J Fam Ther* 2003;**25**:236–62. https://doi.org/10.1111/1467-6427.00247

63. Dallos R. Attachment narrative therapy: integrating ideas from narrative and attachment theory in systemic family therapy with eating disorders. *J Fam Ther* 2004;**26**:40–65. https://doi.org/10.1111/j.1467-6427.2004.00266.x

64. White M. *Maps of Narrative Practice*. New York, NY: WW Norton & Company; 2007.

65. Cecchin G. Constructing therapeutic possibilities. In McNamee S, Gergen KJ, editors. *Therapy as Social Construction*. London: Sage; 1992. pp. 86–95.

66. Masterson C, Barker C, Jackson D, Boston P. Constructing a SHIFT adherence measure (SAM): the development of a family therapy integrity measure for the SHIFT trial. *J Fam Ther* 2016;**38**:274–90. https://doi.org/10.1111/1467-6427.12104

67. NICE. *Self-Harm in Over 8s: Short-Term Management and Prevention of Recurrence*. London: NICE; 2004.

68. NICE. *Depression in Children and Young People*. NICE: London; 2005.

69. 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**:501. http://dx.doi.org/10.1186/s13063-015-1007-4

70. Cottrell DJ, Wright-Hughes A, Collinson M, Boston P, Eisler I, Fortune S, *et al.*. Effectiveness of systemic family therapy versus treatment as usual for young people after self-harm: a pragmatic, phase 3, multicentre, randomised controlled trial. *Lancet Psychiatry* 2018;**5**:203–16.

71. Essau CA, Sasagawa S, Frick PJ. Callous-unemotional traits in a community sample of adolescents. *Assessment* 2006;**13**:454–69. https://doi.org/10.1177/1073191106287354

72. Miller IW, Epstein NB, Bishop DS, Keitner GI. The McMaster family assessment device: reliability and validity. *J Marital Fam Ther 1985*;**11**:345–56. https://doi.org/10.1111/j.1752-0606.1985.tb00028.x

73. Goldberg D. *General Health Questionnaire (GHQ-12)*. Windsor: Nfer-Nelson; 1992.

74. Brooks R, Rabin R, De Charro F. *The Measurement and Valuation of Health Status using EQ-5D: A European Perspective: Evidence from the Euroqol Biomed Research Programme*. The Netherlands: Springer Science & Business Media; 2013.

75. Oluboyede Y, Tubeuf S, McCabe C. Measuring health outcomes of adolescents: report from a pilot study. *Eur J Health Econ* 2013;**14**:11–19. http://dx.doi.org/10.1007/s10198-011-0340-0

76. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, *et al.* Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Med Care* 2002;**40**:113–28. https://doi.org/10.1097/00005650-200202000-00006

77. Ryan CE, Epstein NB, Keitner GI. *Evaluating and Treating Families: The McMaster Approach*. Abingdon: Taylor & Francis; 2005.

78. Fairclough DL. *Design and Analysis of Quality of Life Studies in Clinical Trials*. Boca Raton, FL: CRC Press; 2010.

79. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979;**35**:549–56. https://doi.org/10.2307/2530245

80. Lin DY, Wei L-J, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* 1993;**80**:557–72. https://doi.org/10.1093/biomet/80.3.557

81. Ripatti S, Palmgren J. Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics* 2000;**56**:1016–22. https://doi.org/10.1111/j.0006-341X.2000.01016.x

82. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. New York, NY: Springer Science & Business Media; 2000. https://doi.org/10.1007/978-1-4757-3294-8

83. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat* 1982;**10**:1100–20. https://doi.org/10.1214/aos/1176345976

84. Lin D, Wei L, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *J Royal Stat Soc B* 2000;**62**:711–30. https://doi.org/10.1111/1467-9868.00259

85. Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. *Biometrika* 1981;**68**:373–9. https://doi.org/10.1093/biomet/68.2.373

86. Little RJ, Rubin DB. Single Imputation Methods. *Statistical Analysis with Missing Data*. 2nd edn. Hoboken, NJ: Wiley; 2002. pp. 59–74.

87. Rubin DB. Inference and missing data. *Biometrika* 1976;**63**:581–92. https://doi.org/10.1093/biomet/63.3.581

88. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. *Clin Trials* 2012;**9**:396–407. http://dx.doi.org/10.1177/1740774512450098

89. Faries DE, Obenchain R, Haro JM, Leon AC. *Analysis of Observational Health Care Data Using SAS*. Cary, NC: SAS Institute; 2010.

90. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**:1173. https://doi.org/10.1037/0022-3514.51.6.1173

91. Byford S. The validity and responsiveness of the EQ-5D measure of health-related quality of life in an adolescent population with persistent major depression. *J Ment Health* 2013;**22**:101–10. http://dx.doi.org/10.3109/09638237.2013.779366

92. Group TE. EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:199–208. https://doi.org/10.1016/0168-8510(90)90421-9

93. NICE. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.

94. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. https://doi.org/10.1097/00005650-199711000-00002

95. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health Qual Life Outcomes* 2003;**1**:54. http://dx.doi.org/10.1186/1477-7525-1-54

96. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. *PharmacoEconomics* 1995;**7**:490–502. https://doi.org/10.2165/00019053-199507060-00004

97. Department of Health. *National Schedule of Reference Costs Year 2013–2014 – NHS Trusts PCT Combined.* London: Department of Health; 2014. URL: www.gov.uk/government/uploads/system/uploads/attachment_data/file/380322/01_Final_2013-14_Reference_Costs_publication_v2.pdf (accessed May 2016).

98. Curtis L. *Unit Costs of Health and Social Care 2014*. Canterbury: Personal Social Services Research Unit; 2014. URL: www.pssru.ac.uk/project-pages/unit-costs/2014 (accessed May 2016).

99. British Medical Association (BMA) and Royal Pharmaceutical Society (RPS) of Great Britain. *British National Formulary*. London: BMA and RPS; 2015.

100. 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. *J Am Stat Assoc* 1987;**82**:543–6. https://doi.org/10.2307/2289460

101. Schafer JL. *Analysis of Incomplete Multivariate Data*. Boca Raton, FL: CRC Press; 1997. https://doi.org/10.1201/9781439821862

102. Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, *et al.* Good research practices for cost effectiveness analysis alongside clinical trials: the ISPOR RCT CEA task force report. *Value Health* 2005;**8**:521–33. https://doi.org/10.1111/j.1524-4733.2005.00045.x

103. Al-Janabi H, Flynn TN, Coast J. QALYs and carers. *PharmacoEconomics* 2011;**29**:1015–23. http://dx.doi.org/10.2165/11593940-000000000-00000

104. Hawton K, Fagg J, Simkin S, Bale E, Bond A. Deliberate self-harm in adolescents in Oxford, 1985-1995. *J Adolesc* 2000;**23**:47–55. http://dx.doi.org/10.1006/jado.1999.0290

105. Ougrin D, Zundel T, Ng A, Banarsee R, Bottle A, Taylor E. Trial of Therapeutic Assessment in London: randomised controlled trial of therapeutic assessment versus standard psychosocial assessment in adolescents presenting with self-harm. *Arch Dis Child* 2011;**96**:148–53. http://dx.doi.org/10.1136/adc.2010.188755

106. Harrington R, Kerfoot M, Dyer E, McNiven F, Gill J, Harrington V, *et al.* Randomized trial of a home-based family intervention for children who have deliberately poisoned themselves. *J Am Acad Child Adolesc Psychiatry* 1998;**37**:512–8. https://doi.org/10.1016/S0890-8567(14)60001-0

107. Wood A, Trainor G, Rothwell J, Moore A, Harrington R. Randomized trial of group therapy for repeated deliberate self-harm in adolescents. *J Am Acad Child Adolesc Psychiatry* 2001;**40**:1246–53. https://doi.org/10.1097/00004583-200111000-00003

108. Wilkinson P, Kelvin R, Roberts C, Dubicka B, Goodyer I. Clinical and psychosocial predictors of suicide attempts and nonsuicidal self-injury in the Adolescent Depression Antidepressants and Psychotherapy Trial (ADAPT). *Am J Psychiatry* 2011;**168**:495–501. http://dx.doi.org/10.1176/appi.ajp.2010.10050718

109. Green JM, Wood AJ, Kerfoot MJ, Trainor G, Roberts C, Rothwell J, *et al.* Group therapy for adolescents with repeated self harm: randomised controlled trial with economic evaluation. *BMJ* 2011;**342**:d682. http://dx.doi.org/10.1136/bmj.d682

110. Hazell PL, Martin G, McGill K, Kay T, Wood A, Trainor G, *et al.* Group therapy for repeated deliberate self-harm in adolescents: failure of replication of a randomized trial. *J Am Acad Child Adolesc Psychiatry* 2009;**48**:662–70. https://doi.org/10.1097/chi.0b013e3181a0acec

111. Pineda J, Dadds MR. Family intervention for adolescents with suicidal behavior: a randomized controlled trial and mediation analysis. *J Am Acad Child Adolesc Psychiatry* 2013;**52**:851–62. https://doi.org/10.1016/j.jaac.2013.05.015

112. Kapur N, Cooper J, O'Connor RC, Hawton K. Non-suicidal self-injury v. attempted suicide: new diagnosis or false dichotomy? *Br J Psychiatry* 2013;**202**:326–8. http://dx.doi.org/10.1192/bjp.bp.112.116111

113. Butler AM, Malone K. Attempted suicide v. non-suicidal self-injury: behaviour, syndrome or diagnosis? *Br J Psychiatry* 2013;**202**:324–5. http://dx.doi.org/10.1192/bjp.bp.112.113506

114. Klonsky ED, Oltmanns TF, Turkheimer E. Deliberate self-harm in a nonclinical population: prevalence and psychological correlates. *Am J Psychiatry* 2003;**160**:1501–8. https://doi.org/10.1176/appi.ajp.160.8.1501

115. Edmondson AJ, Brennan CA, House AO. Non-suicidal reasons for self-harm: a systematic review of self-reported accounts. *J Affect Disord* 2016;**191**:109–17. http://dx.doi.org/10.1016/j.jad.2015.11.043

116. Owens D, Kelley R, Munyombwe T, Bergen H, Hawton K, Cooper J, *et al.* Switching methods of self-harm at repeat episodes: findings from a multicentre cohort study. *J Affect Disord* 2015;**180**:44–51. http://dx.doi.org/10.1016/j.jad.2015.03.051

117. Brent D. Nonsuicidal self-injury as a predictor of suicidal behavior in depressed adolescents. *Am J Psychiatry* 2011;**168**:452–4. http://dx.doi.org/10.1176/appi.ajp.2011.11020215

118. Asarnow JR, Emslie G, Clarke G, Wagner KD, Spirito A, Vitiello B, *et al.* Treatment of selective serotonin reuptake inhibitor; resistant depression in adolescents: predictors and moderators of treatment response. *J Am Acad Child Adolesc Psychiatry* 2009;**48**:330–9. http://dx.doi.org/10.1097/CHI.Ob013e3181977476

119. Curry J, Rohde P, Simons A, Silva S, Vitiello B, Kratochvil C, *et al.* Predictors and moderators of acute outcome in the Treatment for Adolescents with Depression Study (TADS). *J Am Acad Child Adolesc Psychiatry* 2006;**45**:1427–39. http://dx.doi.org/10.1097/01.chi.0000240838.78984.e2

120. Harrington R, Kerfoot M, Dyer E, Mcniven F, Gill J, Harrington V, Woodham A. Deliberate self-poisoning in adolescence: why does a brief family intervention work in some cases and not others? *J Adolesc* 2000;**23**:13–20. http://dx.doi.org/10.1006/jado.1999.0293

121. Asarnow JR, Berk M, Hughes JL, Anderson NL. The SAFETY Program: a treatment-development trial of a cognitive-behavioral family treatment for adolescent suicide attempters. *J Clin Child Adolesc Psychol* 2015;**44**:194–203. http://dx.doi.org/10.1080/15374416.2014.940624

122. Barbe RP, Bridge J, Birmaher B, Kolko D, Brent DA. Suicidality and its relationship to treatment outcome in depressed adolescents. *Suicide Life Threat Behav* 2004;**34**:44–55. https://doi.org/10.1521/suli.34.1.44.27768

123. Diamond GS, Wintersteen MB, Brown GK, Diamond GM, Gallop R, Shelef K, *et al.* Attachment-based family therapy for adolescents with suicidal ideation: a randomized controlled trial. *J Am Acad Child Adolesc Psychiatry* 2010;**49**:122–31.

124. Huey SJ, Henggeler SW, Rowland MD, Halliday-Boykins CA, Cunningham PN, Pickrel SG, *et al.* Multisystemic therapy effects on attempted suicide by youths presenting psychiatric emergencies. *J Am Acad Child Adolesc Psychiatry* 2004;**43**:183–90. https://doi.org/10.1097/00004583-200402000-00014

125. Hampson RB, Beavers WR. Measuring family therapy outcome in a clinical setting: families that do better or do worse in therapy. *Fam Process* 1996;**35**:347–61. https://doi.org/10.1111/j.1545-5300.1996.00347.x

126. Eisler I, Dare C, Hodes M, Russell G, Dodge E, Le Grange D. Family therapy for adolescent anorexia nervosa: the results of a controlled comparison of two family interventions. *J Child Psychol Psychiatry* 2000;**41**:727–36. https://doi.org/10.1111/1469-7610.00660

127. Eisler I, Simic M, Russell GF, Dare C. A randomised controlled treatment trial of two forms of family therapy in adolescent anorexia nervosa: a five-year follow-up. *J Child Psychol Psychiatry* 2007;**48**:552–60. https://doi.org/10.1111/j.1469-7610.2007.01726.x

128. Griebsch I, Coast J, Brown J. Quality-adjusted life-years lack quality in pediatric care: a critical review of published cost-utility studies in child health. *Pediatrics* 2005;**115**:e600–14. https://doi.org/10.1542/peds.2004-2127

129. Crane DR. Does family therapy reduce health care costs for more than the identified patient? *Clin Child Psychol Psychiatry* 2011;**16**:3–4. http://dx.doi.org/10.1177/1359104510397607

130. Dunn G, Emsley R, Liu H, Landau S, Green J, White I, Pickles A. Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health: a methodological research programme. *Health Technol Assess* 2015;**19**(93). http://dx.doi.org/10.3310/hta19930

131. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods* 2013;**18**:137. https://doi.org/10.1037/a0031034

132. Valeri L, VanderWeele TJ. SAS macro for causal mediation analysis with survival data. *Epidemiology* 2015;**26**:e23–4. http://dx.doi.org/10.1097/EDE.0000000000000253

133. Cottrell D, Boston P. Practitioner review: the effectiveness of systemic family therapy for children and adolescents. *J Child Psychol Psychiatry* 2002;**43**:573–86. https://doi.org/10.1111/1469-7610.00047

Appendix 1 Health Technology Assessment Programme Commissioning Brief (07/33)

NHS R&D Health Technology Assessment Programme *HTA no 07/33*

Family therapy for adolescents who self-harm

Introduction

The aim of the HTA programme is to ensure that high quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage, provide care in or develop policy for the NHS. Topics for research are identified and prioritised to meet the needs of the NHS. Health technology assessment forms the largest portfolio of work in the NHS Research and Development Programme and each year about fifty new studies are commissioned to help answer questions of direct importance to the NHS. The studies include both primary research and evidence synthesis.

Question

What is the clinical- and cost-effectiveness of family therapy for adolescents who self-harm?

1 **Technology:** Family therapy - researchers to define explicitly the elements of family therapy to be delivered.
2 **Patient group:** Adolescents (11-17 years) who have had more than one episode of deliberate self-harm (including self-poisoning and self-injury), excluding those who are severely depressed and suicidal. Researchers should define the selection of the patient group to exclude those for whom family therapy would be likely to be detrimental. Researchers should consider if there are meaningful subgroups.
3 **Setting:** Any – though patients will have been referred by a specialist service.
4 **Control or comparator treatment:** Conventional care.
5 **Design*:*** A randomised controlled trial.
6 **Primary outcomes:** Repetition of self-harm. Secondary outcomes: suicidal ideation, measures of quality of life for patient and family (researchers to define and justify scales used), adherence, cost, cost-effectiveness.
7 **Minimum duration of follow-up:** One year.

Background to commissioning brief:

Deliberate self-harm (DSH) among young people is a major public health issue in the UK, affecting at least one in 15 young people. As many as 30% of adolescents who self-harm report previous episodes, and a history of DSH is a significant risk factor for suicide. There continues to be insufficient evidence on which to make firm recommendations about the most effective forms of treatment for DSH patients. This is a serious situation given the size of the DSH population and the need to target this population as part of national suicide prevention strategies.

The limited studies to date have been relatively small and hence of limited power. Benefits of family therapy have been reported in one small RCT, but there is a need for robust information from a large trial evaluating family therapy.

Notes to Applicants

For many of the questions posed by the HTA programme, a randomised controlled trial is likely to be the most appropriate method of providing an answer. However, there may be practical or ethical reasons why this might not be possible. Applicants proposing other research methods are invited to justify these choices.

Applicants are asked to:

1. Follow the Medical Research Council's Good Clinical Practice guidelines (http://www.mrc.ac.uk/pdf-ctg.pdf) when planning how studies, particularly RCTs, will be supervised. Further advice specific to each topic will be given by the HTA programme at full proposal and contract stages.

2. Note that trials involving medicinal products must comply with "The Medicines for Human Use (Clinical Trials) Regulations 2004". In the case of such trials, the DH expects the employing institution of the chief investigator to be nominated as the sponsor. Other institutions may wish to take on this responsibility or agree co-sponsorship with the employing institution. The DH is prepared to accept the nomination of multiple sponsors. Applicants who are asked to submit a full proposal will need to obtain confirmation of a sponsor(s) to complete their application. The DH reserve the right to withdraw from funding the project if they are not satisfied with the arrangements put in place to conduct the trial.

The MHRA (info@mhra.gsi.gov.uk, http://www.mhra.gov.uk) can provide guidance as to whether your trial would be covered by the regulations. The DH/MRC website (http://www.ct-toolkit.ac.uk/) also contains the latest information about Clinical Trials regulations and a helpful FAQ page.

Making an application

If you wish to submit an outline proposal on this topic, complete the electronic application form and return it to the HTA Commissioning Manager at the National Coordinating Centre for Health Technology Assessment, Mailpoint 728 Boldrewood, University of Southampton, Southampton SO16 7PX by *Wednesday 25 April 2007.* Outline applications will be considered by the HTA Commissioning Board at its meeting in *July 2007.* If they are acceptable, investigators will be given a minimum of eight weeks to submit a full proposal.

Applications received after 1300 hours on the due date will not be considered.

Please see GUIDANCE ON APPLICATIONS overleaf.

The HTA programme expects, where appropriate, that applicants will work with the relevant research network.

Guidance on applications

Required expertise

HTA is a multidisciplinary enterprise. It needs to draw on the expertise and knowledge of clinicians and of those trained in health service research methodologies such as health economics, medical statistics, study design and qualitative approaches. HTA expects applicants to engage a qualified Trial Manager for appropriate projects. Applicants will need to show a commitment to team working and may wish to consider a collaborative approach between several institutions. It is expected that the research will be undertaken only following a thorough literature review.

Public involvement in research

The HTA programme recognises the increasing active involvement of members of the public in research and would like to support research projects appropriately. The HTA programme encourages applicants to consider *how* the scientific quality, feasibility or practicality of their proposal *might* be improved by involving members of the public. Research teams wishing to involve members of the public should include in their application: the aims of active involvement in this project; a description of the members of the public (to be) involved; a description of the methods of involvement; and an appropriate budget. Applications that involve members of the public will not, for that reason alone, be favoured over proposals that do not but it is hoped that the involvement of members of the public will improve the quality of the application.

Outcomes

Wherever possible, the results of HTA should provide information about the effectiveness and cost-effectiveness of care provided in its usual clinical setting and for the diverse subjects who would be eligible for the interventions under study. The endpoints of interest will in most cases include disease specific measures, health related quality of life and costs (directly and indirectly related to patient management). Wherever possible, these measurements should be made by individuals who are unaware of the treatment allocation of the subjects they are assessing. We encourage applicants to involve users of health care in the preparation of their proposal, for instance in selecting patient-oriented outcomes. A period of follow up should be undertaken which is sufficient to ensure that a wider range of effects are identified other than those which are evident immediately after treatment. These factors should guide applicants in their choice of subjects, settings and measurements made.

Sample size

A formal estimate should be made of the number of subjects required to show important differences in the chosen primary outcome measure. Justification of this estimate will be expected in the application.

Communication

Communication of the results of research to decision makers in the NHS is central to the HTA Programme. Successful applicants will be required to submit a single final report for publication by the HTA programme. They are also required to seek peer-reviewed publication of their results elsewhere and may also be asked to support the NCCHTA in further efforts to ensure that results are readily available to all relevant parties in the NHS. Where findings demonstrate continuing uncertainty, these should be highlighted as areas for further research.

Timescale

There are no fixed limits on the duration of projects or funding and proposals should be tailored to fully address the problem (including long-term follow-up if necessary). Applicants should consider however that there is a pressing need within the NHS for this research, and so the duration of the research needs to be timely.

Appendix 2 Protocol amendments summary

Trial first approved by the Research Ethics Committee in April 2009

Amendment 1: non-substantial – April 2009

- Inclusion in the adult participant information sheet (PIS) that social care records may be reviewed.
- Inclusion in the adult PIS that NHS trusts may also review relevant records.
- Clarification of the exclusion criteria, and inclusion of those in stable foster care.
- Addition of the following minimisation criterion: 'living arrangements (with parents or guardians vs. foster care)'.

Amendment 2: substantial and non-substantial – July 2009

i. Update to inclusion criteria:

   - Where the presenting episode is due to alcohol or recreational drugs, the young person has explicitly stated that he/she was intending self-harm by use of alcohol/recreational drugs.
   - Where it is intended to offer CAMHS follow-up for self-harm.

ii. Update to exclusion criteria:

   - Sibling has been randomised to SHIFT.

iii. Update to *Family therapy* section:

   - Approximately 8 sessions expected.
   - Family therapists will work in teams of 3 or 4 and provide trial FT as a team for a cluster of services. Each local CAMHS will identify a named case manager for the family therapists to link with to provide assurance about clinical governance. Where a family therapist is a full-time employee within the local CAMHS, the family therapist and case manager may be one and the same person.
   - Training and supervision protocols will be developed to ensure consistency of approach throughout the duration of the trial.
   - It is intended that there will be central review (by appropriately qualified members of the Trial Management Group) of a selection of FT tapes to ensure, and allow reporting of, overall adherence to the manual (included in the updated PIS).
   - Collection of therapist data at baseline.

iv. Use of troublesome behaviour sections of the Development and Well-Being Assessment (DAWBA) [(v) updated publication policy and (vi) piloting baseline questionnaires] for suitability for the population (health economics).

Amendment 3: substantial and non-substantial – September 2009

- Inclusion of the EQ-5D in place of the HUI2 for young people (following outcome of health economic pilot work).
- Updated information sheets to include the possibility of long-term follow-up.
- Clarification that related, unexpected SAEs will be collected for the young person only.

Amendment 4: substantial and non-substantial – October 2009

- Inclusion of the ICU in place of the DAWBA.
- Update to secondary end points to clarify that information on all further episodes of self-harm would be collected, not just the events that result in hospital attendance.
- Submission of a 'case management protocol' to the main REC for 'information only' (clarification for sites re: case management processes for families allocated SHIFT FT).

Opened to screening in the first site – November 2009

Amendment 5: non-substantial – November 2009

- Updated wording regarding exclusion of those without sufficient proficiency in English to provide data for the trial.

Amendment 6: non-substantial – November 2009

- Updated wording for the Leeds information sheets – changing 'therapist' to 'CAMHS worker'.

Amendment 7: substantial – February 2010

- Update to process for gaining consent from family members to recording FT sessions. Development of DVD consent form to be signed by each family member at every session.

Amendment 8: substantial – August 2010

- Removal of the 7-day time window for GP referrals. Participants referred by their GP could now enter SHIFT as long as 'recent self-harm was a key feature of referral'.
- Randomisation of family therapists. Participants allocated FT in the Greater Manchester, Calderdale and Bradford CAMHS were also to be randomly allocated a lead FT. It was also clarified that, in the other participating CAMHS, where there was a clear 'lead' therapist, s/he would take that lead role wherever possible.
- Recording FT sessions and transfer of data. Detail added to the protocol regarding the process for obtaining consent for recording sessions and subsequent transfer of recorded data between locations. A revised 'recording FT sessions for research' consent form was approved, which required one-off consent from each participant rather than repeated consent at each session attended.
- Follow-up processes. Researchers to contact families via telephone at 3 and 6 months prior to questionnaires being sent by the CTRU.
- Personnel changes. The main REC were informed of the new trial researchers and the involvement of 'back-up' researchers – both Mental Health Research Network (MHRN) CSOs (in London and Greater Manchester) and the Systemic Therapy for At Risk Teens (START) trial researcher in Leeds.
- Comprehensive Local Research Network (CLRN) support for screening. CLRN support by way of CAMHS records review was noted.

Amendment 9: non-substantial – August 2010

- Change of trial co-ordinator.
- General practitioner letters: CTRU to send all GP letters on CTRU-headed paper on behalf of the CAMHS.
- CLRN support for screening.

Amendment 10: substantial – November 2010

- Consent to researcher contact: this should be written wherever possible, but verbal consent acceptable if documented in the young person's case notes.
- Consent for researcher contact – where parent does not attend the appointment: accept consent to researcher contact from the young person alone where he/she is aged 16 years or older or is aged under 16 years and deemed 'Gillick competent'.
- Full trial consent: where young person does not live with parent – full trial consent will be obtained from young people aged > 16 years, regardless of whether or not they live with a parent. For young people under the age of 16 years, parental consent will be obtained wherever the parent is the primary caregiver and/or part of the 'family group'. Where a parent is not a part of the 'family group', does not live with the young person and is not involved in active parenting, consent will be acceptable from the (competent) young person and from whoever is seen as the primary caregiver, regardless of 'parental responsibility'.
- Blinding of researchers: a different researcher to undertake follow-up assessments, enabling the assessment researcher to take an active role in data collection. (An anticipatory amendment given difficulties obtaining treatment data from clinicians without researcher input.)

Amendment 11: substantial – September 2011

- Family Therapy Leaflet (developed by Leeds Family Therapy team as an additional information resource for those randomised to FT).

Amendment 12: substantial – December 2011

- Sending of Christmas cards by researchers to participants (following suggestion from previous research that this improves engagement).

Amendment 13: substantial – February 2012

- Clarification of eligibility criteria – inclusion of self-harm as a key feature of 'presentation' and exclusion of child protection investigations 'which would make treatment difficult to deliver' and 'siblings receiving family therapy in CAMHS'.
- Different member of the family can complete the SOFTA at session 3 if the consenting primary caregiver is not present.
- Clarification that clinical data can be collected by the researchers or other authorised individuals from the research team.
- Clarification of acute trust (primary outcome) data collection methods.
- Amendments to the consent form.
- Collection of multiple contacts at baseline to aid follow-up contact.
- Reducing GP contact at follow-up.
- Addition of individuals involved in development of the FT adherence protocol.

Amendment 14: substantial – June 2012

- Research Ethics Committee re: informed re: extension possibility.
- Consent form changes.
- Change in 3-month and 6-month follow-up alerting processes.
- Change in 12-month and 18-month follow-up alerting processes.
- Removal of SOFTA completion from TAU arm.

Amendment 15: substantial – August 2013

- Introduction of the use of participant incentives (£20 shopping voucher) at the end of their trial involvement.
- Amendment of the information in the PIS re: trial duration – saying that they *may* be followed up for 18 months (to allow for slightly early completion for those recruited at the end of the recruitment phase, because of grant time constraints).
- Amended wording in the 3-month letter providing clarity that TAU *is* part of SHIFT and that their data are still important even if they have not started (or had much) treatment yet.
- Additional chase and thank-you letters for questionnaires sent postally at 12 and 18 months.
- Inclusion of additional members of the trial team who will be undertaking adherence process development work.
- Submission (for information only) of the revised 'case management protocol' for sites.

Appendix 3 Questionnaire scoring details

Description of questionnaire scores, subscales, cut-off points for categorisation and scoring

Questionnaire: description of scores and cut-off points

Young person Children's Depression Rating Scale – Revised

- Higher scores represent greater levels of depression.
- **Total score:** ranges 17–113.
- **Categorisation:** not depressed (< 30), mild (30–42), moderate (43–57), severe (58–72), very severe depression (≥ 73).
- Scoring with missing data: half-rule used, scores scaled up pro rata if ≥ 50% completed.

Young person Beck Scale for Suicide Ideation

- Higher scores indicate a higher level of suicide ideation.
- **Total score:** ranges 0–38.
- **Categorisation:** suicide ideation indicated if Q4 or Q5 are non-zero, indicating: a weak or moderate desire to kill myself, or would take a chance on life or death, or would not avoid death, if found myself in a life-threatening situation.
- Scoring with missing data: half-rule used, scores scaled up pro rata if ≥ 50% completed.

Young person Hopelessness Scale

- Higher scores reflect greater hopelessness or negative expectations towards the future.
- **Total score:** ranges 0–17.
- Scoring with missing data: half-rule used, scores scaled up pro rata if ≥ 50% completed.

Young person Paediatric Quality of Life Enjoyment and Satisfaction Questionnaire measure

- Higher scores indicative of greater enjoyment and satisfaction.
- **Total score:** ranges 14–70.
- Scoring with missing data: half-rule used, scores scaled up pro rata if ≥ 50% completed.

Young person EuroQoL-5 Dimensions

- Higher utility score reflects greater quality of life.
- **Total score:** ranges 0–1.
- **Subscales:** mobility, self-care, usual activities, pain/discomfort, anxiety/depression.
- **Scoring with missing data:** missing values and ambiguous values (e.g. two boxes are ticked for a single dimension) should be treated as missing and coded as '9'. Various imputation techniques might be considered.

Caregiver Family Questionnaire

- Higher scores indicate greater levels of expressed emotion (negative).
- **Total score:** ranges 20–80.
- **Subscales:** emotional overinvolvement and criticism subscales each ranging 10–40.
- Scoring with missing data: half-rule used, scores scaled up pro rata if ≥ 50% completed.

Caregiver General Health Questionnaire, 12 questions

- Higher scores are indicative of greater psychological distress.
- **Total score:** ranges 0–36 (Likert-type scoring method of 0–1–2–3).
- Scoring with missing data: half-rule used, scores scaled up pro rata if ≥ 50% completed.

Caregiver Health Utilities Index 3

- Higher utility score reflects greater quality of life.
- **Total score:** ranges 0–1.
- **Subscales:** vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain.
- **Scoring with missing data:** respondents can answer 'do not know', for which there is no scoring instruction, to any given question. Various imputation techniques might be considered.

Young person and caregiver Inventory of Callous-Unemotional Traits

- Higher scores represent higher callous and unemotional traits.
- Completed in reference to the young person.
- Total score ranges 0–72.
- **Subscales (range):** callousness (0–33), uncaring (0–24) and unemotional (0–15).
- Scoring with missing data: half-rule used, scores scaled up pro rata if ≥ 50% completed.

Young person and caregiver Strengths and Difficulties Questionnaire

- Higher scores in all but the prosocial score represent greater issues in that category. For the prosocial score lower scores represent greater issues.
- **Total difficulties score:** ranges 0–40 (sum of all but the prosocial subscales).
- **Subscales:** five subscales – prosocial, emotional problems, conduct problems, hyperactivity and peer problems, ranging 0–10. Impact supplement ranging 0–10 in the caregiver version. Externalising (sum of conduct and hyperactivity subscales) and internalising (sum of emotional and peer problems subscales) scores ranging 0–20.
- **Categorisation:**
  - **Young person total difficulties:** close to average (0–14), raised (15–17), high (18–19), very high (20–40).
  - **Caregiver total difficulties:** close to average (0–13), raised (14–16), high (17–19), very high (20–40).
- **Scoring with missing data:** subscales are scaled up pro rata if ≥ 60% completed. Total difficulties score missing if any of the 5 subscales are missing.

Young person and caregiver McMaster Family Assessment Device

- Higher scores are indicative of poorer family functioning.
- **Overall FAD score:** all subscales range 0–4.
- **Subscales**: general functioning, behaviour control, affective involvement, affective responsiveness, roles, communication, problem-solving.
- **Categorisation:** unhealthy on each subscale for – general functioning ≥ 2.00, behaviour control ≥ 1.90, affective involvement ≥ 2.10, affective responsiveness ≥ 2.20, roles ≥ 2.30, communication ≥ 2.20, problem-solving ≥ 2.20.
- Scoring with missing data: overall score and subscales are scaled up pro rata if > 60% completed.

System for Observing Family Therapy Alliances

- Higher scores represent greater alliance.
- **Subscales:** engagement in the process, emotional connection, safety, shared sense of purpose ranging 4–20.
- **Scoring with missing data:** half-rule used, scores scaled up pro rata if ≥ 50% completed.

Appendix 4 Recruitment, retention and characteristics: additional tables

TABLE 102 Recruitment by CAMHS

Hub and trust	CAMHS	Date opened to screening	FT (*N* = 415), *n* (%)	TAU (*N* = 417), *n* (%)	Total (*N* = 832), *n* (%)
Yorkshire			***149 (35.9)***	***151 (36.2)***	***300 (36.1)***
Y trust 1	Y1	23 November 2009	29 (7.0)	29 (7.0)	58 (7.0)
	Y2	23 November 2009	11 (2.7)	10 (2.4)	21 (2.5)
	Y3	23 November 2009	18 (4.3)	17 (4.1)	35 (4.2)
Y trust 2	Y4	14 June 2010	16 (3.9)	15 (3.6)	31 (3.7)
	Y5	14 June 2010	2 (0.5)	1 (0.2)	3 (0.4)
Y trust 3	Y6	14 June 2010	9 (2.2)	6 (1.4)	15 (1.8)
	Y7	14 June 2010	18 (4.3)	19 (4.6)	37 (4.4)
Y trust 4	Y8	23 September 2010	5 (1.2)	9 (2.2)	14 (1.7)
	Y9	18 October 2010	11 (2.7)	10 (2.4)	21 (2.5)
	Y10	18 October 2010	1 (0.2)	1 (0.2)	2 (0.2)
Y trust 5	Y11	27 September 2010	14 (3.4)	15 (3.6)	29 (3.5)
Y trust 6	Y12	28 September 2010	7 (1.7)	9 (2.2)	16 (1.9)
	Y13	28 September 2010	1 (0.2)	3 (0.7)	4 (0.5)
	Y14	28 September 2010	7 (1.7)	7 (1.7)	14 (1.7)
Manchester			***143 (34.5)***	***143 (34.3)***	***286 (34.4)***
M trust 1	M1	19 July 2010	6 (1.4)	5 (1.2)	11 (1.3)
	M2	19 July 2010	9 (2.2)	7 (1.7)	16 (1.9)
	M3	19 July 2010	5 (1.2)	5 (1.2)	10 (1.2)
	M4	19 July 2010	7 (1.7)	9 (2.2)	16 (1.9)
M trust 2	M5	19 July 2010	12 (2.9)	13 (3.1)	25 (3.0)
	M6	19 July 2010	12 (2.9)	13 (3.1)	25 (3.0)
	M7	19 July 2010	6 (1.4)	6 (1.4)	12 (1.4)
	M8	19 July 2010	4 (1.0)	4 (1.0)	8 (1.0)
	M9	19 July 2010	6 (1.4)	7 (1.7)	13 (1.6)
	M10	17 February 2011	13 (3.1)	9 (2.2)	22 (2.6)
M trust 3	M11	19 July 2010	22 (5.3)	21 (5.0)	43 (5.2)
	M12	19 July 2010	20 (4.8)	19 (4.6)	39 (4.7)
	M13	19 July 2010	6 (1.4)	5 (1.2)	11 (1.3)
	M14	19 July 2010	8 (1.9)	12 (2.9)	20 (2.4)
	M15	19 July 2010	7 (1.7)	8 (1.9)	15 (1.8)

continued

TABLE 102 Recruitment by CAMHS (*continued*)

Hub and trust	CAMHS	Date opened to screening	FT (*N* = 415), *n* (%)	TAU (*N* = 417), *n* (%)	Total (*N* = 832), *n* (%)
London			***123 (29.6)***	***123 (29.5)***	***246 (29.6)***
L trust 1	L1	14 January 2010	18 (4.3)	18 (4.3)	36 (4.3)
	L2	14 January 2010	18 (4.3)	18 (4.3)	36 (4.3)
	L3	14 January 2010	9 (2.2)	10 (2.4)	19 (2.3)
	L4	14 January 2010	9 (2.2)	7 (1.7)	16 (1.9)
L trust 2	L5	14 January 2010	7 (1.7)	7 (1.7)	14 (1.7)
	L6	14 January 2010	12 (2.9)	10 (2.4)	22 (2.6)
	L7	14 January 2010	18 (4.3)	19 (4.6)	37 (4.4)
L trust 3	L8	6 September 2011	9 (2.2)	7 (1.7)	16 (1.9)
L trust 4	L9	5 May 2011	3 (0.7)	3 (0.7)	6 (0.7)
L trust 5	L10	5 May 2011	17 (4.1)	17 (4.1)	34 (4.1)
L trust 6	L11	22 January 2013	3 (0.7)	7 (1.7)	10 (1.2)

TABLE 103 All self-harm methods, treatment and suicide attempts reported over the past year (SASII timeline)

Self-harm method/s (not mutually exclusive)	FT (*N* = 415), *n* (%)	TAU (*N* = 417), *n* (%)	Total (*N* = 832), *n* (%)
Alcohol	0 (0.0)	2 (0.5)	2 (0.2)
Drugs/medications	142 (34.2)	149 (35.7)	291 (35.0)
Poison/caustic substance	5 (1.2)	4 (1.0)	9 (1.1)
Burning	17 (4.1)	15 (3.6)	32 (3.8)
Scratch/cut	338 (81.4)	352 (84.4)	690 (82.9)
Stabbing, puncture	10 (2.4)	8 (1.9)	18 (2.2)
Hanging	7 (1.7)	14 (3.4)	21 (2.5)
Strangling	17 (4.1)	8 (1.9)	25 (3.0)
Asphyxiation	0 (0.0)	1 (0.2)	1 (0.1)
Jumping	11 (2.7)	5 (1.2)	16 (1.9)
Drowning	5 (1.2)	7 (1.7)	12 (1.4)
Hitting body	32 (7.7)	29 (7.0)	61 (7.3)
Stopped required medical treatments or medications	1 (0.2)	2 (0.5)	3 (0.4)
Transportation-related injury	0 (0.0)	1 (0.2)	1 (0.1)
Stepped into traffic	6 (1.4)	7 (1.7)	13 (1.6)
Method not specified	8 (1.9)	3 (0.7)	11 (1.3)
Other	39 (9.4)	53 (12.7)	92 (11.1)
Missing	13 (3.1)	2 (0.5)	15 (1.8)
Most serious medical treatment required			
Hospital medical treatment	127 (30.6)	126 (30.2)	253 (30.4)
Non-hospital medical treatment	132 (31.8)	134 (32.1)	266 (32.0)
No medical treatment	145 (34.9)	156 (37.4)	301 (36.2)
Missing	11 (2.7)	1 (0.2)	12 (1.4)

Appendix 5 Primary outcome: further details and tables

Proportional hazards assumption

Figures 47–62 present the log-cumulative hazard plots of time to self-harm for treatment and covariates included in the modelling, and are used to demonstrate the proportional hazards assumption is met if lines are parallel.

Investigation of *Figure 47* for treatment suggests some crossover in hazards around log(time) = 0 – 2.2; however, at this point there is little differentiation between the curves as they are very close. With the exception of trust, the remaining covariates adequately satisfy the proportional hazards assumptions, as, although there is some early cross-over in lines for some covariates, this occurs prior to log(time) = 1 corresponding to the first few months of follow-up when the Kaplan–Meier curves are very close, after which there is differentiation between roughly parallel lines.

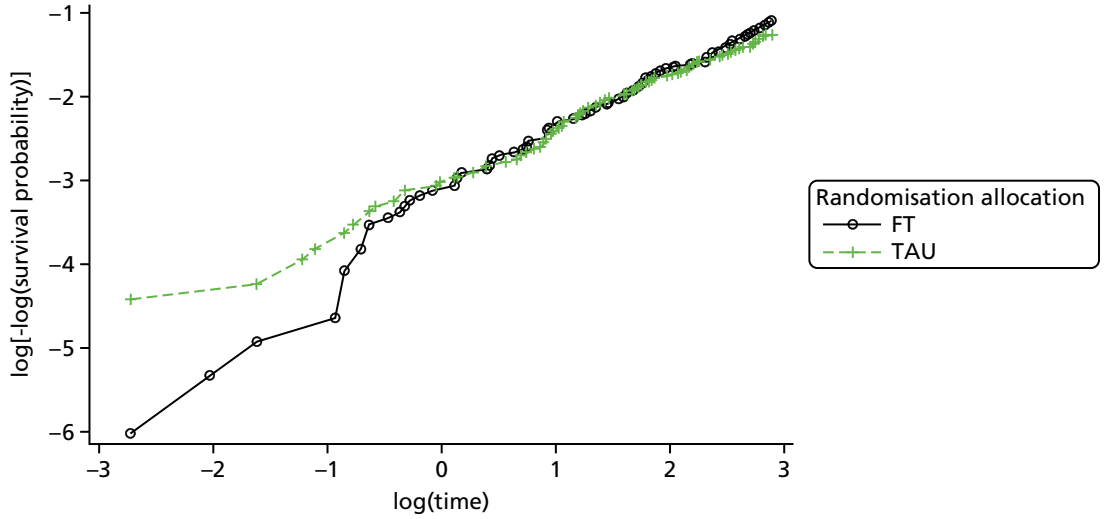


FIGURE 47 Proportional hazards assumption for treatment: log of negative log of estimated survivor functions.

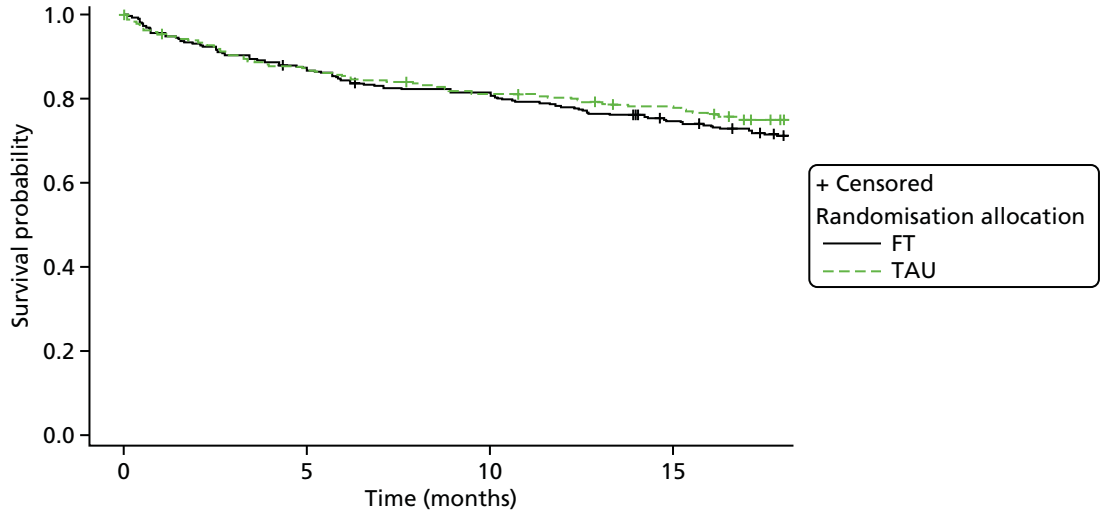


FIGURE 48 Kaplan–Meier plot of time to self-harm by treatment: product limit survival estimates.

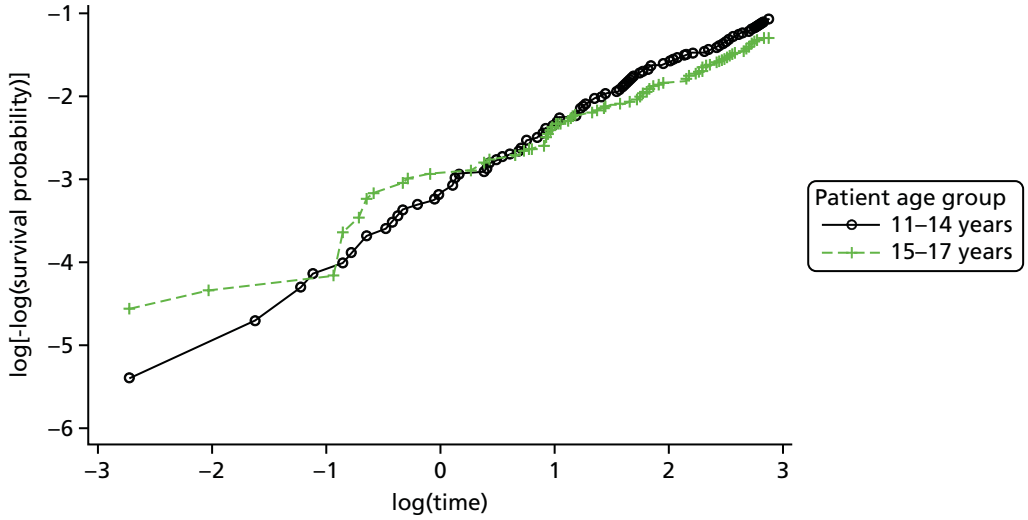


FIGURE 49 Proportional hazards assumption for age group: log of negative log of estimated survivor functions.

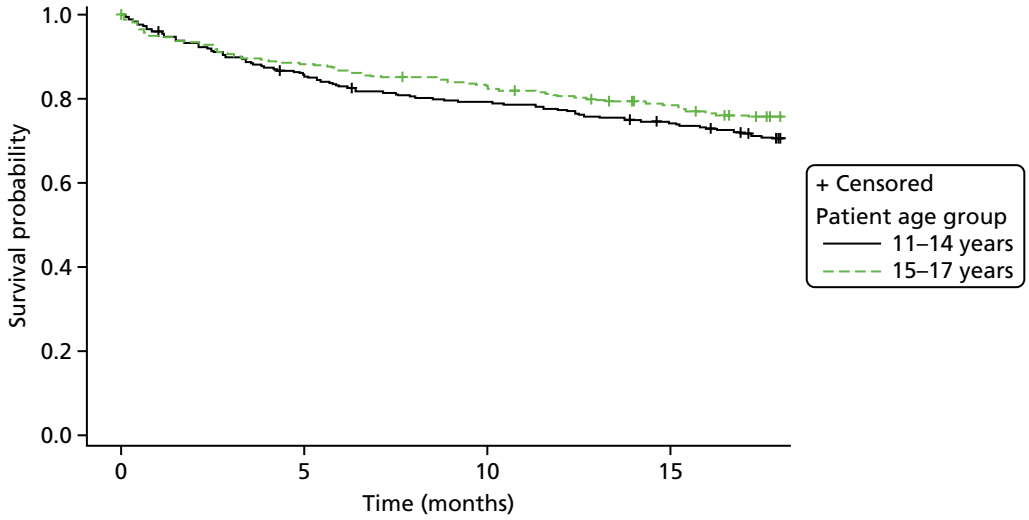


FIGURE 50 Kaplan–Meier plot of time to self-harm by age group: product limit survival estimates.

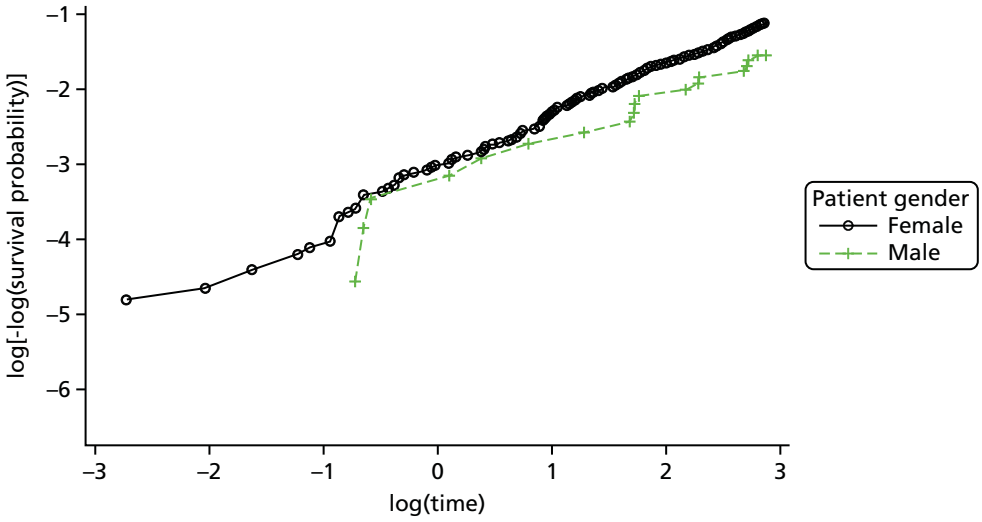


FIGURE 51 Proportional hazards assumption for gender: log of negative log of estimated survivor functions.

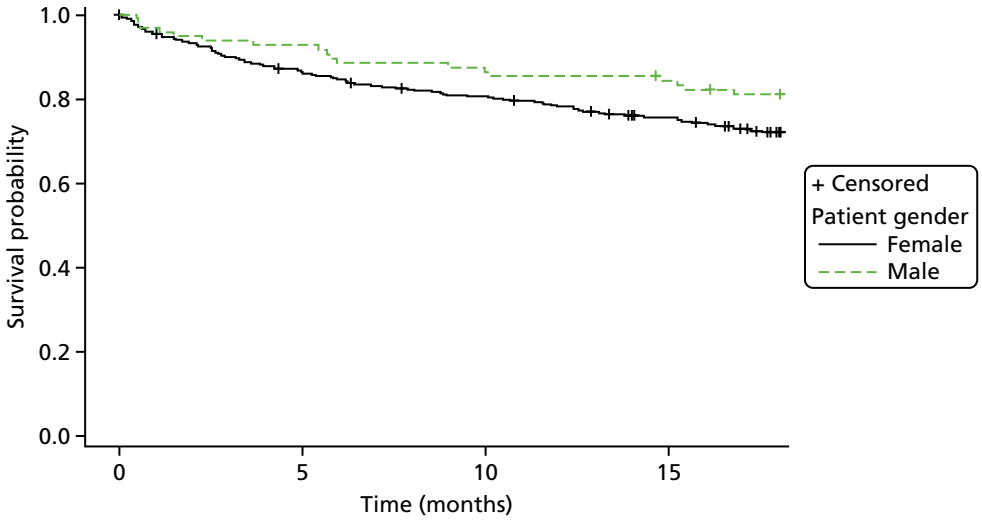


FIGURE 52 Kaplan–Meier plot of time to self-harm by gender: product limit survival estimates.

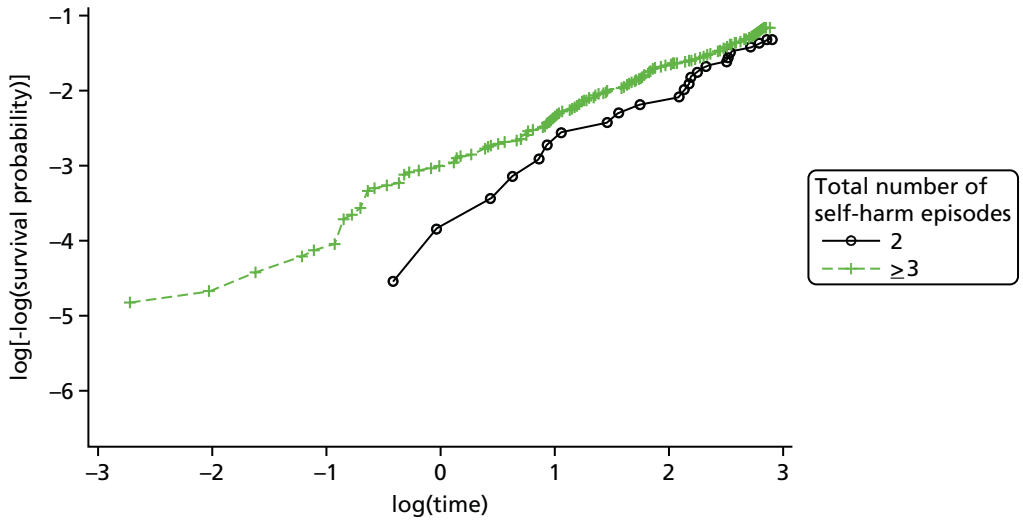


FIGURE 53 Proportional hazards assumption for number of previous self-harm episodes at baseline: log of negative log of estimated survivor functions.

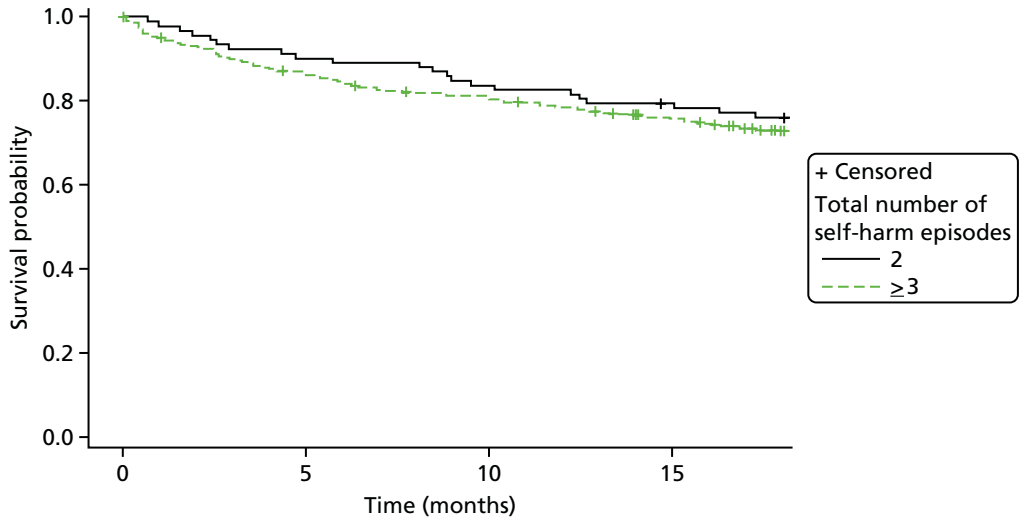


FIGURE 54 Kaplan–Meier plot of time to self-harm by the number of previous self-harm episodes at baseline: limit survival estimates.

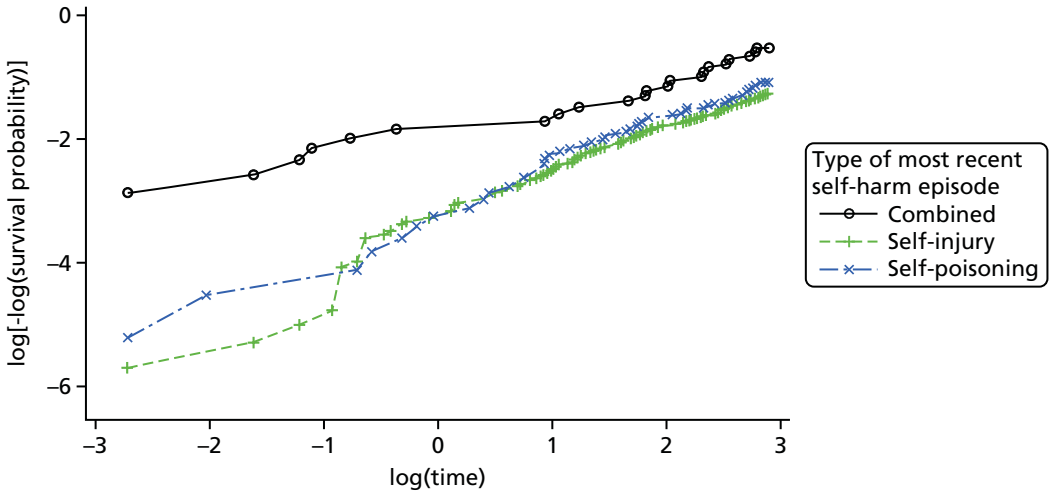


FIGURE 55 Proportional hazards assumption for the type of index self-harm episode: log of negative log of estimated survivor functions.

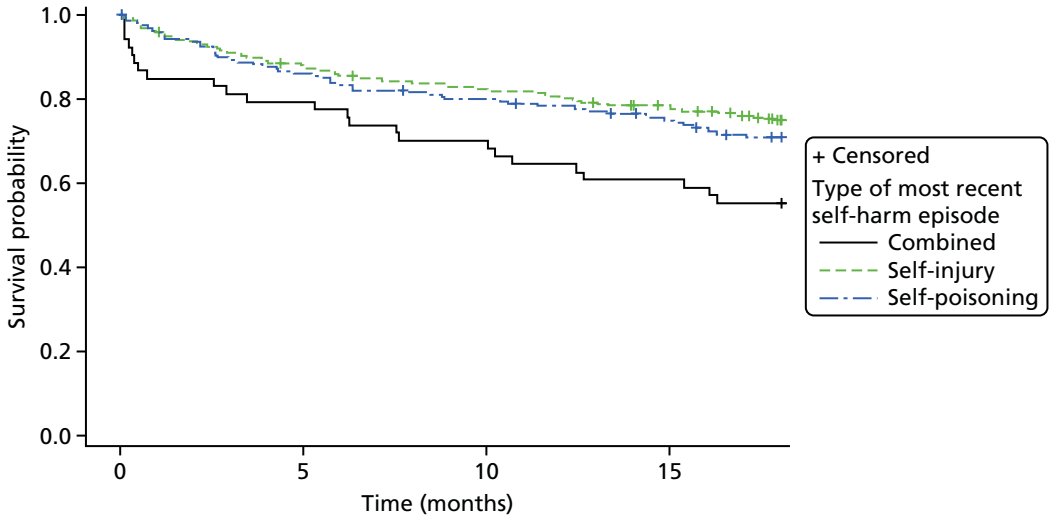


FIGURE 56 Kaplan–Meier plot of time to self-harm by the type of index self-harm episode: product limit survival estimates.

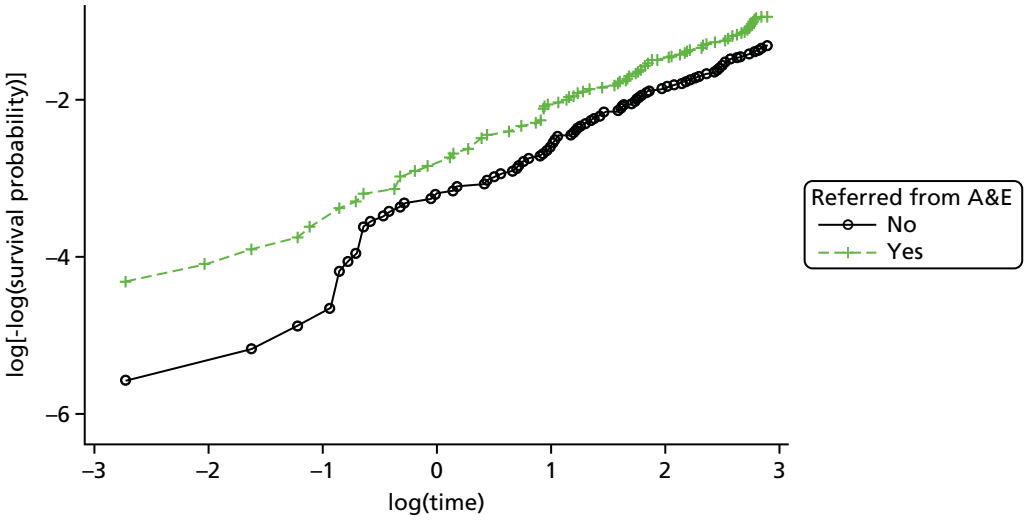


FIGURE 57 Proportional hazards assumption for referral from A&E: log of negative log of estimated survivor functions.

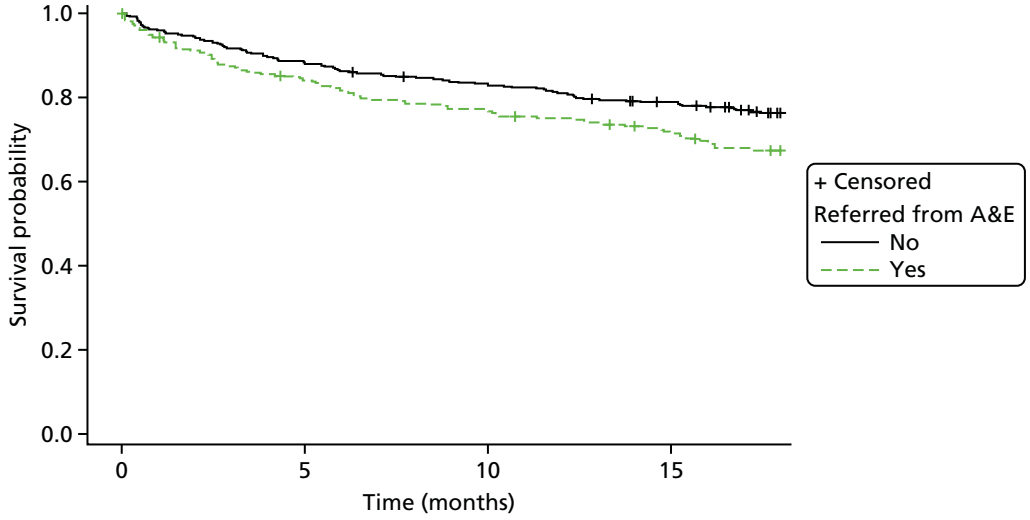


FIGURE 58 Kaplan–Meier plot of time to self-harm by referral from A&E: product limit survival estimates.

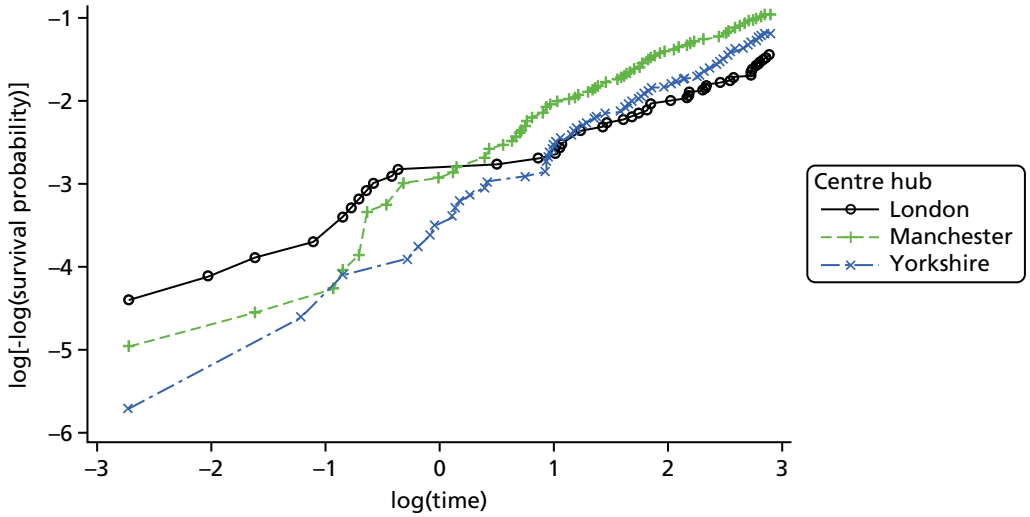


FIGURE 59 Proportional hazards assumption for hub: log of negative log of estimated survivor functions.

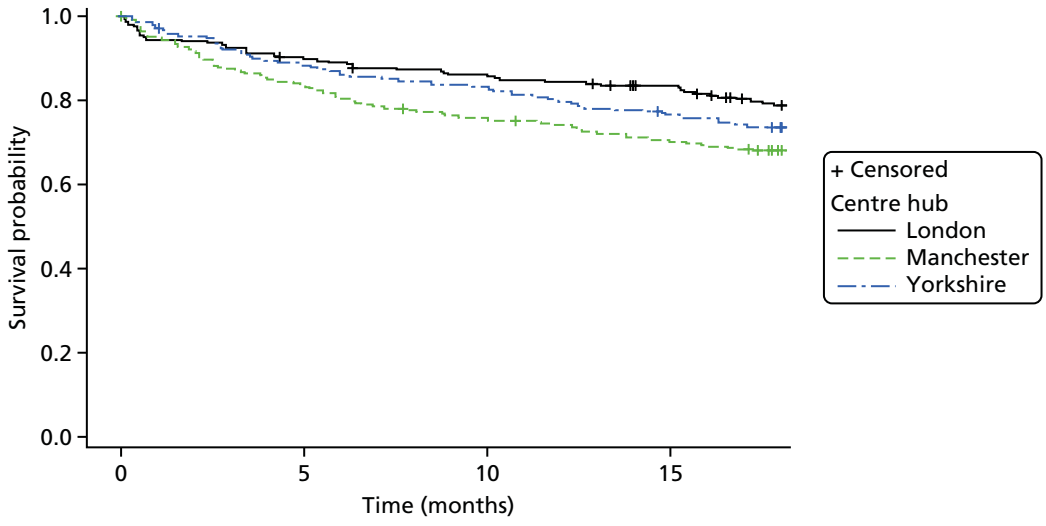


FIGURE 60 Kaplan–Meier plot of time to self-harm by hub: product limit survival estimates.

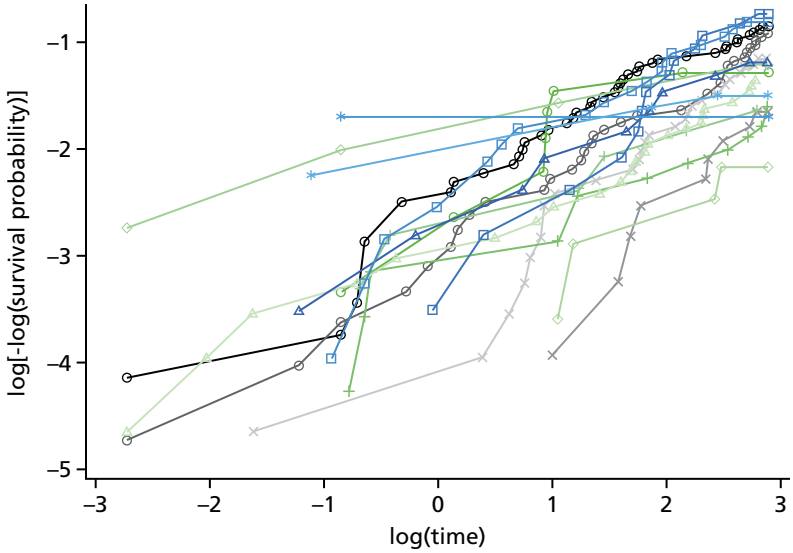


FIGURE 61 Proportional hazards assumption for trust: log of negative log of estimated survivor functions.

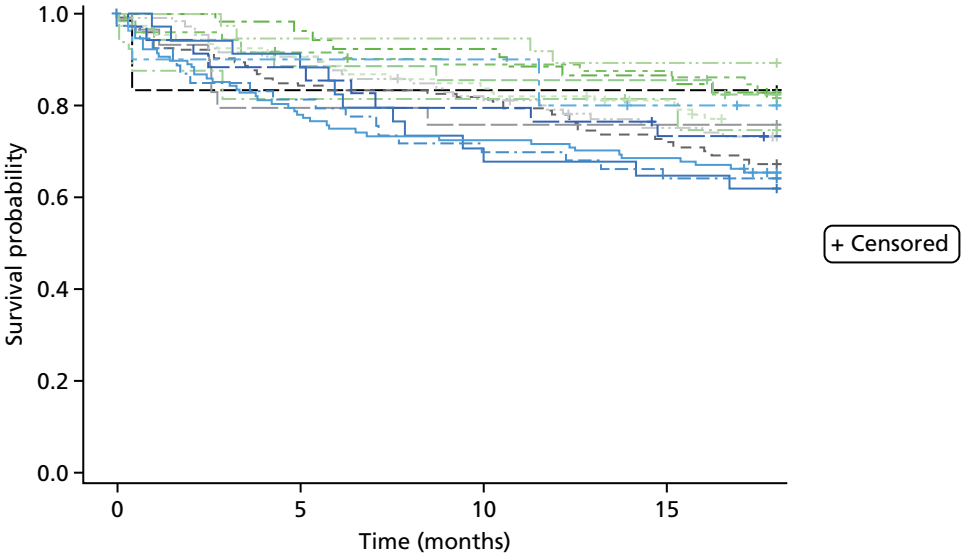


FIGURE 62 Kaplan–Meier plot of time to self-harm by trust: product limit survival estimates.

Table 104 further presents the results of the Kolmogorov-type supremum test of the validity of the proportional hazards assumption for treatment and covariates based on 1000 simulations. There is no evidence to reject the assumption of proportional hazards for treatment ($p = 0.2950$) or other covariates, again with the exception of trust for which it can be seen that the proportional hazards assumption is violated for trusts 5 Boroughs Partnership NHS Foundation Trust ($p = 0.0610$) and York ($p = 0.0490$).

Figures 63 and *64* further investigate the adequacy of the model. In *Figure 63* (index plot of deviance residuals) it can be seen that, although there are outlying values [cases poorly predicted by the model lying outside (–2 to 2)], there is a similar pattern of outliers among the two treatment groups. As per the log-cumulative hazard plot and Kolmogorov-type supremum test, *Figure 64* (scatterplot of Schoenfeld residuals) assesses the proportional hazard assumption for treatment, in which a horizontal line represents proportional hazards. There is a slight non-horizontal trend, suggesting that the HR varies with respect to time; however, as per the previous tests, this is not a significant trend.

TABLE 104 Supremum test for proportional hazards assumption: model 1

Variable	Max. absolute value	Replications	Probability > max. absolute value
Treatment: TAU vs. FT	0.9381	1000	0.2950
Sex: male vs. female	0.4704	1000	0.9120
Age group (years): 11–14 vs. 15–17	0.7532	1000	0.5880
Self-harm number: 2 vs. ≥ 3	0.9464	1000	0.2710
Self-harm type: self-injury vs. combined	1.1758	1000	0.1910
Self-harm type: self-injury vs. self-poisoning	0.6618	1000	0.8840
Hospital referral: no vs. yes	0.8251	1000	0.7340
Centre trust: Y trust 1 vs.			
M trust 3	1.6729	1000	0.0610
L trust 5	0.6133	1000	0.6730
Y trust 3	1.1154	1000	0.1560
Y trust 2	0.5226	1000	0.8460
M trust 1	0.8751	1000	0.4580
L trust 6	0.6387	1000	0.5030
Y trust 4	0.6800	1000	0.4720
L trust 2	0.8643	1000	0.4390
M trust 2	1.1367	1000	0.2790
L trust 1	0.7393	1000	0.7450
Y trust 6	1.2182	1000	0.1750
L trust 4	0.9399	1000	0.2980
L trust 3	0.8893	1000	0.3550
Y trust 5	1.4832	1000	0.0490

Max., maximum.

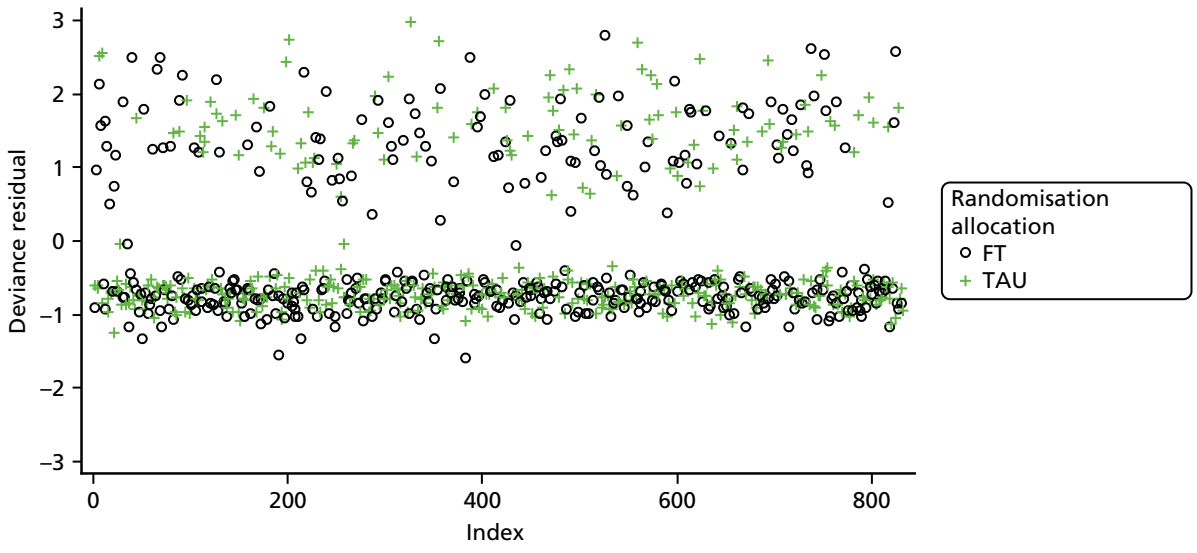


FIGURE 63 Index plot of deviance residuals for model 1 by treatment.

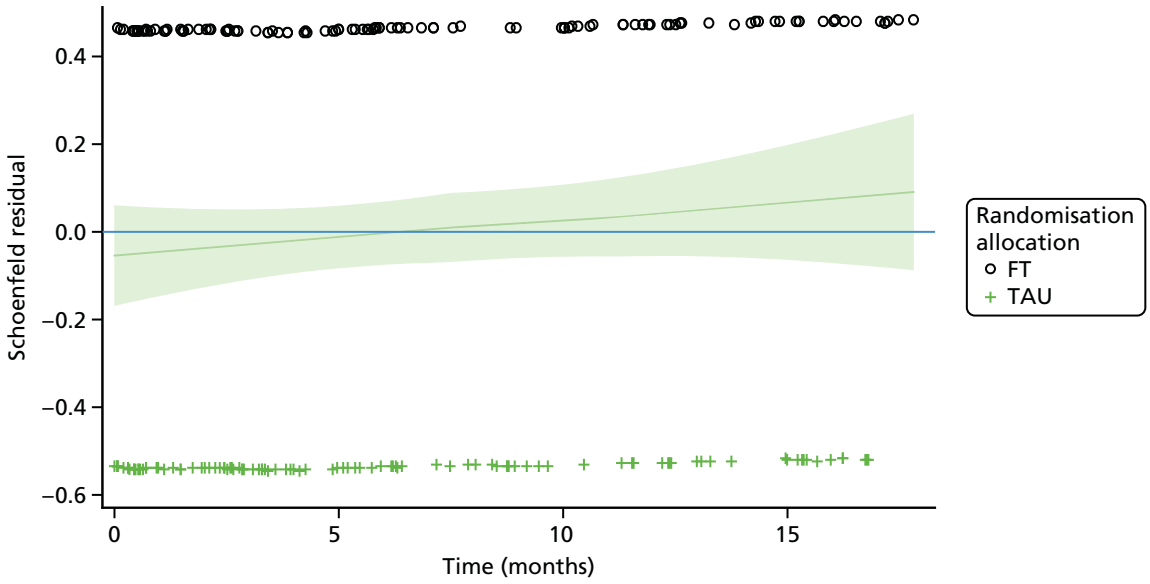


FIGURE 64 Scatterplot of Schoenfeld residuals for model 1 by treatment.

TABLE 105 Gender by recruiting trust

Trust	Gender, *n* (%)	
	Male	Female
Y trust 1 (*n* = 114)	14 (12.3)	100 (87.7)
Y trust 2 (*n* = 34)	5 (14.7)	29 (85.3)
Y trust 2 (*n* = 52)	11 (21.2)	41 (78.8)
Y trust 2 (*n* = 37)	2 (5.4)	35 (94.6)
Y trust 2 (*n* = 29)	3 (10.3)	26 (89.7)
Y trust 2 (*n* = 34)	5 (14.7)	29 (85.3)
M trust 1 (*n* = 53)	7 (13.2)	46 (86.8)
M trust 1 (*n* = 105)	8 (7.6)	97 (92.4)
M trust 1 (*n* = 128)	21 (16.4)	107 (83.6)
L trust 1 (*n* = 107)	9 (8.4)	98 (91.6)
L trust 2 (*n* = 73)	3 (4.1)	70 (95.9)
L trust 3 (*n* = 16)	1 (6.3)	15 (93.8)
L trust 4 (*n* = 6)	0 (0.0)	6 (100.0)
L trust 5 (*n* = 34)	5 (14.7)	29 (85.3)
L trust 6 (*n* = 10)	1 (10.0)	9 (90.0)
Total (*n* = 832)	95 (11.4)	737 (88.6)

TABLE 106 Referral into CAMHS via hospital by trust

Trust	Referral into CAMHS via hospital, *n* (%)	
	Yes	No
Y trust 1 ($n = 114$)	49 (43.0)	65 (57.0)
Y trust 2 ($n = 34$)	20 (58.8)	14 (41.2)
Y trust 2 ($n = 52$)	17 (32.7)	35 (67.3)
Y trust 2 ($n = 37$)	5 (13.5)	32 (86.5)
Y trust 2 ($n = 29$)	9 (31.0)	20 (69.0)
Y trust 2 ($n = 34$)	7 (20.6)	27 (79.4)
M trust 1 ($n = 53$)	22 (41.5)	31 (58.5)
M trust 1 ($n = 105$)	36 (34.3)	69 (65.7)
M trust 1 ($n = 128$)	60 (46.9)	68 (53.1)
L trust 1 ($n = 107$)	42 (39.3)	65 (60.7)
L trust 2 ($n = 73$)	27 (37.0)	46 (63.0)
L trust 3 ($n = 16$)	6 (37.5)	10 (62.5)
L trust 4 ($n = 6$)	0 (0.0)	6 (100.0)
L trust 5 ($n = 34$)	2 (5.9)	32 (94.1)
L trust 6 ($n = 10$)	2 (20.0)	8 (80.0)
Total ($n = 832$)	304 (36.5)	528 (63.5)

Chi-squared test for an association between recruiting trust and referral into CAMHS via hospital: $p = 0.0000148314$ (14 degrees of freedom).

Sensitivity analysis

TABLE 107 Cox proportional hazards model including unclassified events (adjusted for covariates)

Covariate	Parameter estimate	SE	HR (95% CI)	df	*p*-value
Treatment: FT (vs. TAU)	0.14	0.13	1.15 (0.89 to 1.49)	1	0.2736
Gender: female (vs. male)	0.50	0.24	1.65 (1.03 to 2.65)	1	0.0388
Age group (years): 15–17 (vs. 11–14)	–0.28	0.14	0.75 (0.58 to 0.98)	1	0.0382
Number of previous self-harm episodes: ≥ 3 (vs. 2)	0.18	0.22	1.20 (0.78 to 1.85)	1	0.4107
Type of index episode				***2***	***0.0200***
Combined (vs. self-injury)	0.64	0.24	1.90 (1.20 to 3.02)	1	–
Self-poisoning (vs. self-injury)	0.08	0.20	1.09 (0.74 to 1.60)	1	–
Referred via A&E: yes (vs. no)	0.22	0.17	1.24 (0.88 to 1.74)	1	0.2121
Trust (vs. Y trust 1)				14	0.0767

df, degrees of freedom.

Clustering because of therapists: intraclass correlation coefficient

Frailty model

The frailty model with log-normal frailty for therapist effects provides an estimate of covariance, that is, the between-cluster variance of 0.08535 (SE 0.09254) for participants treated by the same therapist. There is, however, no standard formula available or literature supporting the calculation of the ICC from this model based on validated methods implementable in standard statistical software. This estimate of covariance was not found to be statistically significant and did not improve the fit of the model (based on the marginal likelihood) compared with that without the random therapist effect.

Logistic model

To allow us to investigate the likely ICC, we dichotomised time to self-harm according to whether or not an event was observed. Estimates were obtained using a logistic regression random intercept model, both with no covariates and separately including covariates (with hub rather than centre-fixed effect) in the linear predictor as fixed effects. We obtained an estimate of the logistic ICC for the overall data set and estimated the ICC for each study arm separately; for the latter, we applied the method to subsets of the data corresponding to the study arm. The latent intraclass correlation was calculated based on (as the standard logistic distribution has variance $\pi^2/3$):

$$\rho_{\text{logit}} = \frac{\sigma_u^2}{\sigma_u^2 + \pi^2/3}. \tag{7}$$

The unadjusted ICC was 0.0415 overall across both treatment arms, 0.0395 in the FT arm and 0.0108 in the TAU arm. Furthermore, the adjusted ICC over both treatment arms reduced to 0.0185.

Adjusted ICCs by arm were not calculated, as in the presence of covariates it is not appropriate to obtain these independently for subsets of the data. Estimates of the ICC from this model should be interpreted with caution as the logistic random effect model assumes the random effect follows a normal distribution, however, this is violated when there are lots of clusters of size one (as we have here with many therapists seeing only a single participant), leading to a high proportion of clusters within which the proportion of self-harm is zero or one. The ICCs from the logistic model are therefore included to report only on the plausible range of the ICC, with all estimates reported to be < 0.05.

Appendix 6 Secondary outcomes: further details and tables

Recurrent event analysis

Figure 65 (scatterplot of Schoenfeld residuals) assesses the proportional hazard assumption for treatment in the recurrent event model, in which a horizontal line represents proportional hazards. There is a slight non-horizontal trend, suggesting that the HR varies with respect to time; however, this is in line with the primary analysis model for which further tests confirmed this was not a significant trend.

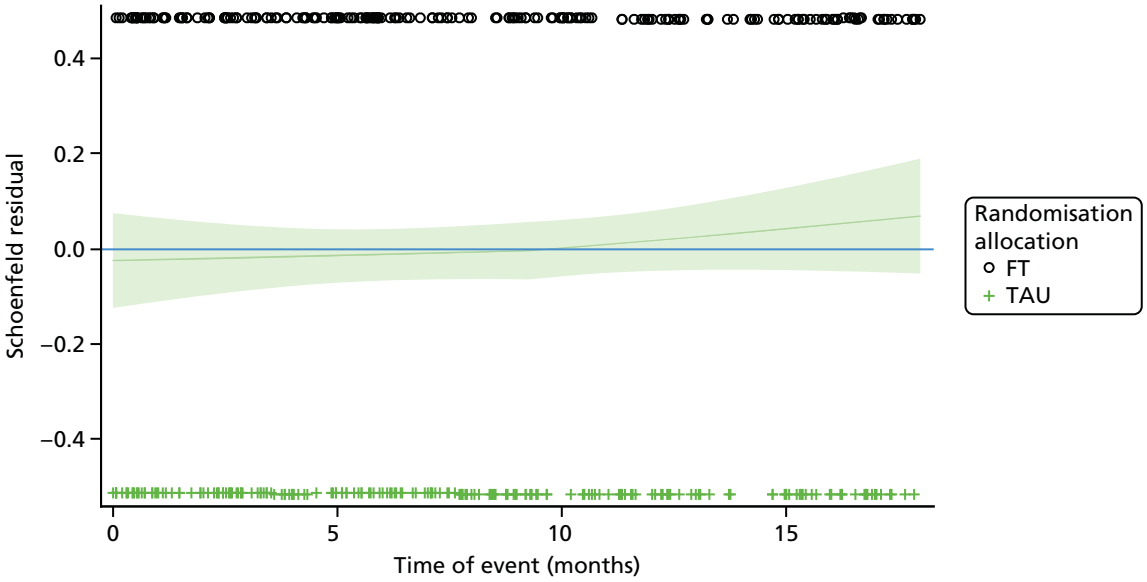


FIGURE 65 Model checking for the counting process model with robust sandwich variance estimator for recurrent events: scatterplot of Schoenfeld residuals for treatment; Andersen–Gill model with robust sandwich variance estimator for recurrent events.

Missing data patterns

TABLE 108 Missing data pattern for questionnaire outcomes

Missing data pattern				Young person, *n* (%)						Caregiver, *n* (%)			Caregiver (3 and 6 months), *n* (%)
Pattern	Baseline	12 months	18 months	BSS	PQ-LES-Q[a]	CDRS-R	SDQ[b]	Hopelessness Scale	McMaster FAD[a]	SDQ[b]	GHQ-12	McMaster FAD[a]	Family Questionnaire[c]
1	Yes	Yes	Yes	313 (37.6)	320 (38.5)	301 (36.2)	327 (39.3)	312 (37.5)	315 (37.9)	317 (38.1)	316 (38.0)	313 (37.6)	316 (38.0)
2	Yes	Yes	No	136 (16.3)	133 (16.0)	129 (15.5)	136 (16.3)	134 (16.1)	135 (16.2)	130 (15.6)	129 (15.5)	128 (15.4)	123 (14.8)
3	Yes	No	Yes	69 (8.3)	67 (8.1)	67 (8.1)	66 (7.9)	70 (8.4)	67 (8.1)	70 (8.4)	75 (9.0)	74 (8.9)	44 (5.3)
4	Yes	No	No	298 (35.8)	299 (35.9)	334 (40.1)	299 (35.9)	299 (35.9)	292 (35.1)	310 (37.3)	309 (37.1)	308 (37.0)	348 (41.8)
5	No	Yes	Yes	8 (1.0)	4 (0.5)	1 (0.1)	1 (0.1)	8 (1.0)	7 (0.8)	0 (0.0)	0 (0.0)	2 (0.2)	0 (0.0)
6	No	Yes	No	2 (0.2)	3 (0.4)	0 (0.0)	1 (0.1)	2 (0.2)	3 (0.4)	0 (0.0)	2 (0.2)	1 (0.1)	0 (0.0)
7	No	No	Yes	2 (0.2)	2 (0.2	0 (0.0)	1 (0.1)	2 (0.2)	3 (0.4)	1 (0.1)	0 (0.0)	0 (0.0)	0 (0.0)
8	No	No	No	4 (0.5)	4 (0.5)	0 (0.0)	1 (0.1)	5 (0.6)	10 (1.2)	4 (0.5)	1 (0.1)	6 (0.7)	1 (0.1)
Missing items over all time points				829 (33.2)	828 (33.2)	865 (34.7)	808 (32.4)	833 (33.4)	835 (33.5)	834 (33.4)	829 (33.2)	840 (33.7)	866 (34.7)

a Total score.
b Total difficulties score.
c Total score (at 3 and 6 months rather than 12 and 18 months).

Responder characteristics

Two models were fitted to explore participants' response status (followed up or lost to follow-up) by participant baseline characteristics and treatment arm. The first (model 1) modelled response status with main effects for treatment arm, the participant characteristic and the interaction between the two to evaluate whether or not the response profile differed by the participant characteristic overall and further differentially by arm. As tests for interaction do not provide a very powerful test for moderation, a 10% significance level is used and we therefore also modelled response status by participant characteristics separately for each treatment arm (model 2), providing separate tests of the main characteristic effect in each arm to examine any differential effects.

Results are presented in *Tables 109–112* and green cells are used to highlight significant effects at the 10% level; interpretation is detailed in *Chapter 3*, *Responder characteristics*.

TABLE 109 Distribution of 12-month questionnaire response status by participant characteristics and treatment arm with tests for main and interaction effects

	FT: responder?		TAU: responder?		Model 1: interaction test[a]		Model 2: main effect test[b]	
Variable	Yes ($N = 261$), *n* (%)	No ($N = 154$), *n* (%)	Yes ($N = 204$), *n* (%)	No ($N = 213$), *n* (%)	Main effect *p*-value	Interaction effect *p*-value	FT *p*-value	TAU *p*-value
Self-harm by 12 months (yes)	56 (21.5)	35 (22.7)	35 (17.2)	47 (22.1)	0.2674	0.4932	0.7624	0.2084
Patient gender (female)	230 (88.1)	138 (89.6)	181 (88.7)	188 (88.3)	0.8145	0.6618	0.6443	0.8824
Total number of self-harm episodes (≥ 3)	233 (89.3)	136 (88.3)	183 (89.7)	187 (87.8)	0.5187	0.8310	0.7634	0.5375
Type of most recent self-harm episode						–	–	–
Self-poisoning	59 (22.6)	34 (22.1)	39 (19.1)	52 (24.4)		–	–	–
Self-injury	187 (71.6)	110 (71.4)	156 (76.5)	141 (66.2)	0.0559	0.1042	0.8407	0.0115
Combined	15 (5.7)	10 (6.5)	9 (4.4)	20 (9.4)	0.1598	0.3408	0.7508	0.0926
Referred from A&E (yes)	90 (34.5)	66 (42.9)	68 (33.3)	80 (37.6)	0.0654	0.5627	0.0894	0.3676

a Response status ≈treatment + characteristic + (treatment × characteristic).
b Response status ≈characteristic. Interaction and main effect test via logistic regression (Wald chi-squared tests).
Green cells are used to highlight significant effects at the 10% level.

TABLE 110 Distribution of 18-month questionnaire response status by participant characteristics and treatment arm with tests for main and interaction effects

Variable	FT: responder?		TAU: responder?		Model 1: interaction test[a]		Model 2: main effect test[b]	
	Yes ($N = 261$), *n* (%)	No ($N = 154$), *n* (%)	Yes ($N = 204$), *n* (%)	No ($N = 213$), *n* (%)	Main effect *p*-value	Interaction effect *p*-value	FT *p*-value	TAU *p*-value
Self-harm by 18 months (yes)	57 (26.8)	61 (30.2)	41 (22.5)	62 (26.4)	0.2340	0.8995	0.4380	0.3657
Patient gender (female)	180 (84.5)	188 (93.1)	158 (86.8)	211 (89.8)	0.0089	0.1791	0.0073	0.3465
Total number of self-harm episodes (≥ 3)	193 (90.6)	176 (87.1)	161 (88.5)	209 (88.9)	0.4882	0.3644	0.2606	0.8789
Type of most recent self-harm episode						–	–	–
Self-poisoning	43 (20.2)	50 (24.8)	37 (20.3)	54 (23.0)		–	–	–
Self-injury	158 (74.2)	139 (68.8)	133 (73.1)	164 (69.8)	0.2611	0.7993	0.3375	0.5323
Combined	12 (5.6)	13 (6.4)	12 (6.6)	17 (7.2)	0.8305	0.9804	0.8693	0.8906
Referred from A&E (yes)	71 (33.3)	85 (42.1)	54 (29.7)	94 (40.0)	0.0045	0.7741	0.0664	0.0293

a Response status ≈TRT + characteristic + TRT*characteristic.
b Response status ≈characteristic. Interaction and main effect test via logistic regression (Wald chi-squared tests).
Green cells are used to highlight significant effects at the 10% level.

TABLE 111 Distribution of 12-month questionnaire response status by participant characteristics and treatment arm with tests for main and interaction effects

Variable	FT: responder? Mean (SD)		TAU: responder? Mean (SD)		Model 1: interaction test[a]		Model 2: main effect test[b]	
	Yes ($N = 261$), *n* (%)	No ($N = 154$), *n* (%)	Yes ($N = 204$), *n* (%)	No ($N = 213$), *n* (%)	Main effect *p*-value	Interaction effect *p*-value	FT *p*-value	TAU *p*-value
Age	14.3 (1.36)	14.4 (1.42)	14.3 (1.34)	14.4 (1.37)	0.3749	0.8086	0.4288	0.6452
Young person Hopelessness Scale	7.6 (4.31)	7.9 (4.29)	7.3 (4.08)	7.3 (4.30)	0.7817	0.6210	0.5861	0.8776
Young person BSS	10.7 (8.81)	11.1 (9.18)	10.2 (9.11)	10.5 (9.72)	0.6183	0.9238	0.6865	0.7663
Young person CDRS-R	47.8 (14.25)	48.3 (14.12)	48.5 (13.70)	50.4 (12.85)	0.1901	0.4270	0.7101	0.1439
Young person PQ-LES-Q	41.6 (9.45)	40.6 (9.23)	41.4 (9.31)	41.0 (9.60)	0.2894	0.6877	0.3113	0.6350
Young person SDQ total difficulties	19.3 (5.76)	20.0 (5.60)	20.1 (5.41)	20.1 (5.79)	0.4265	0.3728	0.2389	0.9451
Young person SDQ prosocial	7.0 (1.86)	7.3 (1.80)	7.3 (1.87)	7.1 (1.87)	0.5762	0.1190	0.1463	0.4659
Young person SDQ impact	3.3 (2.30)	3.4 (2.60)	3.5 (2.40)	3.6 (2.63)	0.4251	0.8282	0.4883	0.6702
Young person FAD total score	2.4 (0.33)	2.5 (0.34)	2.4 (0.36)	2.5 (0.36)	0.0280	0.9352	0.1559	0.0873
Young person ICU total score	28.0 (9.15)	28.7 (9.02)	28.1 (8.93)	28.9 (9.25)	0.2635	0.8999	0.4870	0.3743
Caregiver SDQ total difficulties	19.3 (6.36)	19.6 (6.91)	19.2 (6.70)	20.4 (6.91)	0.1233	0.3696	0.6601	0.0734
Caregiver SDQ emotional problems	6.2 (2.47)	6.2 (2.25)	6.1 (2.49)	6.2 (2.70)	0.7376	0.8490	0.9231	0.6921
Caregiver SDQ conduct problems	4.1 (2.28)	4.5 (2.63)	3.9 (2.43)	4.6 (2.47)	0.0028	0.4613	0.1185	0.0072
Caregiver SDQ peer problems	3.6 (2.05)	3.4 (2.16)	3.6 (2.14)	3.6 (2.15)	0.3824	0.3560	0.2193	0.9710
Caregiver SDQ impact	4.4 (2.70)	4.5 (2.79)	4.0 (2.64)	4.6 (2.78)	0.0429	0.2476	0.5431	0.0231
Caregiver SDQ externalising	9.5 (4.10)	10.1 (4.61)	9.4 (4.55)	10.5 (4.41)	0.0072	0.5050	0.1672	0.0138
Caregiver SDQ internalising	9.9 (3.73)	9.6 (3.64)	9.7 (3.98)	9.8 (4.07)	0.7721	0.5132	0.5289	0.7833
Caregiver FAD total score	2.2 (0.36)	2.2 (0.35)	2.2 (0.37)	2.3 (0.35)	0.0039	0.4234	0.1456	0.0082
Caregiver FAD roles	2.5 (0.41)	2.6 (0.43)	2.5 (0.42)	2.5 (0.42)	0.0069	0.5409	0.0206	0.1341

continued

TABLE 111 Distribution of 12-month questionnaire response status by participant characteristics and treatment arm with tests for main and interaction effects (*continued*)

Variable	FT: responder? Mean (SD)		TAU: responder? Mean (SD)		Model 1: interaction test[a]		Model 2: main effect test[b]	
	Yes ($N = 261$), *n* (%)	No ($N = 154$), *n* (%)	Yes ($N = 204$), *n* (%)	No ($N = 213$), *n* (%)	Main effect *p*-value	Interaction effect *p*-value	FT *p*-value	TAU *p*-value
Caregiver FAD affective involvement	2.2 (0.45)	2.2 (0.45)	2.2 (0.50)	2.3 (0.51)	0.0321	0.0821	0.7877	0.0034
Caregiver FAD behaviour control	1.8 (0.42)	1.8 (0.40)	1.8 (0.42)	1.9 (0.39)	0.3875	0.8717	0.6252	0.4609
Caregiver Family Questionnaire total score	52.5 (10.32)	53.6 (11.24)	51.4 (10.80)	54.4 (10.71)	0.0064	0.2353	0.2850	0.0048
Caregiver GHQ-12 (Likert)	16.9 (6.77)	19.2 (7.30)	17.9 (7.28)	19.2 (7.17)	0.0003	0.2303	0.0010	0.0815
Caregiver ICU total score	31.9 (10.98)	33.7 (12.51)	31.3 (12.10)	34.5 (10.55)	0.0023	0.3416	0.1363	0.0047

a Response status ≈TRT + characteristic + TRT*characteristic.
b Response status ≈characteristic. Interaction and main effect test via logistic regression (Wald chi-squared tests).
Green cells are used to highlight significant effects at the 10% level.

TABLE 112 Distribution of 18-month questionnaire response status by participant characteristics and treatment arm with tests for main and interaction effects

Variable	FT: responder? Mean (SD)		TAU: responder? Mean (SD)		Model 1: interaction test[a]		Model 2: main effect test[b]	
	Yes ($N = 213$), *n* (%)	No ($N = 202$), *n* (%)	Yes ($N = 182$), *n* (%)	No ($N = 235$), *n* (%)	Main effect *p*-value	Interaction effect *p*-value	FT *p*-value	FT *p*-value
Age	14.3 (1.29)	14.3 (1.48)	14.2 (1.42)	14.5 (1.29)	0.1299	0.2128	0.8465	0.0554
Young person Hopelessness Scale	7.8 (4.30)	7.6 (4.31)	7.4 (4.14)	7.2 (4.24)	0.6111	0.9967	0.7164	0.7218
Young person BSS	11.2 (8.62)	10.5 (9.28)	10.3 (9.27)	10.4 (9.55)	0.6277	0.5638	0.4630	0.9469
Young person CDRS-R	47.9 (13.97)	48.0 (14.45)	48.2 (14.02)	50.4 (12.66)	0.2085	0.2500	0.9368	0.1015
Young person PQ-LES-Q	41.2 (9.23)	41.2 (9.55)	42.0 (9.54)	40.6 (9.35)	0.2965	0.2958	1.0000	0.1414
Young person SDQ total difficulties	19.3 (6.01)	19.8 (5.37)	20.2 (5.40)	20.0 (5.77)	0.6946	0.4419	0.4073	0.7923
Young person SDQ prosocial	7.0 (1.81)	7.3 (1.87)	7.3 (1.90)	7.1 (1.85)	0.7799	0.1823	0.2565	0.4537
Young person SDQ impact	3.2 (2.28)	3.4 (2.54)	3.6 (2.47)	3.5 (2.56)	0.8034	0.3524	0.4140	0.6229
Young person FAD total score	2.4 (0.33)	2.5 (0.33)	2.4 (0.36)	2.5 (0.35)	0.1048	0.3383	0.6483	0.0601
Young person ICU total score	28.1 (8.89)	28.4 (9.33)	28.0 (9.26)	28.9 (8.96)	0.3487	0.6416	0.7358	0.3268
Caregiver SDQ total difficulties	19.0 (6.56)	19.9 (6.55)	19.3 (6.90)	20.2 (6.76)	0.0604	0.9676	0.1832	0.1851
Caregiver SDQ emotional problems	6.1 (2.41)	6.3 (2.36)	6.2 (2.56)	6.1 (2.63)	0.5938	0.3454	0.3152	0.7623
Caregiver SDQ conduct problems	4.0 (2.40)	4.5 (2.42)	4.0 (2.50)	4.5 (2.42)	0.0061	0.8851	0.0674	0.0400
Caregiver SDQ peer problems	3.6 (2.10)	3.5 (2.08)	3.6 (2.12)	3.7 (2.16)	0.9464	0.3659	0.4965	0.5504
Caregiver SDQ impact	4.4 (2.77)	4.5 (2.70)	4.0 (2.59)	4.6 (2.81)	0.0765	0.1734	0.7694	0.0285
Caregiver SDQ externalising	9.3 (4.22)	10.1 (4.36)	9.5 (4.64)	10.4 (4.37)	0.0068	0.9349	0.0535	0.0579
Caregiver SDQ internalising	9.7 (3.78)	9.8 (3.61)	9.8 (4.03)	9.8 (4.02)	0.7796	0.9094	0.7888	0.9026
Caregiver FAD total score	2.2 (0.36)	2.2 (0.35)	2.2 (0.37)	2.2 (0.35)	0.0190	0.9633	0.1048	0.0900

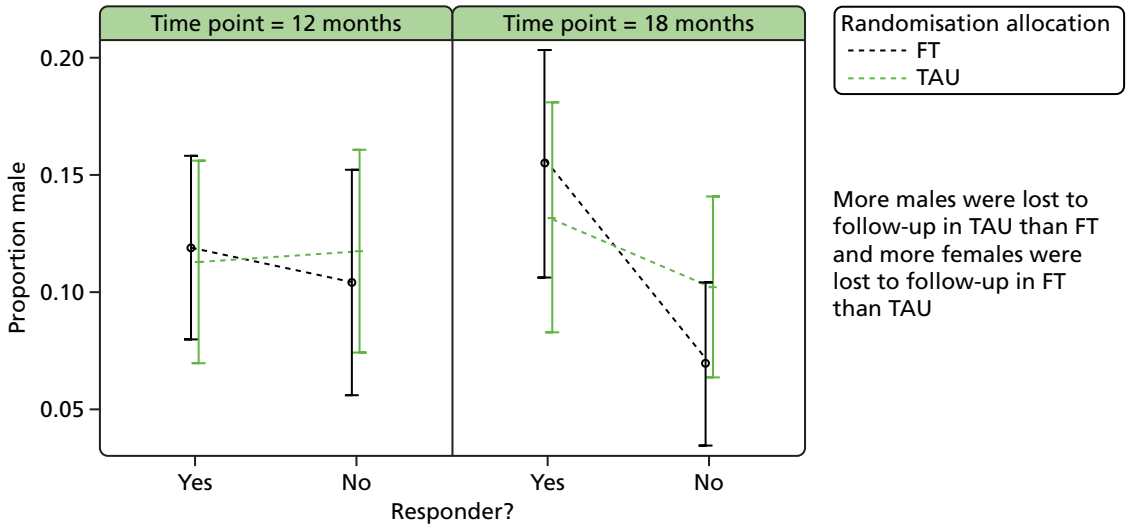
continued

TABLE 112 Distribution of 18-month questionnaire response status by participant characteristics and treatment arm with tests for main and interaction effects (*continued*)

Variable	FT: responder? Mean (SD)		TAU: responder? Mean (SD)		Model 1: interaction test[a]		Model 2: main effect test[b]	
	Yes (*N* = 213), *n* (%)	No (*N* = 202), *n* (%)	Yes (*N* = 182), *n* (%)	No (*N* = 235), *n* (%)	Main effect *p*-value	Interaction effect *p*-value	FT *p*-value	FT *p*-value
Caregiver FAD roles	2.5 (0.41)	2.6 (0.43)	2.5 (0.42)	2.5 (0.41)	0.0047	0.5262	0.0143	0.1223
Caregiver FAD affective involvement	2.2 (0.45)	2.3 (0.45)	2.2 (0.49)	2.3 (0.52)	0.0037	0.2644	0.2256	0.0030
Caregiver FAD behaviour control	1.8 (0.42)	1.8 (0.39)	1.8 (0.43)	1.9 (0.39)	0.1784	0.7043	0.2222	0.4948
Caregiver Family Questionnaire total score	52.4 (10.38)	53.4 (10.97)	51.2 (10.57)	54.3 (10.89)	0.0058	0.1857	0.3084	0.0040
Caregiver GHQ-12 (Likert)	17.2 (6.97)	18.3 (7.12)	18.0 (6.99)	19.1 (7.41)	0.0227	0.9004	0.0920	0.1244
Caregiver ICU total score	31.7 (11.18)	33.5 (11.97)	31.1 (12.22)	34.4 (10.57)	0.0016	0.2978	0.1273	0.0037

a Response status ≈TRT + characteristic + TRT*characteristic.
b Response status ≈characteristic. Interaction and main effect test via logistic regression (Wald chi-squared tests).
Green cells are used to highlight significant effects at the 10% level.

(a) Gender by questionnaire response and treatment



(b) Mean age (95% CI) by questionnaire response and treatment

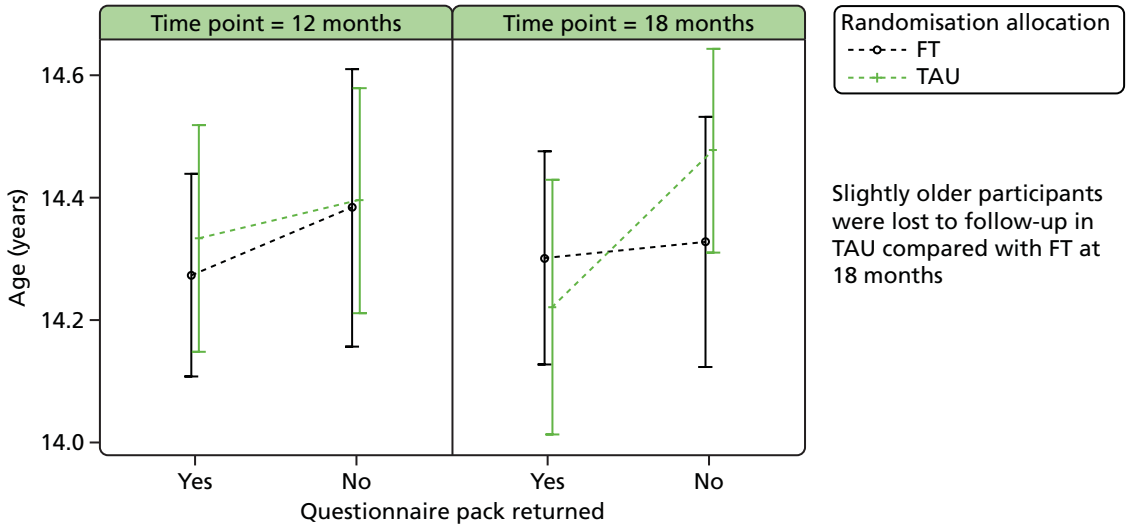
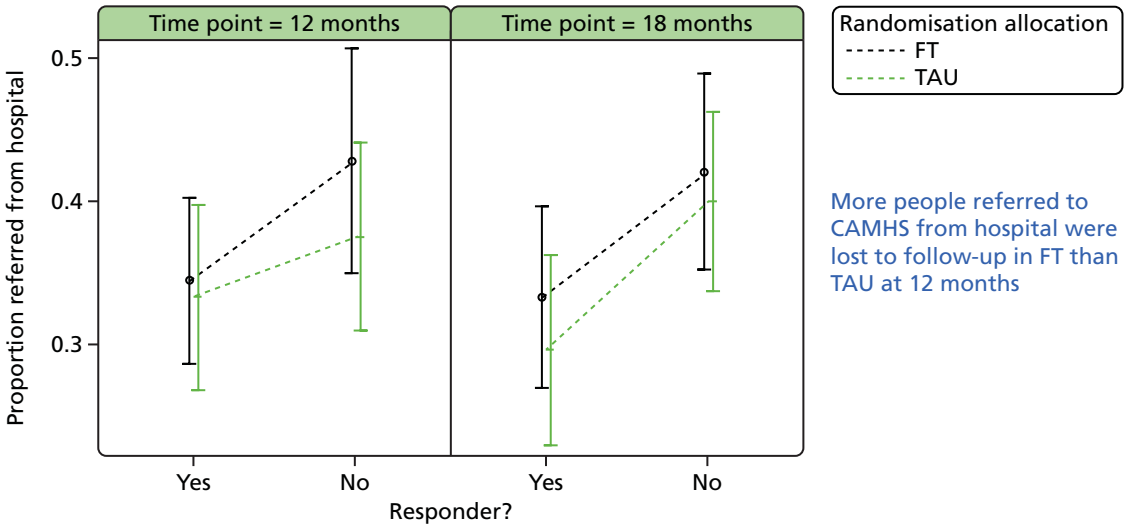


FIGURE 66 Characteristics of questionnaire responders and non-responders by arm at 12 and 18 months, where some differences (main effect in each arm and/or interaction) were detected. CG, caregiver; YP, young person. Black text, particularly concerning cases not indicated; blue text, particularly concerning cases lost to follow up in FT/particularly concerning cases followed up in TAU; green text, particularly concerning cases lost to follow up in TAU/particularly concerning cases followed up in FT. (*continued*)

(c) Referral via hospital by questionnaire response and treatment



(d) Mean caregiver FAD roles (95% CI) by questionnaire response and treatment

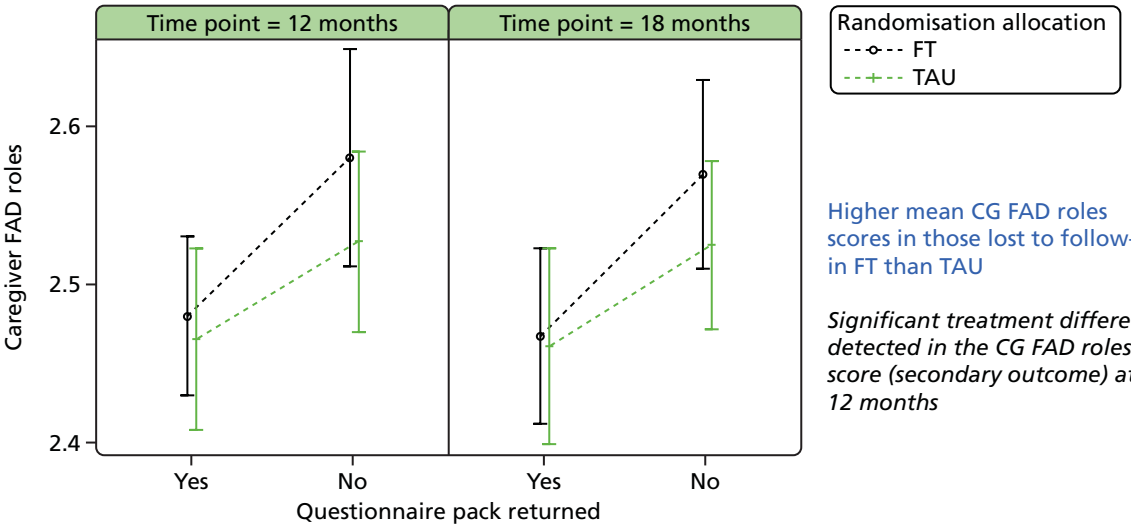
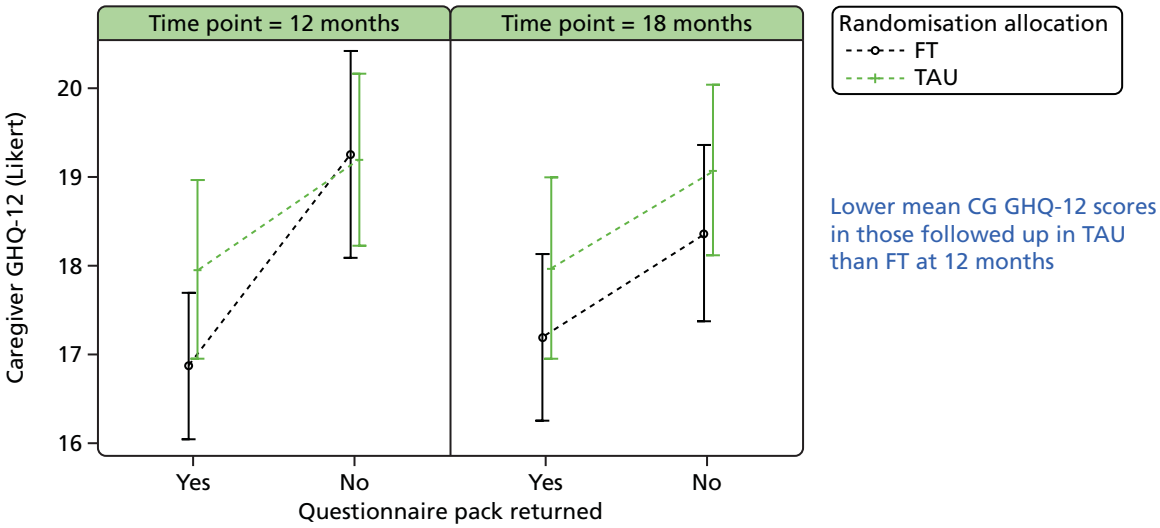


FIGURE 66 Characteristics of questionnaire responders and non-responders by arm at 12 and 18 months, where some differences (main effect in each arm and/or interaction) were detected. CG, caregiver; YP, young person. Black text, particularly concerning cases not indicated; blue text, particularly concerning cases lost to follow up in FT/particularly concerning cases followed up in TAU; green text, particularly concerning cases lost to follow up in TAU/particularly concerning cases followed up in FT. (*continued*)

(e) Mean caregiver GHQ-12 (Likert) (95% CI) by questionnaire response and treatment



(f) Self-harm during follow-up by questionnaire response and treatment

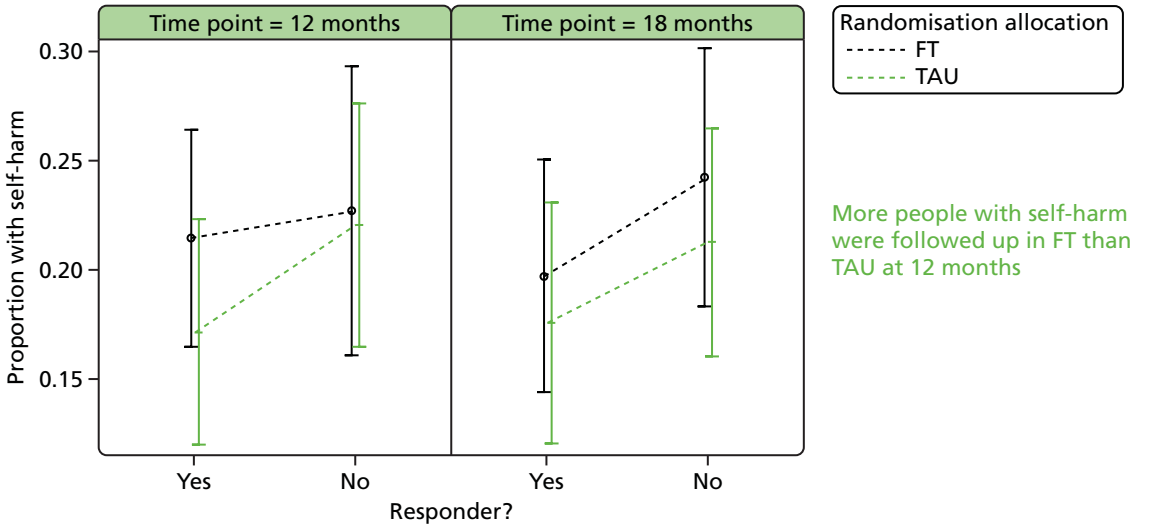
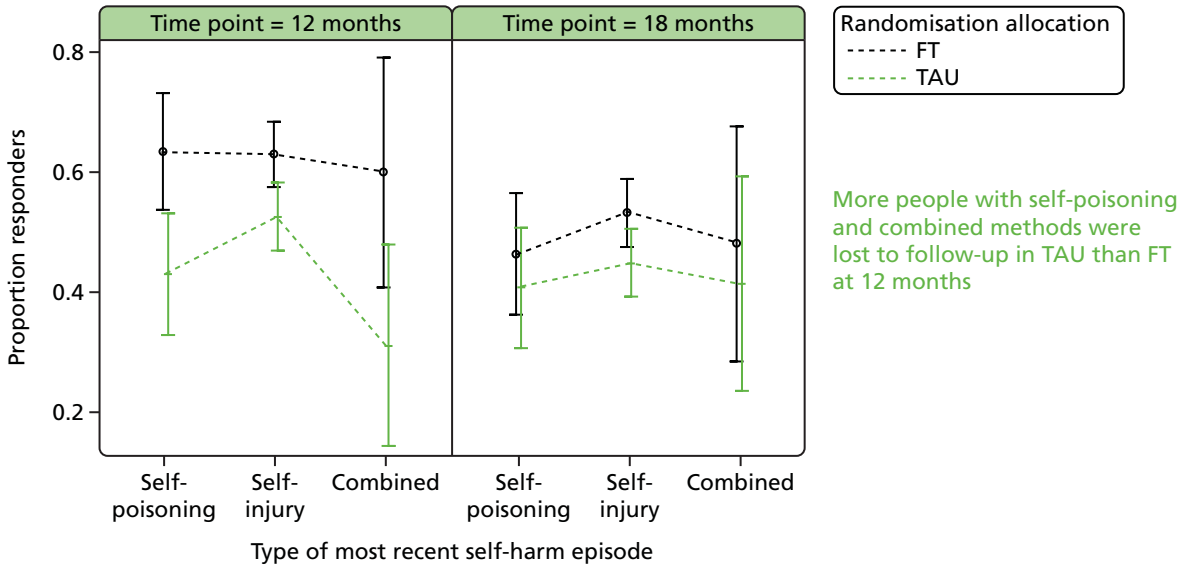


FIGURE 66 Characteristics of questionnaire responders and non-responders by arm at 12 and 18 months, where some differences (main effect in each arm and/or interaction) were detected. CG, caregiver; YP, young person. Black text, particularly concerning cases not indicated; blue text, particularly concerning cases lost to follow up in FT/particularly concerning cases followed up in TAU; green text, particularly concerning cases lost to follow up in TAU/particularly concerning cases followed up in FT. (*continued*)

(g) Questionnaire response by type of index self-harm and treatment



(h) Mean young person FAD total score (95% CI) by questionnaire response and treatment

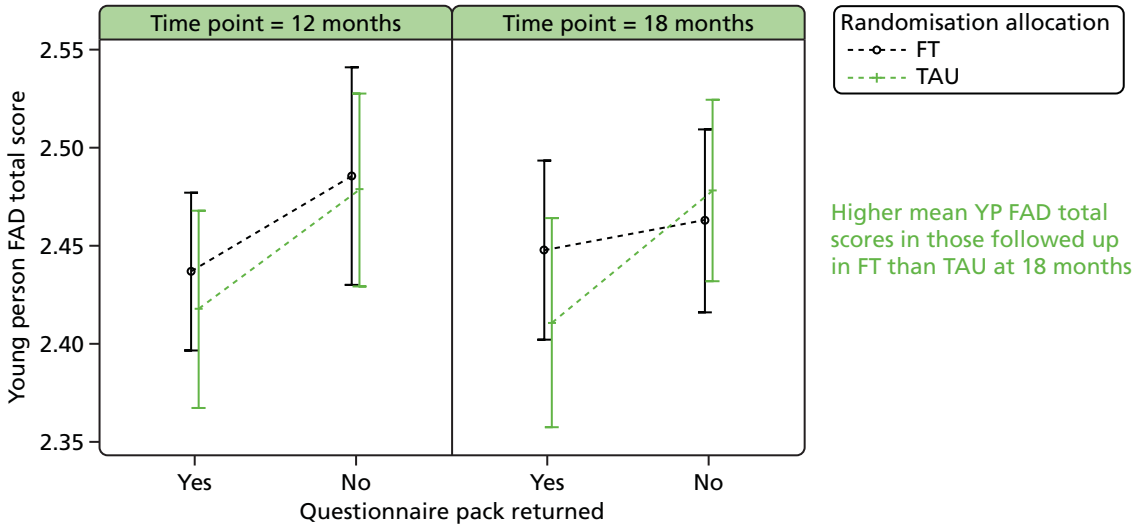
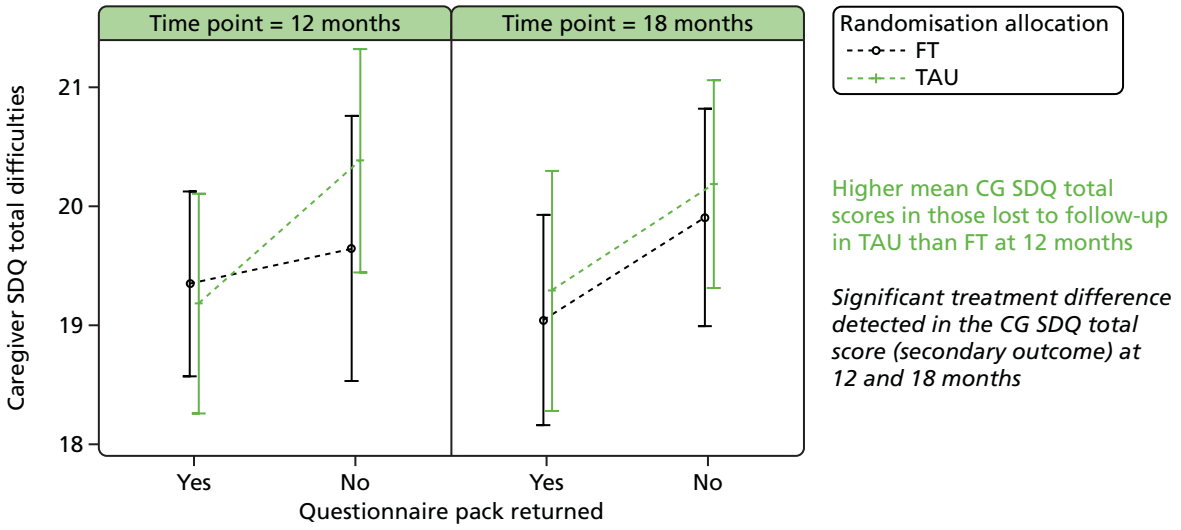


FIGURE 66 Characteristics of questionnaire responders and non-responders by arm at 12 and 18 months, where some differences (main effect in each arm and/or interaction) were detected. CG, caregiver; YP, young person. Black text, particularly concerning cases not indicated; blue text, particularly concerning cases lost to follow up in FT/particularly concerning cases followed up in TAU; green text, particularly concerning cases lost to follow up in TAU/particularly concerning cases followed up in FT. (*continued*)

(i) Mean caregiver SDQ total difficulties (95% CI) by questionnaire response and treatment



(j) Mean caregiver SDQ conduct problems (95% CI) by questionnaire response and treatment

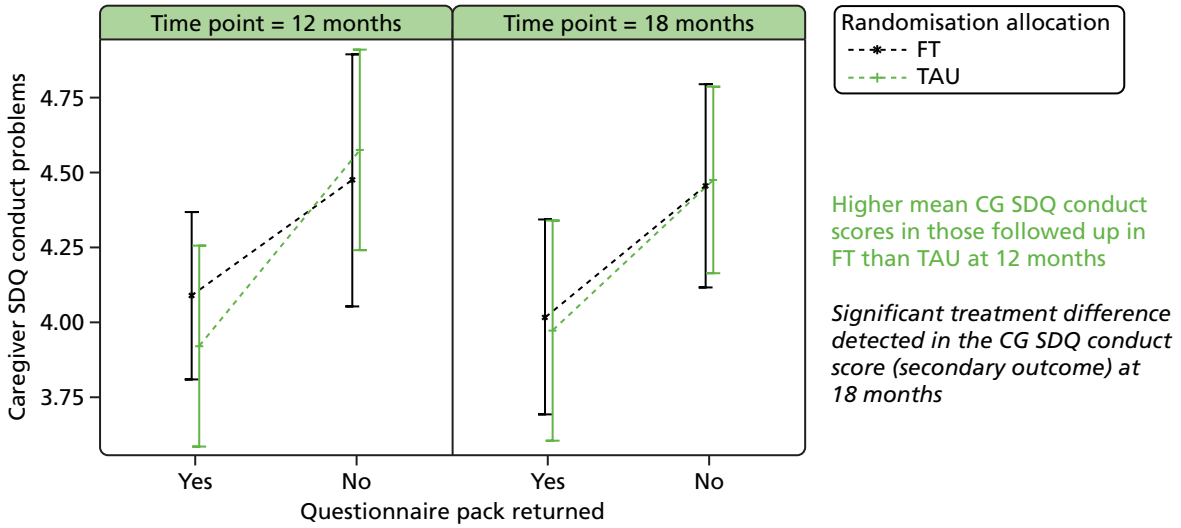
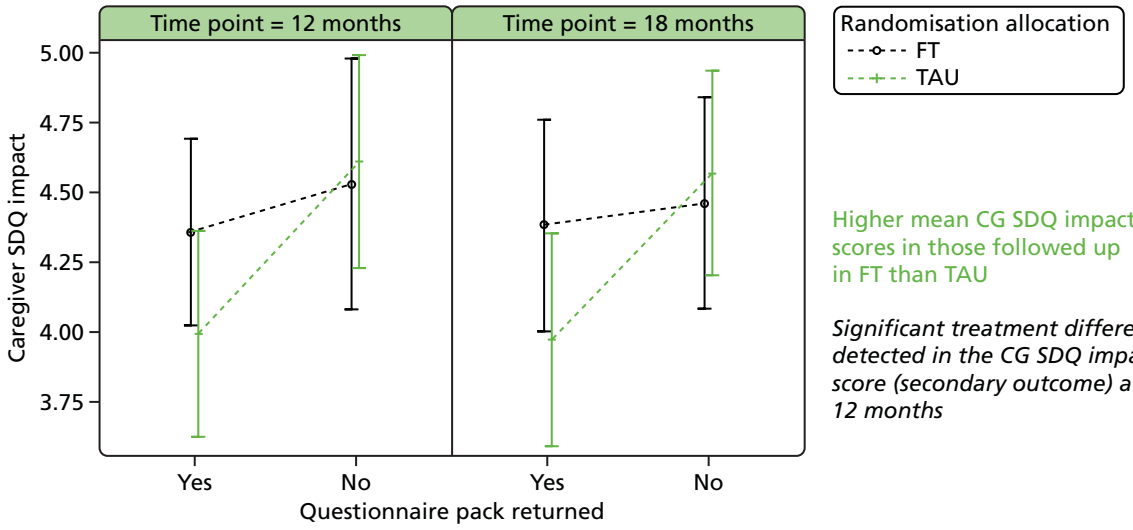


FIGURE 66 Characteristics of questionnaire responders and non-responders by arm at 12 and 18 months, where some differences (main effect in each arm and/or interaction) were detected. CG, caregiver; YP, young person. Black text, particularly concerning cases not indicated; blue text, particularly concerning cases lost to follow up in FT/particularly concerning cases followed up in TAU; green text, particularly concerning cases lost to follow up in TAU/particularly concerning cases followed up in FT. (*continued*)

(k) Mean caregiver SDQ impact (95% CI) by questionnaire response and treatment



(l) Mean caregiver SDQ externalising (95% CI) by questionnaire response and treatment

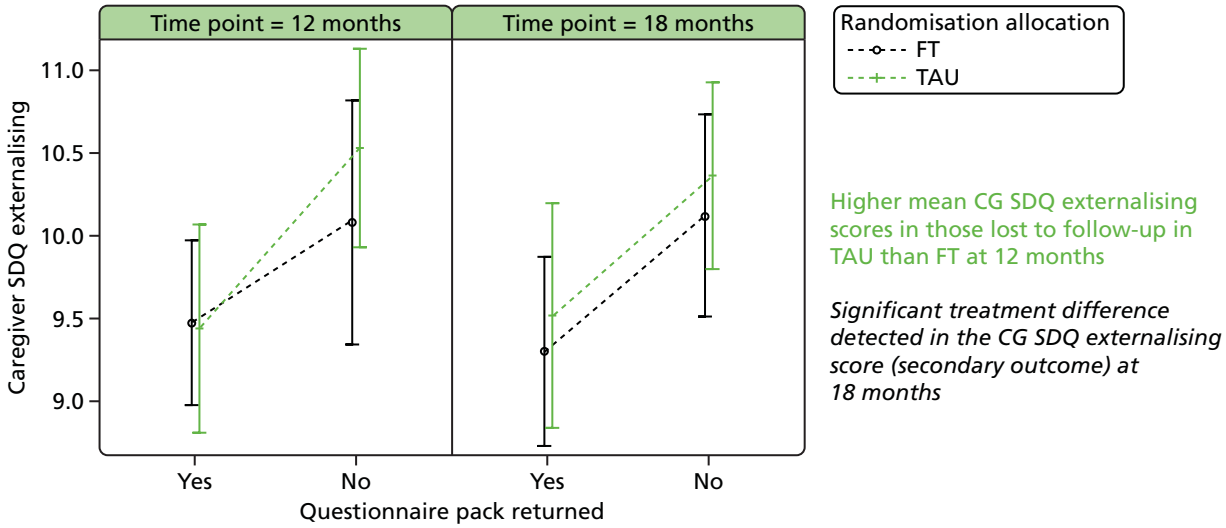
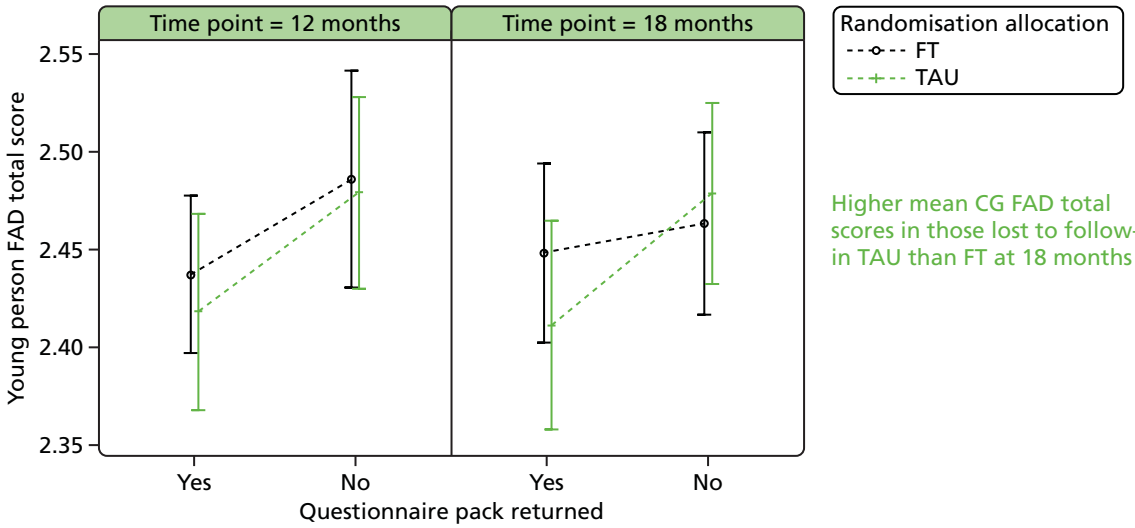


FIGURE 66 Characteristics of questionnaire responders and non-responders by arm at 12 and 18 months, where some differences (main effect in each arm and/or interaction) were detected. CG, caregiver; YP, young person. Black text, particularly concerning cases not indicated; blue text, particularly concerning cases lost to follow up in FT/particularly concerning cases followed up in TAU; green text, particularly concerning cases lost to follow up in TAU/particularly concerning cases followed up in FT. (*continued*)

(m) Mean caregiver FAD total score (95% CI) by questionnaire response and treatment



(n) Mean caregiver FAD affective involvement (95% CI) by questionnaire response and treatment

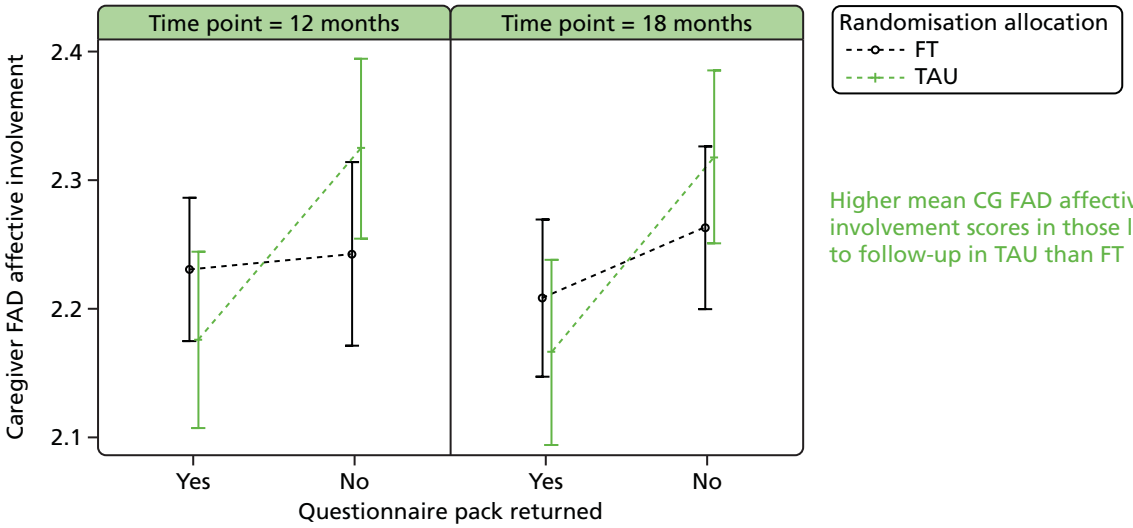
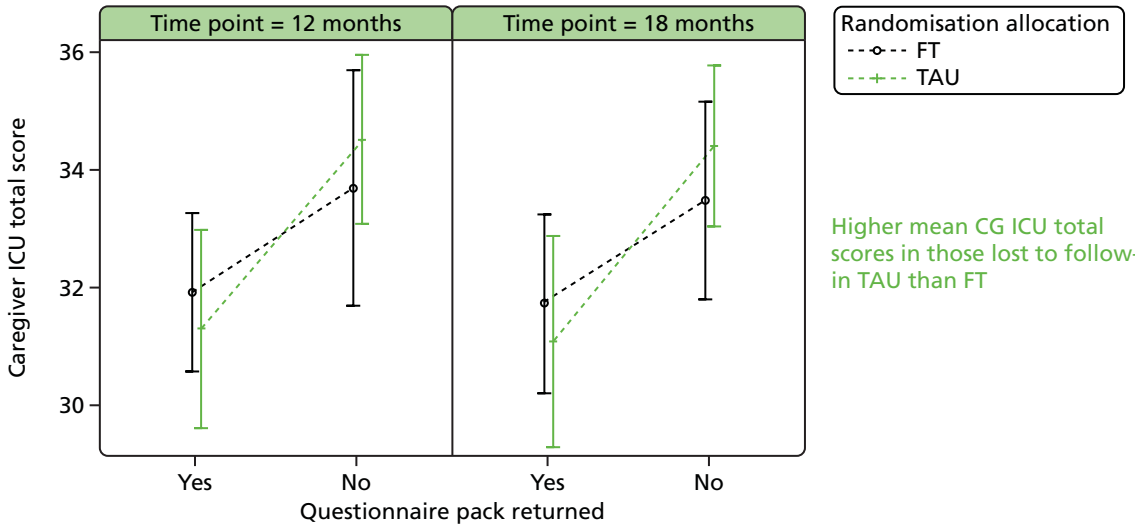


FIGURE 66 Characteristics of questionnaire responders and non-responders by arm at 12 and 18 months, where some differences (main effect in each arm and/or interaction) were detected. CG, caregiver; YP, young person. Black text, particularly concerning cases not indicated; blue text, particularly concerning cases lost to follow up in FT/particularly concerning cases followed up in TAU; green text, particularly concerning cases lost to follow up in TAU/particularly concerning cases followed up in FT. (*continued*)

(o) Mean caregiver ICU total score (95% CI) by questionnaire response and treatment;



(p) Mean caregiver Family Questionnaire total score (95% CI) by questionnaire response and treatment

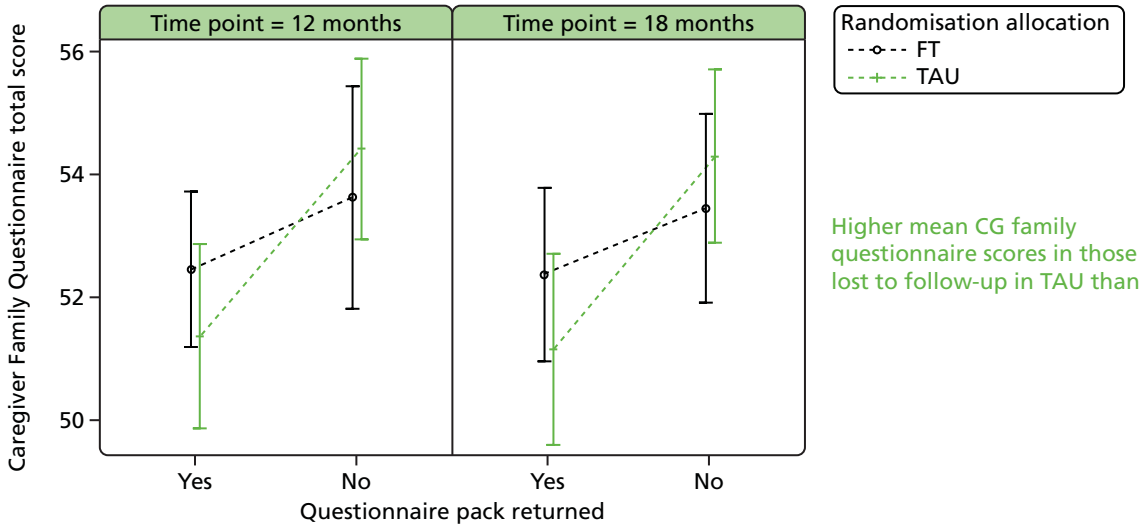


FIGURE 66 Characteristics of questionnaire responders and non-responders by arm at 12 and 18 months, where some differences (main effect in each arm and/or interaction) were detected. CG, caregiver; YP, young person. Black text, particularly concerning cases not indicated; blue text, particularly concerning cases lost to follow up in FT/particularly concerning cases followed up in TAU; green text, particularly concerning cases lost to follow up in TAU/particularly concerning cases followed up in FT.

Appendix 7 Health economics analysis: additional information

Figure 67 shows the cost-effectiveness plane for FT compared with TAU based on 10,000 bootstrapped estimates of costs and aggregate QALYs. The average costs from the bootstrapped estimates were £3751.21 (SD £197.67) and £4942.7 (SD £197.8) for the TAU and FT arms, respectively. The corresponding mean (aggregate) QALYs were 2.236 (SD 0.020) for the TAU arm and 2.288 (SD 0.020) for the FT arm. The simulation estimates are above the x-axis, showing that FT is always more costly than TAU. As with the main analysis, most estimates are spread in the north-east quadrant, suggesting that FT will probably lead to better health outcomes when caregivers' QALY gains are taken into account. However, some estimates are also in the north-west quadrant, where the FT arm is likely to have lower QALYs than the TAU arm.

The cost-effectiveness acceptability curve of FT compared with TAU when caregivers' quality of life is taken into account is presented in *Figure 68*. At a threshold of £20,000, FT has a 41% chance of being cost-effective and this percentage increases to 64% when the threshold is £30,000.

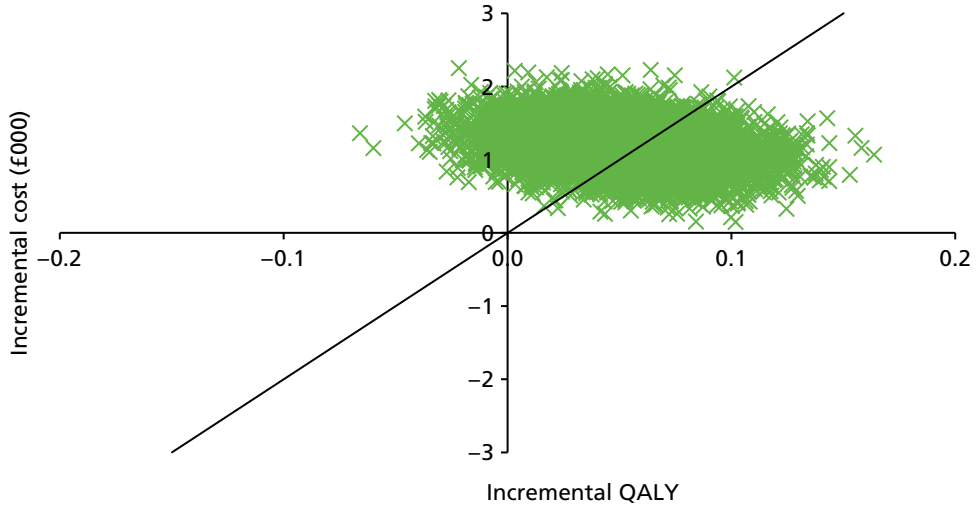


FIGURE 67 Cost-effectiveness plane of FT compared with TAU (outcome measure: aggregate QALY, NHS perspective).

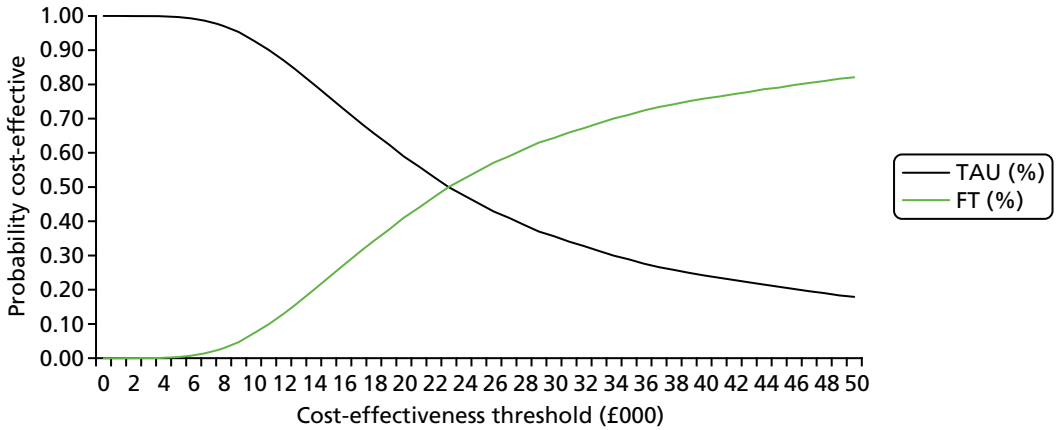


FIGURE 68 Cost-effectiveness acceptability curve of FT compared with TAU (outcome measure: aggregate QALY, NHS perspective).

Appendix 8 End-of-trial information sheet for participants

What were you trying to do?

The aim of the SHIFT study was to find out more about how to help young people who have self-harmed, and in particular to see what could be done to support them not self-harming again. By self-harm we meant taking an overdose, cutting or some other form of hurting yourself. We wanted to look at how young people got on after family therapy compared to usual treatment in CAMHS (child and adolescent mental health services).

What was the main finding?

We found that those who had family therapy went to hospital after self-harm about the same number of times as those who had treatment as usual. This means that there was no real difference between the treatments; they were about as good as each other.

What did you do?

We worked with around 40 child and adolescent mental health teams in Yorkshire, Manchester and London. We invited over 800 young people aged between 11 and 17 years old who had self-harmed at least twice to take part in the research. As well as the young person, their main caregiver (for example, Mum or Dad) also took part. If they both agreed they were then randomly allocated to either family therapy or the usual treatment in their local CAMHS.

We collected lots of information about the young person and their family at the beginning, before treatment started. Treatment lasted around 6–8 sessions although some people had more and some less. We collected more information over the next 18 months including whether the young person had self-harmed again and had to go to hospital.

Did you find out anything else?

Yes, we found that:

- Young people who were older (aged 15–17), who combined an overdose with cutting, and who went to hospital rather than directly to CAMHS were more likely to self-harm again.
- Those who had been in the family therapy group reported improvements in sociability (kindness and caring for others), and their caregivers reported improvements in their levels of emotional and behavioural problems.
- There was some evidence to suggest that family therapy might be more effective in reducing self-harm where caregivers reported poor family functioning at the start of the study (particularly in relation to talking about feelings). Family therapy also seemed to be helpful where young people reported that they found it easy to discuss their emotions.
- On the other hand, where the young person themselves reported difficulty in expressing emotion, or caregivers reported that they already discussed emotion quite well, family therapy was not as effective as usual treatment.

We also looked at the cost-effectiveness (value for money) of the two treatments. There was no evidence that family therapy was cost-effective when compared with treatment as usual when we focused on just the young people who took part. However, there was a suggestion that family therapy may be cost-effective if the health benefits to caregivers are also taken into account.

What are you going to do next?

We collected lots of information from our participants and there is still more work to do analysing all the information and understanding more about self-harm in general.

We have also received funding from the National Institute for Health Research (the part of the government's Department of Health that funds research) to extend the study and see if any differences between the family therapy and treatment as usual groups can be seen if we look again at the end of 2016. To do this we will be collecting data directly from NHS Digital (where data is held from all the hospitals in England about people's visits to hospital). You will have received an earlier information sheet explaining a bit more about this, but if you have any questions please contact the research team (see below). Please note that we will not collect any more information if you have already asked us not to.

How can I find out more?

You can look at our web site (***TBC***) or contact the research team directly:

David Cottrell Liz Graham

E-mail: XXXX E-mail: XXXX

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library