Systematic Review and Individual Participant Data Meta-Analysis: Reducing Self-Harm in Adolescents: Pooled Treatment Effects, Study, Treatment and Participant Moderators

Alex Wright-Hughes, MSc,[a] Amanda J. Farrin, MSc,[a] Peter Fonagy, CBE, FMedSci,[b] Dennis Ougrin, PhD,[c] Daniel Stahl, PhD,[d] Judy Wright, MSc,[a] Donna Irving, BA,[e] Faraz Mughal, FRCGP,[f] Alex Truscott, BA,[b] Emma Diggins, MA,[a] Andrew Chanen, PhD,[g] Emily Cooney, PhD,[h] Greg Carter, FRANZCP,[i] Kerrie Clover, PhD,[i] Mark Dadds, PhD,[j] Guy Diamond, PhD,[k] Christianne Esposito-Smythers, PhD,[l] Jonathan Green FRCPsych, FMedSc,[m] Helen Griffiths, DClinPsychol,[n] Hossein Hassanian-Moghaddam, MD, FACMT,[o] Simon Hatcher, MD,[p] Philip Hazell, PhD,[q] Nusrat Husein, MD,[m] Michael Kaess, MD,[r,s] Cheryl King, PhD,[t] Britt Morthorst, PhD,[u] Rory C O'Connor, PhD,[v] Pilar Santamarina-Perez, PhD,[w] Peter Tyrer, FMedSci,[x] Rebecca Walwyn, PhD,[a]* David Cottrell, FRCPsych[a]*

[a] The University of Leeds, Leeds, United Kingdom.
[b] University College London, London, United Kingdom.
[c] Queen Mary University of London, London, United Kingdom.
[d] King's College London, London, United Kingdom.
[e] Leeds Trinity University, Leeds, United Kingdom.
[f] Keele University, Keele, United Kingdom.
[g] The University of Melbourne, Melbourne, Australia.
[h] University of Otago, Wellington, New Zealand.
[i] University of Newcastle, New South Wales, Australia.
[j] University of Sydney, Sydney, Australia.
[k] ABFT International Training Institute, Plantation, Florida.
[l] George Mason University, Fairfax, Virginia.
[m] University of Manchester, Manchester, United Kingdom.
[n] University of Edinburgh, Edinburgh United Kingdom.
[o] At the time of the study: Shahid Beheshti University of Medical Sciences, Tehran, Iran.
[p] The University of Ottawa, Ontario, Canada.
[q] The University of Sydney School of Medicine, Sydney, Australia.
[r] University of Bern, Bern, Switzerland.
[s] University of Heidelberg, Heidelberg, Germany.
[t] University of Michigan, Ann Arbor, Michigan.
[u] University of Copenhagen, Copenhagen, Denmark.
[v] University of Glasgow, Glasgow, United Kingdom.
[w] Hospital Clinic Barcelona, Barcelona, Spain.
[x] Imperial College London, London, United Kingdom.

*Rebecca Walwyn and David Cottrell are joint last authors of this work.

Correspondence to Alex Wright-Hughes, MSc, Complex Interventions Division, Clinical Trial Research Unit, Leeds Institute of Clinical Trials Research, School of Medicine, University of Leeds, UK; e-mail: a.wright-hughes@leeds.ac.uk

ABSTRACT

Objective

Self-harm is common in adolescents and a major public health concern. Evidence for effective interventions that stop repetition is lacking. This individual-participant-data (IPD) meta-analysis of randomised controlled trials (RCTs) aimed to provide robust estimates of therapeutic intervention effects and explore which treatments are best suited to different subgroups.

Method

We searched databases and trial registers, to January-2022. RCTs compared therapeutic intervention to control, targeted adolescents aged 11-18 with a history of self-harm and receiving clinical care and reported on outcomes related to self-harm or suicide attempt. Primary outcome was repetition of self-harm at 12 months post-randomization . Two-stage random-effects IPD meta-analyses were conducted overall and by intervention. Secondary analyses incorporated aggregate data (AD) from RCTs without IPD. PROSPERO registration: CRD42019152119.

Results

We identified 39 eligible studies; 26 provided IPD (3,448 participants), 7 provided AD (698 participants). There was no evidence that intervention/s were more or less effective than controls at preventing repeat self-harm by 12 months in IPD (odds ratio (OR)=1.06 [95% CI 0.86, 1.31], studies=20, n=2,949) or IPD+AD (OR=1.02 [95% CI 0.82, 1.27], studies=22, n=3,117) meta-analyses and no evidence of heterogeneity of treatment effects on study and treatment factors. Across all interventions, participants with multiple prior self-harm episodes showed evidence of improved treatment effect on self-harm repetition 6-12 months after randomization (OR=0.33 [95% CI 0.12, 0.94], studies=9, n=1,771).

Conclusion

This large-scale meta-analysis of RCTs provided no evidence that therapeutic intervention was more, or less, effective than control for reducing repeat self-harm. We observed evidence indicating more effective interventions within youth with two or more self-harm incidents. Funders and researchers need to agree on a core set of outcome measures to include in subsequent studies.

Clinical trial registration information: Reducing self-harm in adolescents: an individual participant data meta-analysis; https://www.crd.york.ac.uk/; CRD42019152119.

Key words: self-harm; suicide; adolescents; systematic review; ipd meta-analysis

INTRODUCTION

Self-harm in adolescents is a major public health concern in the UK and globally,[1] with lifetime prevalence of 16.9%, and increasing rates.[2] Following self-harm, suicide attempt is common[3] and is the second commonest cause of death in 10-24-year-olds.[4] All-cause mortality shows a four-fold increase, and death by suicide, a ten-fold increase[5] following self-harm. Non-fatal repetition of self-harm is common with a one-year hospital re-attendance rate of 18%.[6]

Interpretation of the literature is complicated by the lack of clarity regarding definitions of self-harm. Harris et al[7] conducted a meta-analysis of self-injurious thoughts and behaviours (SITB) in which definitions of self-harm were based on the presence or absence of suicidal intent. Intentional self-directed harm without suicidal intent was considered as non-suicidal self-injury (NSSI), suicide-related cognitions and plans as suicidal ideation, self-directed harm with intent to die as a suicide attempt, and where no information regarding suicidal intent is provided, as self-harm. However, determining suicidal intent is not straightforward. Classifying self-harm by suicidal intent and method, for example NSSI or not, may be clinically misleading given so many hospital attending patients switch method.[8]

In this study we have used the generic term 'self-harm' defined by the UK National Institute for Health Care Excellence (NICE) as any form of non-fatal self-poisoning or self-injury (including cutting, taking excess medication, attempted hanging, self-strangulation, jumping from height, running into traffic), regardless of suicidal intent.[9] It thus includes NSSI as defined by Harris et al[7] but excludes suicidal ideation.

Effective interventions to prevent repeat self-harm have not yet been identified despite several published studies, systematic reviews, and meta-analyses. The most recent Cochrane Review in 2021[10] used tight eligibility criteria and included only randomised controlled trials (RCTs) comparing specific interventions with a comparator, in children and adolescents with a recent (within six months of trial entry) episode of self-harm resulting in presentation to hospital or clinical services. They identified only 17 studies (2280 participants) and found "only uncertain evidence regarding a number of psychosocial interventions in children and adolescents who engage in SH", whilst suggesting further evaluation of dialectical behaviour therapy (DBT-A) was warranted. Kothgassner et al's 2020 review[11] used similar eligibility criteria but past self-harm was only necessary at some point in the past, not within 6 months as with Cochrane. They identified 25 studies with 2,962 participants. In both cases around a third of participants were from one study.[12] The Kothgassner review concluded that therapeutic interventions overall fared slightly better than active controls but could not identify any one specific intervention although there were small positive effects for DBT-A and family centred therapy. The most recent UK NICE guideline[13], using similar eligibility criteria also recommended 'consideration' of DBT-A for children and young people.

Harris et al[7] took a different approach with much wider definitions of self-harm. They included studies that explored only impact on suicidal ideation, and also studies where the primary target of intervention was a mental disorder such as depression but where self-harm had been measured as an outcome. However, this approach produced similar findings with "Nearly all interventions produced nonsignificant reductions in SITBs."

In adults, the picture is similar. The most recent Cochrane review[14] found "only uncertain evidence regarding a number of psychosocial interventions for adults who engage in self-

harm." Fox et al[15] found only small intervention effects "no intervention appeared significantly and consistently stronger than others".

Reviews in both adults and children highlight the relatively low number of high quality RCTs alongside the generally poor quality of many of the studies reviewed despite the significant scale of the problem and the lack of recognised effective interventions. One plausible hypothesis for the lack of effective interventions is that samples aggregate young people who have self-harmed for different reasons and apply a 'one size fits all' intervention. Identification of relevant sub-groups within those who self-harm may offer a more promising path forward.

An individual participant data (IPD) meta-analysis, by re-analysing pooled data from eligible studies, can provide more robust estimates of the effects of therapeutic interventions for self-harm than conventional meta-analyses that rely on aggregated data (AD) and reported analyses.[16] IPD allows for the inclusion of subsets of participants, thereby increasing the number of studies and participants contributing to analyses, which enhances the power to detect effects. In many of the reviews described above, good quality RCTs had to be excluded because authors could not disaggregate data when not all participants met the review inclusion criteria, for example, such as by age or type of self-harm. Factors at the participant, treatment, and study levels may all influence the effectiveness and outcomes of interventions.[10,11,17] IPD enables the analyses of the potential moderating effects of participant-level factors on outcomes, which is often limited in AD meta-analyses.

We hypothesised that an IPD meta-analysis would allow us to retain the stringent eligibility criteria of reviews like Cochrane, while including significantly more studies and participants. In cases where authors of eligible studies were unable to share data, we could still incorporate

published AD on eligible participants, as in standard meta-analysis, alongside our IPD analyses. This approach would not only provide more accurate information about the effectiveness of interventions to prevent self-harm but also allow us to explore study and treatment moderators. Additionally, it would enable us to investigate participant-level moderators of treatment effects, potentially identifying subgroups of adolescents who might benefit from specific therapeutic interventions for self-harm.

The protocol[18] and a description of the search methods, how we accessed IPD, risk of bias assessment and outputs are published elsewhere.[19]

METHOD

This study is registered (PROSPERO-CRD42019152119),[20] and the report follows PRISMA-IPD reporting guidelines (Supplement 1, available online).[21] The study was conducted following a successful bid for a commissioned call from National Institute for Health and Care Research, which specified the need to conduct an IPD meta-analysis.

Eligibility Criteria and Search Strategy

Self-harm was defined as any form of non-fatal self-poisoning or self-injury (including cutting, taking excess medication, attempted hanging, self-strangulation, jumping from height, running into traffic), regardless of suicidal intent.[9] This includes definitions of non-suicidal self-injury commonly used by US researchers and suicidal behaviour, where lack of intent is assumed by reference to the method of self-harm. Self-harm could be self-reported, reported via interview or identified through hospital or medical records.

The inclusion criteria were:

1) RCTs comparing any intervention delivered in outpatient/community settings aiming to reduce subsequent self-harm against any control (studies where the intervention primarily targeted underlying pathology, for example depression, and self-harm was only measured as a secondary outcome, were excluded);
2) Adolescents aged 11–18 who had (i) self-harmed at least once prior to randomization and (ii) presented to clinical services for self-harm (includes suicide attempt but excludes suicidal ideation without explicit self-harm);
3) Data collected relating to self-harm or suicide attempt.

We incorporated studies where only a subset of participants met our criteria if IPD could be obtained for them: (i) aged 11–18 and (ii) having self-harmed at least once before randomization . We excluded studies of intensive inpatient and prevention-based interventions not targeted specifically at adolescents who presented to clinical services with self-harm, and studies with less than 20 eligible participants.

We identified eligible studies (Supplement 2, available online) via:

1) A search for systematic reviews of self-harm in adolescents conducted in June 2019;
2) Searches for recent publications from 2015 (the date of the last comprehensive search reported in the most recent systematic review) to August-2019, updated in February-2021 and January-2022. This search included unpublished trials (no date restriction), and ongoing RCTs.

Our literature search is described in full elsewhere[19].

Two authors (AWH, DC) independently reviewed titles, abstracts and full texts in Covidence. Disagreements were resolved by discussion, and where necessary adjudicated by a further author (RW).

Interventions

Therapeutic interventions were grouped by consensus (DC, DO, PF), according to study published descriptions, theoretical underpinnings, supplementary material and manuals. The categories were: cognitive–behavioural therapy (CBT); dialectical behaviour therapy (DBT); family therapy; group therapy; mentalisation based, psychodynamic, cognitive analytic therapy (MBT/CAT); multi-systemic therapy (MST); problem solving, psychoeducation, support (PST); postcards, tokens, documents (postcards/tokens); other single session, brief-interventions. Control treatments were treatment as usual (TAU), enhanced TAU, or active control.

Outcome Measures

The primary outcome was repetition of self-harm, from randomization to the last available follow-up within 3, 6, 12, 18 and 24-months post-randomization . The primary time-period was 12-months, including studies where the follow-up assessment of self-harm was >6 and ≤ 12 months.

Secondary outcomes were time to repetition of self-harm; pattern of self-harm repetition between 6-12 months, 12-18 months, 18-24 months post-randomization ; general psychopathology (aggregated symptom scores of any mental disorder); depression; and suicidal ideation.

Additional outcomes (descriptive analysis only): quality of life, and death of adolescent.

Follow-up was grouped into short-term (up to 3-months), and 6 (>3 to ≤6), 12 (>6 to ≤12), 18 (>12 to ≤18), 24 (>18 to ≤24), and ≥24-months post-randomization .

Data Collection

We sought IPD, including baseline participant demographics and clinical data, details of therapeutic intervention, and outcomes, prioritising the primary outcome. More detail about how IPD were accessed is available elsewhere.[19] Where IPD were not available, AD were extracted from study reports and publications, where possible, by AWH (verified by DS). Study-level variables relating to study conduct and design, methodology and clinical factors were extracted from publications by AWH (verified and categorised by DC).

Risk of Bias (ROB) in Individual Studies

Pairs of authors (DC, FM, AT, ED) independently assessed the quality of included studies using version 2 of the Cochrane risk-of-bias tool for randomized trials (RoB2):[22] providing up to two RoB2 assessments for each study, for each method of data collection. Disagreements were resolved through discussion and consultation with a third reviewer (AWH).

Statistical Analysis

Data were analysed using SAS 9.4[23] and STATA 17.[24] Analyses were based on intention-to-treat, including all participants as randomised, regardless of withdrawal or protocol compliance. Primary analyses were based on available data, conducted separately for each time-period, overall and by therapeutic intervention.

Treatment effects are expressed as odds ratios (OR), hazard ratios (HR) and standardised mean differences (SMD). ORs of <0.59 or >1.68, <0.29 or >3.47 and <0.15 or>6.71[25], and absolute SMDs of 0.2, 0.5 and 0.8 are used to describe small, moderate, and large (positive or negative) effects[26]. For repetition of self-harm outcomes, ORs<1 represent reduced odds of self-harm in the intervention compared to control, whereas ORs>1 represent increased odds of self-harm. SMDs<0 represent a reduction in general psychopathology, depression and suicidal ideation symptoms in the intervention compared to control, whilst SMDs>0 represent increased symptoms. Where outcomes or baseline moderators comprised continuous data from different scales, scores were standardised in each study and time-period. This was done using the mean and pooled standard deviation (SD), applying an approximate Hedges g adjustment, before conducting the meta-analysis.

Further details of analysis and changes from the protocol are in Supplement 2, available online.

Pooled Treatment Effects

We used a two-stage approach,[27] estimating treatment effects in each study separately, then pooling aggregate results. Step-one (SAS 9.4[23]) used logistic, linear and Cox proportional hazards regression, adjusted for age (continuous) and sex at birth as appropriate to the outcome. To address the bias caused by rare events in logistic regression, Firth's Penalized Likelihood[28] was applied. Analyses accounted for cluster randomization , using multilevel mixed-effects regression with a random cluster effect. Step-two (STATA 17.[24]) used random-effects meta-analysis allowing for statistical heterogeneity (variability in the intervention effects between studies being evaluated) which may be explained by clinical and methodological diversity of studies. Estimation used restricted maximum likelihood (REML),[29] and confidence intervals (CIs) were derived using the Hartung-Knapp-Sidik-Jonkmann[30] approach to allow for

uncertainty in variance estimates. Confidence intervals are also provided without the Hartung-Knapp-Sidik-Jonkmann adjustment to support comparison with previous meta-analyses for the primary outcome and secondary outcomes at the primary 12-month timepoint.

We report estimated pooled treatment effects, 95% CIs and 95% prediction intervals alongside forest plots with study-specific treatment effect estimates. Statistical heterogeneity, indicating inconsistent treatment effects was assessed using τ^2 (indicating the variability of true effect sizes under a random-effects model) and I^2 (representing the proportion of total variability due to between-study heterogeneity) with Cochran's Q test assessed using a p-value of $p<0.1$ rather than 0.05 as it is widely accepted that the test has poor power when there are few studies.

Secondary analysis incorporated available published AD[31] for studies where IPD were not available. Here, IPD were reduced to AD in step-one without adjustment for baseline covariates and estimates were combined with existing AD from studies without IPD.

Sensitivity analyses were conducted to test the robustness of our conclusions to our analysis methods by imputing missing IPD using multiple imputation, undertaking one-stage IPD meta-analysis, and excluding studies of high risk of bias (for the primary outcome).

Moderating Study and Treatment-Level Effects

Random effects subgroup analysis (categorical variables) and meta-regression (continuous variables) were used to explore sources of between-study heterogeneity on treatment effect estimates, and specifically the impact of i) clinical diversity in participant populations, ii) intervention delivery and iii) methodological diversity of study conduct and design.

We used two-stage IPD+AD meta-analyses to maximise the number of studies included. Analysis was conducted separately for each moderator, across all interventions (due to limited numbers of studies for each intervention). To minimise the potential for spurious findings, analyses were conducted on the primary outcome only, and repeated using secondary outcomes only if evidence of moderating effects was detected to ensure consistent findings across outcomes.

Candidate moderators were:

- full vs partial sample eligible
- pilot/feasibility vs effectiveness design
- study sample size powered vs not
- US vs other (differences in self-harm definition)
- low risk of bias vs some concerns vs high
- self-report data/researcher interview vs hospital/medical records
- control TAU/standard care/assessment vs enhanced TAU/good clinical care vs active
- group element to intervention
- family element to intervention
- low, medium, high intervention intensity
- years since primary publication
- number of eligible participants
- planned treatment duration
- number of planned treatment sessions

Moderating Adolescent-Level Effects

We examined whether treatment effects remained uniform across adolescent moderators. Analyses of predetermined key moderators were conducted on primary and secondary outcomes. Additional moderator analyses were performed on the primary outcome only if the potential moderator were documented in at least half of the trials. These analyses were extended to secondary outcomes only when subgroup effects were observed on the primary outcome, to verify consistency across different outcomes. Results for primary and secondary outcomes related to self-harm repetition are presented together to facilitate comparison.

We used two-stage IPD meta-analyses, extending the modelling approach for the primary analysis of pooled treatment effects. Each adolescent moderator was included as a fixed effect alongside a moderator-by-treatment interaction in step-one. Step-two pooled the moderator-by-treatment interaction estimates, representing the change in the treatment effect for different levels of the moderator (i.e. a 1-year increase in age, male vs female participants)[32]. These analyses considered but did not account for cluster-randomization (Ougrin 2013) since no clustering effects were observed in the primary analysis of pooled treatment effects. We present CIs with and without this Hartung-Knapp-Sidik-Jonkmann adjustment in sensitivity analyses to aid hypothesis generation. Due to the exploratory nature of the analysis, prediction intervals were not calculated.

We present forest plots displaying estimated pooled interaction effects and their 95% confidence intervals (CIs), alongside study-specific estimates and estimates by categorical moderators for easier understanding.

Data Availability Bias, Small Study Effects and Publication Bias

We examined whether studies with and without IPD were systematically different by comparing study and participant baseline characteristics. We investigated small study effects and publication bias by inspecting funnel plots and via Egger's test assessed using $p<0.1$ due to low power.[33] For potential asymmetry, we report results of fixed one-stage IPD meta-analysis (sensitivity analysis) in which less weight is given to smaller studies.

Data Sharing

Individual-level data cannot be made available to others, due to confidentiality agreements in the original studies.

RESULTS

Detailed results are provided for data availability, derivations and IPD integrity in Supplement 3, available online; additional tables and figures associated with overall pooled results in Supplement 4, available online, and participant moderators in Supplement 5, available online; and tables of all meta-analysis results in Supplement 6, available online.

Selected Studies

Up to January 21, 2022, we screened 3,690 citations and 286 full texts to identify studies. We sought IPD for 4,600 participants from 39 eligible studies (18 full sample eligible n=2,383; 21 partial sample eligible n=2,217) (Figure 1). We obtained IPD from 26 (66.7%) studies including 3,448 (75.0%) participants and further AD from 7 (17.9%) studies including 698

(15.2%) participants. Data were missing for 6 (15.4%) studies and 454 (9.9%) participants.[19] One study provided IPD but repetition of self-harm outcomes were unavailable.

Study Characteristics and Risk of Bias

Tables 1 and 2 detail and summarise eligible study characteristics (see also[19]). Studies were conducted across a wide range of countries and interventions with the largest proportion in the US (38.5%). Of eligible studies, CBT was evaluated in 10 (25.6%, 7 with IPD), PST in 6 (15.4%, 5 with IPD), family therapy and postcards/tokens in 5 (12.8%, 3 with IPD), DBT in 4 (10.3%, 2 with IPD), CAT/MBT in 3 (7.7%, all with IPD), group therapy in 3 (7.7%, 2 with IPD), brief-intervention in 2 (5.1%, 1 with IPD) and MST in one study (2.6%, no IPD). Just over half the eligible studies evaluated interventions of medium intensity, whilst 8 (20.5%) were classed as low and 10 (25.6%) as high intensity. Controls predominantly consisted of TAU in 25 eligible studies (64.1%, 18 with IPD), enhanced-TAU in 8 (20.5%, 5 with IPD) and active intervention in 6 (15.4%, 3 with IPD). Outcomes were most commonly reported at 3, 6, and 12-months post-randomization . Further details can be found in[19].

Risk of bias for studies providing IPD on repetition of self-harm outcomes was rated as low, some concerns, and high for 6/25 (24%), 16/25 (64%), and 3/25 (12%) studies. Additional AD were available for 5 studies with some concerns and 2 at high risk. No data were available for 4 studies with some concerns and 3 at high risk. The number of studies with concerns or at high risk largely arose from outcomes being self-reported from non-blinded participants and studies not having pre-specified published analysis plans.[19]

Table 1 Characteristics of Included Studies

Study	Data available[a]	Country	Design	Aim[b]	Study powered	Total sample size (eligible)	Eligibility	Mean age	Female participants (%)	Socio demographics[c] %	Intervention	Control Group	Treatment Intensity	Treatment Duration (weeks)	N treatment sessions	Any group element in treatment	Any family element in treatment	Method of self-harm data collection[d]	Overall Risk of Bias
Asarnow 2011[34]	No Data	US	2-arm RCT	Effectiveness	Yes	181	Partial	14.7	69	African American 13, Hispanic 45, Other 9, White non-Hispanic 33	CBT	TAU	Low	4	5	No	Yes	Self-report	Some concerns
Asarnow 2017[35]	Aggregate	US	2-arm RCT	Pilot	No	42	Full	14.6	88.1	African American 5, Asian 12, Hispanic/Latino 21, Other 7, White non-Hispanic 83	CBT	E-TAU	Medium	12	10	No	Yes	Self-report	Some concerns
Brent 2009[36]	No Data	US	3-arm RCT/ preference	Pilot	No	124 (22 randomised)	Partial	15.7	77.4	Not possible to determine for randomised	CBT	Active	High	26	24	No	Yes	Self-report	High
Carter 2005[37]	IPD	Australia	2-arm Zelen RCT	Effectiveness	Yes	772 (68 aged 11-18)	Partial	17.6	82.4	Not reported	Postcards/tokens	TAU	Low	52	0	No	No	Medical records	High
Chanen 2008[38]	IPD	Australia	2-arm RCT	Effectiveness	Yes	86 (72 prior SH)	Partial	16.4	79.2	Not reported	CAT/MBT	E-TAU	High	24	24	No	No	Self-report	Some concerns
Cooney 2010[39]	IPD	New Zealand	2-arm RCT	Pilot	No	29	Full	16	75.9	NZ European 77, NZ Māori 3, Other European 3, South African 7, UK 10,	DBT	TAU	High	26	52	Yes	Yes	Self-report	High
Cotgrove 1995[40]	Aggregate	UK	2-arm RCT	Effectiveness	No	105	Full	14.9	84.8	Not reported	Postcards/tokens	TAU	Low	0	0	No	No	Medical records	High
Cottrell 2018[41]	IPD	UK	2-arm RCT	Effectiveness	Yes	832	Full	14.8	88.6	Asian 3%, Black 7%, Other ethnic group 5, White 84, Missing <1	Family Therapy	TAU	Medium	26	8	No	Yes	Medical records	Low
Diamond 2010[42]	IPD	US	2-arm RCT	Effectiveness	No	66 (41 prior SH)	Partial	15	92.7	African American 74	Family Therapy	E-TAU	Medium	12	12	No	Yes	Self-report	Some concerns
Diamond 2014[e]	No Data	US	2-arm RCT	Pilot	No	20	Full	14.9	80	African American 65	Family Therapy	E-TAU	.	16	16	No	Yes	Self-report	High

Study	Data available[a]	Country	Design	Aim[b]	Study powered	Total sample size (eligible)	Eligibility	Mean age	Female participants (%)	Socio demographics[c] %	Intervention	Control Group	Treatment Intensity	Treatment Duration (weeks)	N treatment sessions	Any group element in treatment	Any family element in treatment	Method of self-harm data collection[d]	Overall Risk of Bias
Diamond 2019[43]	IPD	US	2-arm RCT	Effectiveness	Yes	129 (90 with prior SH)	Partial	15	84.4	African American 50, American Indian/Alaskan Native 2, Asian 2 Hispanic 16, Multiracial 8 Native Hawaiian/Pacific Islander 1, Other 9, White 29%	Family Therapy	Active	Medium	16	16	No	Yes	Self-report	Some concerns
Donaldson 2005[44]	IPD	US	2-arm RCT	Pilot	No	44[f]	Full	14.9	79.5	African American 5%, Hispanic 10, White 85%	PST	Active	Medium	26	10	No	Yes	Self-report	High
E-Smythers 2011[45]	IPD	US	2-arm RCT	Pilot	No	40 (35 prior SH)	Partial	15.7	65.7	Race: White 89 Ethnicity: Hispanic 14	CBT	E-TAU	High	52	54	No	Yes	Self-report	Some concerns
E-Smythers 2017[46]	IPD	US	2-arm RCT	Pilot	No	81 (37 prior SH)	Partial	15.5	67.6	Race: Black 37, Other 20, White 42 Ethnicity: Hispanic 17	CBT	E-TAU	Medium	4	3	Yes	Yes	Self-report	Some concerns
E-Smythers 2019[47]	IPD	US	2-arm RCT	Effectiveness	Yes	147 (133 prior SH)	Partial	14.8	79.7	Race: Asian/Pacific Islander 3, Black/African American 2, Multiracial 10, White 86 Ethnicity: Non-Hispanic/Latino 83	CBT	E-TAU	High	52	54	No	Yes	Combined approach	Low
Green 2011[48]	IPD	UK	2-arm RCT	Effectiveness	Yes	366	Full	15.1	88.5	BAME 6	Group Therapy	TAU	Medium	6	10	Yes	No	Self-report	Some concerns
Griffiths 2019[49]	IPD	UK	2-arm RCT	Pilot	No	53	Full	15.5	79.2	White/Scottish 69	CAT/MBT	TAU	Medium	12	12	Yes	No	Medical records	Some concerns
H-Moghaddam 2017[50]	IPD	Iran	2-arm RCT	Effectiveness	Yes	2300 (549 aged 11-18)	Partial	16.7	74.3	Not reported	Postcards/tokens	TAU	Low	12	0	No	No	Self-report	Some concerns

Study	Data available[a]	Country	Design	Aim[b]	Study powered	Total sample size (eligible)	Eligibility	Mean age	Female participants (%)	Socio demographics[c] %	Intervention	Control Group	Treatment Intensity	Treatment Duration (weeks)	N treatment sessions	Any group element in treatment	Any family element in treatment	Method of self-harm data collection[d]	Overall Risk of Bias
Harrington 1998[51]	Aggregate	UK	2-arm RCT	Effectiveness	Yes	162	Full	14.5	89.5	Not reported	Family Therapy	TAU	Medium	.	5	No	Yes	Self-report	Some concerns
Hatcher 2011[52]	IPD	New Zealand	2-arm Zelen RCT	Effectiveness	Yes	1094 (89 aged 11-18)	Partial	18.3	67.4	Asian 3, Māori 16, NZ European 61, Other 14, Pacific Island 6	CBT	TAU	Medium	12	9	No	No	Medical records	Low
Hatcher 2015[53]	IPD	New Zealand	2-arm Zelen RCT	Effectiveness	Yes	1474 (98 aged 11-18)	Partial	17.8	77.6	Asian 6, Māori 7, NZ European 79, Other 2, Pacific Island 7	CBT	TAU	Medium	52	8	No	No	Medical records	Low
Hazell 2009[54]	IPD	Australia	2-arm RCT	Effectiveness	No	82	Full	14.5	90.2	Not reported	Group Therapy	TAU	Medium	6	6	Yes	No	Self-report	Some concerns
Huey 2004[55]	No Data	US	2-arm RCT	Effectiveness	No	156 (70 with prior SH)	Partial	12.9	35	African American 65, European American 33, Other 1	MST	Active	High	16	.	No	Yes	Self-report	Some concerns
Husain 2014[56]	IPD	Pakistan	2-arm RCT	Effectiveness	Yes	221 (53 aged 11-18)	Partial	17.4	77.4	Not reported	PST	TAU	Medium	12	6	No	No	Self-report	Some concerns
Kaess 2019[57]	IPD	Germany	2-arm RCT	Effectiveness	Yes	74	Full	14.9	95.9	Migration status: Germany 92	CBT	TAU	Medium	16	12	No	No	Self-report	Some concerns
King 2006[58]	No Data	US	2-arm RCT	Effectiveness	No	289 (190 prior SH))	Partial	15.3	68.2	Black 10, Other 7, White 82	PST	TAU	Medium	26	26	No	Yes	Self-report	Some concerns
King 2009[59]	IPD	US	2-arm RCT	Effectiveness	No	448 (331 prior SH)	Partial	15.6	73.1	African-American 6, Caucasian 84, Hispanic 2, Other 8	PST	TAU	Medium	12	12	No	Yes	Self-report	Some concerns
McCauley 2018[60]	Aggregate	US	2-arm RCT	Effectiveness	Yes	173	Full	14.9	94.8	African American 7, Asian American 6, Hispanic 27, Native American <1, Other 2, White 56	DBT	Active	High	26	52	Yes	Yes	Self-report	Some concerns
Mehlum 2014[61]	Aggregate	Norway	2-arm RCT	Effectiveness	Yes	77	Full	15.6	88.3	Norwegian Ethnicity 85	DBT	E-TAU	High	19	38	Yes	Yes	Medical records	Some concerns

Study	Data available[a]	Country	Design	Aim[b]	Study powered	Total sample size (eligible)	Eligibility	Mean age	Female participants (%)	Socio demographics[c] %	Intervention	Control Group	Treatment Intensity	Treatment Duration (weeks)	N treatment sessions	Any group element in treatment	Any family element in treatment	Method of self-harm data collection[d]	Overall Risk of Bias
Morthorst 2012[62]	IPD	Denmark	2-arm RCT	Effectiveness	Yes	243 (46 aged 11-18)	Partial	16.1	95.7	Danish 67, European/American 8, Middle Eastern 14, Other 10	PST	TAU	Medium	26	20	No	Yes	Medical records	Some concerns
O'Connor 2017[63]	IPD	UK	2-arm RCT	Effectiveness	Yes	518 (39 aged 11-18)	Partial	17.2	74.4	Not reported	Postcards/tokens	TAU	Low	8	1	No	No	Medical records	Low
Ougrin 2013[64]	IPD	UK	2-arm cluster-RCT	Effectiveness	Yes	70	Full	15.6	80	Asian 11, Black 20, Mixed 13, Other 3, White 53	Brief-Intervention	TAU	Low	1	1	No	Yes	Medical records	Some concerns
Pineda 2013[65]	IPD	Australia	2-arm RCT	Effectiveness	Yes	48 (48 with prior SH[g])	Partial	15.1	79.2	Aboriginal 8, Anglo-saxon 58, CALD/ NESB 35	PST	TAU	Medium	8	4	No	Yes	Self-report	Some concerns
Robinson 2012[66]	No Data	Australia	2-arm RCT	Effectiveness	Yes	164 (~56 aged 11-18 prior SH)	Partial	18.6	64.6	Born in Australia 89	Postcards/tokens	TAU	Low	52	0	No	No	Self-report	High
Rossouw 2012[67]	IPD	UK	2-arm RCT	Effectiveness	Yes	80	Full	15.1	85	Asian 10, Black 5 , Mixed 7.5, Other 2.5, White 75	CAT/MBT	TAU	High	52	64	No	Yes	Self-report	Some concerns
Santamarina 2017, 2020[68,69]	IPD	Spain	2-arm RCT	Effectiveness	No	35	Full	15.3	88.6	White 91	DBT	Active	High	16	40	Yes	Yes	Combined approach	Some concerns
Spirito 2002[70]	Aggregate	US	2-arm RCT	Effectiveness	No	76	Full	15	90	African American 11, Hispanic 13, Mixed ancestry 3, White 73,	Brief-Intervention	TAU	Low	8	5	No	Yes	Self-report	High
Tyrer 2003[71]	IPD	UK	2-arm RCT	Effectiveness	Yes	480 (54 aged 11-18)	Partial	17.8	88.9	White 90	CBT	TAU	Medium	12	7	No	No	Combined approach	Low
Wood 2001[72]	Aggregate	UK	2-arm RCT	Pilot	No	63	Full	14.2	78	Not reported	Group Therapy	TAU	Medium	26	26	Yes	No	Self-report	Some concerns

Note: Active = active control; AD = aggregate data; CALD = culturally and linguistically diverse; CBT = cognitive–behavioural therapy; DBT = dialectical behavior therapy; E-TAU = enhanced treatment as usual; IPD = individual participant data; MBT/CAT = mentalisation based, psychodynamic, cognitive analytic therapy; MST=multi-systemic therapy; NESB = non–English-speaking background; PST = problem solving, psychoeducation, support; TAU = treatment as usual.

[a] The reasons IPD were not provided or excluded were: the data had been lost and were no longer available for Cotgrove 1995[40], Harrington 1998[51], Spirito 2002[70], Wood 2001[72] and King 2009[59]; authors did not agree to share data or felt they did not have ethical approval to share for Asarnow 2011[34], Asarnow 2017[35], McCauley 2018[60], Mehlum 2014[61], and Robinson 2012[66]; there was no response to our request for Huey 2004[55]; IPD

were obtained but excluded for Diamond 2014 (unpublished data) as the IPD were not consistent with limited aggregate results detailed in the list of excluded studies in Hawton 2015[73]; and IPD were obtained but excluded for Brent 2009[36] as eligible participants could not be identified.
[b]Pilot studies includes pilot or feasibility studies.
[c]Ethnicity as reported in original studies, with the exception of Santamarina 2020[68,69] as based on IPD. Caucasian as reported in the original study.
[d] Self-report includes researcher interview; medical records include hospital records, studies that collected the primary outcome using a combination of methods primarily relied on self-report verified by medical record.
[e] Unpublished data; ClinicalTrials.gov NCT01195740
[f] Donaldson 2005[44] reported 39 randomised but IPD confirmed 44 randomised.
[g] Pineda 2013[65]: Partially eligible as eligibility based on suicidal behaviour including ideation only, however authors confirmed all met RISA eligibility criteria. IPD did not include the RISA primary outcome (aggregate data also unavailable).

Available Adolescent-Level Moderators

Age and gender were the only baseline characteristics available in all 26 studies with IPD (Table 3). Most participants were female (2,823/3,448, 82.0%) with mean age 15.7 (SD=1.6) years. In 14 (53.8%) studies with IPD, 1,471/2,779 (52.9%) participants presented to services with self-poisoning, 1,052 (37.9%) with self-injury, and 245 (8.8%) used a combination. In five of these studies, the majority or all participants had self-poisoned. In seven (26.9%) studies, including all three studies of CAT/MBT, 163/921 (17.7%) participants identified as having borderline personality disorder (BPD). Over half the participants were reported as depressed (1,138/1,989 (57.2%), studies=15) and had self-harmed multiple times (1,874/3,213 (59.3%), studies=21). Just over half the participants had anxiety (clinically diagnosed or questionnaire indicated, 917/1,682 (54.5%), studies=14). Over three-quarters were white (1,947/2,482 (78.4%), studies=18) and just under three-quarters exhibited suicidal ideation (1,620/2,221 (72.9%), studies=15); in one study, all participants showed suicidal ideation as it was a criterion for eligibility. Family dysfunction was indicated in 747/966 participants (77.3%, studies=4) and 325/2,198 participants (14.8%, studies=12) were on psychotropic medication at baseline. Whilst data were available in less than half the trials, these were included in analyses as additional moderators as available in all three family therapy studies. Other baseline characteristics available in fewer than half of the studies included unemotional and callous traits, lesbian, gay, bisexual, transgender and queer (LGBTQ) status, autistic spectrum disorder, abuse, presence of an eating disorder, intellectual disability, out of home placement, and presence of a physical health problem. See Supplement 3, available online, for further details.

Primary Outcome – Repetition of Self-Harm

Overall, 2,949 (93.7% eligible) participants from 20 studies and 3,117 (89.5% eligible) participants from 22 studies were included in IPD and IPD+AD meta-analyses respectively at 12 months (Figure 2, Table 3) in whom 973 (33.0%) and 995 (31.9%) participants were reported to have repeat self-harmed, with repetition rates ranging from 6.1% to 92.3% (Supplement 4, Table S4.2.2, available online). There was no evidence that interventions were more or less effective than controls at reducing repeat self-harm at 12-months using IPD (OR=1.06 [95% CI 0.86, 1.31], studies=20, n=2,949) or IPD+AD (OR=1.02 [95% CI 0.82, 1.27], studies=22, n=3,117), or at other timepoints. Whilst the confidence intervals include the null effect, the upper and lower bounds include both small positive and negative adverse effects. Due to high levels of within-study variability in outcome, between-study heterogeneity in treatment effects was relatively low (IPD I^2=1.1%, p=0.443; IPD+AD I^2=12.1%, p=0.299) and there was no evidence of heterogeneity in treatment effects between groups of interventions (IPD p=0.779; IPD+AD p=0.759) i.e. intervention specific effects were consistent, and did not differ significantly, across studies.

Except CBT at 12 months, most intervention specific pooled effect estimates were based on four or fewer studies, and there was no evidence that any were more or less effective than control at reducing repeat self-harm at any time-period in primary IPD or IPD+AD meta-analysis.

There was evidence of between-study heterogeneity in treatment effects at the 10% level for the 2-3 studies of group therapy (IPD I^2=72.2%, p=0.058; IPD+AD I^2=76.2%, p=0.015) at 12-months; and between two IPD studies of CBT at 18-months (I^2=71.7%, p=0.06) and CAT/MBT at 3-months (I^2=64.6%,p= 0.093) in which contrasting treatment effects were observed.

Funnel plots showed no evidence of small study effects or publication bias. There were few notable differences in results and no changes to conclusions from sensitivity analyses using multiple imputation, one-stage random or fixed-effects IPD meta-analysis, or IPD+AD meta-analysis excluding studies of high risk of bias. Compared to random effects, one-stage fixed-effect IPD meta-analysis provided similar estimates but with tighter confidence intervals.

Heterogeneity and Study-Level Moderating Effects. There was no statistical evidence of between-study heterogeneity in treatment effect when all interventions were compared with control at any time point. In IPD+AD meta-analyses, heterogeneity was estimated to be zero at all timepoints with exception of 12- (I^2=12.1% [95% CI 0, 48.2%], studies=22) and 18-months (I^2=14% [95% CI 0, 66.8%], studies=6). Given the lack of heterogeneity and small number of studies, meta-regression was conducted only at 12-months. No candidate moderators were significantly associated with treatment effect (Figure 2).

Secondary Outcomes

Time to Repetition of Self-Harm. Overall, 1539 (98.8% eligible) participants (8 studies) were included in IPD meta-analyses (Supplement 4, Figure S4.4.1, Table 2, available online). Median follow-up was 43 months (range 0-82.5 months) and ranged 6-60 months across studies. There was no evidence that interventions (overall or by intervention) were more or less effective than control on time-to-repetition in IPD meta-analyses. There was no evidence of between-study heterogeneity overall or by intervention, or heterogeneity between groups of interventions.

Pattern of Self-Harm Repetition. Overall, 2,083 (92.9% eligible) participants (14 studies) and 2 212 (91.6% eligible) participants (15 studies) were included in IPD and IPD+AD meta-analyses of self-harm repetition between 6 to 12-months post-randomization respectively (Supplement 4, Figure S4.3.1a, Table 3). One study was excluded due to zero events in both arms. There was no evidence that interventions (overall or by intervention) were more or less effective than controls at reducing self-harm between 6 to 12 months, or other time-periods using IPD or IPD+AD; all confidence intervals included the null effect and spanned mainly small (positive and negative) effects overall, and small to large effects by intervention.

Overall, between-study heterogeneity was low at 6 to 12 months (IPD I^2=4.2%, p=0.405; IPD+AD I^2=9%, p=0.352) and 12 to 18-months (IPD I^2=20.1%, p=0.282). There was no evidence of heterogeneity between groups of interventions; however there was some evidence at the 10% level of between-study heterogeneity for two group therapy studies with contrasting effects at 6 to 12-months (IPD I^2=69.0%, p=0.072; IPD+AD I^2=67.1%, p=0.081), and two CBT studies at 12 to 18-months in adjusted IPD meta-analyses (I^2=64.9%, p=0.092).

General Psychopathology. Overall, 1369 (71.9% eligible) participants (8 studies) and 1564 (73.0% eligible) participants (10 studies) were included in IPD and IPD+AD meta-analyses respectively at 12-months (Figure 3, Table 2). There was good evidence of a small positive effect, indicating that interventions overall were more effective than control at reducing general psychopathology using IPD (SMD=-0.13 [95% CI -0.25, -0.01], studies=8, n=1,369) and IPD+AD (SMD=-0.13 [95% CI -0.25, -0.02], studies=10, n=1,564). Between-study heterogeneity was low (IPD I^2=5.7%, p=0.386; IPD+AD I^2=0%, p=0.458) and there was no evidence of heterogeneity between studies or groups of interventions. Effects were consistently positive, but not statistically significant, at other timepoints.

There was insufficient evidence to detect a statistically significant reduction compared to control at 12-months using IPD or IPD+AD for specific interventions; confidence intervals were wide spanning small to large positive and negative effects. There was evidence of a statistically significant reduction in general psychopathology for specific interventions compared to control from IPD when only single studies were available for (Supplement 4 and Supplement 6, available online): PST at 3 and 6-months; DBT at 6-months but not when including two additional studies with AD; and CBT at 18-months.

There was again some evidence of between-study heterogeneity for two group therapy studies in IPD (I^2=67.0%, p=0.082) but not IPD+AD analyses at 12-months. There was further statistical evidence of between-study heterogeneity in the overall treatment effect in IPD but not IPD+AD meta-analyses at 3-months due to the large treatment effect observed for one study of PST in adjusted analysis; and 6-months due to larger treatment effect estimates observed for one DBT and PST study.

Depression. Overall, 1,481 (67.9% eligible) participants (12 studies) and 1,672 (69.1% eligible) participants (14 studies) were included in IPD and IPD+AD meta-analyses, at 12 months (Supplement 4 Figure S4.6.3, Table 2). There was no evidence that therapeutic interventions (overall or by intervention) were more or less effective than controls at reducing depression using IPD or IPD+AD at any timepoint; all confidence intervals included the null effect and spanned mainly small (positive and negative) effects overall, and small to large effects by intervention.

There was no evidence of heterogeneity between groups of interventions at any timepoint. There was statistical evidence of between-study heterogeneity in the overall treatment effect at 18 months in IPD and IPD+AD meta-analyses, driven by conflicting effect estimates between two studies of CBT. There was also statistical evidence of heterogeneity between two studies of family therapy at 3 months and some evidence at 6-months in IPD meta-analysis only, and between three studies of PST at 6-months in adjusted IPD meta-analysis only.

Suicidal Ideation. Overall, 1,323 (70.0% eligible) participants (9 studies) and 1,510 (72.2% eligible) participants (11 studies) were included in IPD and IPD+AD meta-analyses respectively at 12 months (Supplement 4 and Figure S4.7.3, Table 2). There was no evidence that any therapeutic intervention (overall or by intervention) was more or less effective than control at reducing suicidal ideation using IPD or IPD+AD, and no evidence of between-study heterogeneity, or heterogeneity between groups of interventions at 3, 12- or 18-months.

Evidence of a small positive effect of intervention compared to control was observed at 6 months (Supplement 4 and Figure S4.7.3) in IPD+AD meta-analysis (SMD=-0.17 [95% CI -0.32, -0.02], studies=15, n=1,418), however this effect was not supported by IPD meta-analysis

with fewer studies (two DBT, one FT). The positive overall effect observed was driven largely by small to moderate effects for DBT (SMD=-0.43 [95% CI -0.90, 0.03], studies=4, n=258), CBT (SMD=-0.24 [95% CI -0.90, 0.42], studies=3, n=170) and family therapy (SMD=-0.15 [95% CI -0.95, 0.65], studies=3, n=261).

There was evidence at the 10% level of heterogeneity between the two family therapy studies at 6-months in IPD meta-analysis but not IPD+AD meta-analysis including an additional study. There was also evidence of heterogeneity between groups of interventions at 6-months in IPD+AD meta-analysis (p=0.094) driven by the range of effects observed for DBT, CBT and family therapy, and null effects found for group therapy and PST.

Assessment of Publication Bias and Small Study Effect. Funnel plots showed no evidence of small study effects or publication bias, with a few exceptions. Outlying studies with positive treatment effects and a lack of symmetry were observed for IPD+AD meta-analysis of general psychopathology at 6 and 12-months. Pooled treatment effect estimates from one-stage fixed-effect IPD meta-analysis, were comparable to random-effects IPD meta-analysis at 12-months but extenuated the treatment effect at 6 months.

Lack of symmetry was observed for IPD+AD meta-analysis of depression at 6 months (Egger's test p=0.024), and suicidal ideation at 3 and 6-months (Egger's test p=0.063 and p=0.021 respectively), however this was less pronounced for adjusted IPD meta-analysis and one-stage fixed-effect and two-stage random-effects IPD meta-analysis were comparable.

Sensitivity Analysis. Two-stage IPD meta-analysis with multiple imputation gave similar results overall and by intervention compared with primary complete case analysis except for 3-month depression outcomes for family therapy interventions, where there was a reduction in the magnitude of the pooled treatment effect estimate and no longer significant between-study heterogeneity. One-stage random-effect IPD meta-analysis gave similar pooled treatment effect estimates compared to the primary two-stage approach, with generally wider confidence intervals both overall and by intervention.

One-stage fixed-effect IPD meta-analysis resulted in pooled treatment effect estimates which were generally comparable or closer to the null with tighter confidence intervals compared with estimates obtained in primary random-effects IPD meta-analyses. Nevertheless, there were few changes to conclusions, except for general psychopathology where smaller CIs for fixed-effect estimates gave significant evidence that interventions overall and group therapy interventions were more effective than control at 6-months, as were group therapy interventions at 12-months.

Additional sensitivity analysis of two-stage IPD+AD meta-analysis on pattern of self-harm outcomes to explore the impact of zero events in one arm found no change to conclusions.

Additional Outcomes

Quality of life was available for 670 (59.3% eligible) participants (5 studies) with IPD and no studies with AD at 12-months. There was no evidence that interventions were more or less effective than controls in individual studies; estimated effects were small to moderate and positive in all but one study however confidence intervals all included the null effect and spanned positive and negative small to large effects.

Adolescent deaths were reported for 11 eligible participants in 7 studies (6 IPD, 1 AD; 7 intervention, 4 control) and for 2 participants (1 control, 1 unknown) in 2 partially eligible studies where eligibility could not be determined. No eligible participants died in 10 studies and no deaths were reported in the remaining 20 studies.

Participant Moderators

Our analysis focused on the 6 and 12-month timepoints as they encompassed all studies with available outcomes (further details supplement-3 and 5, available online).

We found no evidence of moderating effects on primary or secondary outcomes, overall or in specific groups of interventions, according to key moderators of participants age, gender, depression, method of self-harm, or BPD diagnosis; or on repeated self-harm outcomes, according to additional moderators of anxiety, family dysfunction, ethnicity, psychotropic medication, or suicidal ideation.

There was evidence of an improved treatment effect in participants with multiple previous self-harm episodes compared to those with fewer (≤two) on repeated self-harm within 6-12 months (OR=0.33 [95% CI 0.12, 0.94], studies=9, n=1,771; Figure 4). This interaction should be considered alongside the overall treatment effect, which showed a non-significant 7% reduction in the likelihood of repeat self-harm in intervention vs control (OR=0.93 [95% CI 0.71, 1.23], studies=14; Figure 2). A more favourable, but not statistically significant, treatment effect was similarly indicated in participants with multiple previous self-harm episodes on primary repeat self-harm outcomes at 6 and 12 months, general psychopathology, and suicidal ideation outcomes, and on 12 but not 6-month depression outcomes. There was no evidence of

variability between studies, heterogeneity among different groups of interventions, or variability within groups of interventions.

Assessment of Publication Bias and Small Study Effect of Participant Moderators

Funnel plots for all moderators and outcomes, and regression-based Egger tests (Figures S4.1.3–S4.13.2; Table S2, available online), generally indicated no evidence of small study effects or publication bias. The notable exception was the moderating effect of the level of depression on the primary outcome at 12-months, where a marginal indication of potential bias was observed (p=0.0638).

Sensitivity Analysis of Participant Moderators. In sensitivity analysis where confidence intervals for combined treatment effects were calculated without the Hartung-Knapp-Sidik-Jonkmann[20] adjustment, results appeared more precise due to the exclusion of uncertainty in estimated between-study heterogeneity. This change in methodology identified PST as more effective for older participants, on 6-months repetition of self-harm (OR=0.56 [95% CI* 0.26, 0.90], studies=4, n=400; Figure S5.1.1) and suicidal ideation (SMD=-0.19 [95% CI* -0.36, -0.03], studies=3, n=319; Figure S5.1.2) outcomes; and CBT as more effective for male participants on depression outcomes at 6- (SMD=-0.70 [95% CI* -1.39, -0.02], studies=3, n=190; Figure S5.2.2) and 12-months (SMD=-0.69 [95% CI* -1.39, -0.0], studies=4, n=216; Figure S5.2.2). These effects were primarily influenced by a more pronounced treatment effect in older participants in the King-2009[59] study, and in male participants in the Esposito-Smythers-2019[47]study however there was no evidence of variability between studies.

Heterogeneity of Participant Moderators. Heterogeneity, both between studies, and by and within studies of specific interventions, was observed across a range of examined moderators

and outcomes where no evidence of a participant-level moderating effect was observed. These included moderators of age, self-harm method, BPD, depression, family dysfunction; and particularly between study variability in two group therapy studies, and varying CBT and PST studies.

DISCUSSION

Summary of Evidence

In summary of our findings: We identified 39 studies that met our inclusion criteria. Our IPD analysis set comprised 26 studies with 3,448 eligible participants and additional AD from 7 studies with 698 participants were included in IPD+AD meta-analysis. For our primary outcome, repetition of self-harm, only 6/39 studies were rated as low risk of bias with 8/39 rated as high risk.

Primary Outcome. There was no evidence that any therapeutic intervention (overall or by intervention) was more, or less, effective than control for reducing repeat self-harm between randomization and 3, 6, 12, 18, 24 or >24 months. There were high levels of within-study variability in outcomes, and generally low between-study heterogeneity in treatment effects, with the exception of a few studies of CAT/MBT at 3-months, group therapy at 12-months and CBT at 18-months. We found no evidence for candidate moderator study or treatment effects, small study effects or publication bias, and sensitivity analyses led to no changes to our conclusions.

Secondary Outcomes. We found no evidence that any therapeutic intervention (overall or by intervention) was more or less effective than control on time-to-repetition of self-harm, pattern

of self-harm repetition (within 6-month periods post-randomization), or depression at any timepoint.

There was good evidence that interventions overall had a small positive effect on general psychopathology at 12-months in IPD and IPD+AD meta-analysis, however, there was insufficient evidence for specific interventions (with only 1-3 studies available per intervention). There was no further evidence for any therapeutic intervention overall at other timepoints, however there was some limited evidence from single studies in support of CBT,[45] DBT[68,69] and PST[65] at certain time points using IPD, but this was not supported for DBT when including additional AD.

There was good evidence that interventions overall had a small positive effect on suicidal ideation at 6-months in IPD+AD meta-analysis.[39,60,61,68,69] There was insufficient evidence for specific interventions (with only 2-4 studies available per intervention), however, the overall treatment effect was subject to heterogeneity between groups of interventions driven by the moderate and small positive (non-significant) effects observed for DBT and CBT compared to other interventions. This finding was not detected at other timepoints or in IPD meta-analysis in which IPD for some DBT studies were not available.

Participant Moderators. There was a notable improvement in treatment effect for participants with multiple previous self-harm episodes, compared with those with fewer episodes on repetition of self-harm 6-12 months post-randomization , accompanied by a consistent trend across other outcomes.

There was more limited evidence for potential differential treatment effects based on participants' age and gender in PST and CBT respectively, as indicated by confidence intervals not adjusted for uncertainty in variance estimates.

The strengths of this study are that an IPD meta-analysis provides more robust estimates of the effects of therapeutic interventions for self-harm compared to conventional meta-analyses that rely on aggregated data and reported analyses.[16] The IPD approach enables a consistent and comprehensive analysis across a wide range of studies, outcomes, and moderators, significantly enhancing the power to detect interaction effects. This capability surpasses what can be achieved through single trials or AD meta-analysis. This IPD meta-analysis represents a pioneering effort in this specific population. Prior to this, no similar IPD meta-analysis had been conducted, marking this study as the first of its kind to explore the critical clinical problem of adolescent self-harm.

Rigorous inclusion criteria and inclusion of trial registrations in the search allowed us to minimise selection and publication bias. All records were screened independently by two authors, with a third adjudicating if an agreement could not be reached. Participants in eligible studies must have self-harmed leading to contact with services, prior to randomization , thus excluding studies using non-clinical samples and making our participants generally more troubled and in need of services. We had a broad geographical spread of studies. The IPD approach allowed us to include studies where only a part of the sample was eligible. We were able to identify 21 additional studies (~2,217 participants) and obtain IPD for 1,783 participants from 16 of these studies; providing more studies and participants than were able to obtain IPD for from studies in which the full sample were eligible. The most recent Cochrane review[10] included only 17 trials with a total of 2,280 participants. The inclusion of substantially more

studies and participants than in previous reviews has narrowed confidence intervals, increased precision and given more confidence in our findings. The use of sub samples from studies not included in other reviews has also enabled us to include and report on interventions, such as postcards, that have either not been included or included only in small numbers in other meta-analyses.

Finally, an additional strength came from inviting the original study authors who shared data, to join our study collaborative group. This collaboration allowed for their contributions to the development of the analysis plan and the interpretation of results, thus enhancing the trustworthiness of findings.

A limitation of this study is that the eligible RCTs we identified were conducted by various groups worldwide over several decades, differing in therapeutic orientation, intensity, and duration of intervention, controls, aims, outcomes, and targeted subsets of adolescents. Given the seriousness of the clinical problem, it is surprising how few good quality studies we identified. The wide variety of interventions within the eligible studies meant that for any single intervention, there were few good quality studies, and for some intervention types, very few, with some subgroup analyses based on only a single study. Replication was lacking. On the one occasion it occurred, Hazell et al's[54] replication of Wood et al[72] resulted in the earlier findings being contradicted.

A significant limitation was missing IPD with some authors being unable to share data despite our willingness to receive reduced datasets addressing concerns about participant identification. In some cases, data had been lost or destroyed. Whilst IPD+AD meta-analysis was able to incorporate AD from studies lacking IPD, this was not possible for studies where

only a subset of participants was eligible and where published AD pertained to the full sample, or where insufficient AD were reported in studies where the full sample were eligible.

We were unable to obtain IPD for two important studies (both rated as some concerns of risk of bias) that are often cited in systematic reviews and meta-analyses as showing evidence for effectiveness of dialectical behaviour therapy.[60,74] Accordingly, our IPD meta-analysis includes only two small underpowered studies of DBT.[39,68,69] This is a clear limitation, however we were able to incorporate aggregate data from the published results of these two studies in our IPD+AD meta-analysis. This did not lead to any change in our conclusions.

A major challenge is the variability in the definitions of outcomes, timing, methods, and measures used for data collection across studies of self-harm interventions, which potentially increase between study heterogeneity and limit the ability to meaningfully pool studies and interpret pooled treatment effects. Regarding timing, we defined multiple follow-up time points in order to pool treatment effects over consistent periods of time across studies. Whilst ensuring consistency in timing, this increased the number of separate meta-analyses each containing fewer studies than had we selected a single but variable post-intervention timepoint for each study. Methods of data collection also varied from self-report, parent-report, clinical interview or rating, and medical record review. At least eight different measures were used to collect self-harm, general psychopathology, and depression outcomes.

Studies also used variable age ranges, perhaps reflecting different patterns of care in the locations where studies were conducted. Our choice of 11-18 years of age (up to the 19th birthday) will have implications for interpretation in places where different age cut points are

used by local services. Our choice was largely determined by the cut points used in other large reviews (8-10).

The paucity of data across common outcomes and follow-up durations significantly constrained our exploration of potential participant moderators. Apart from age and gender, other baseline characteristics such as ethnicity, family dysfunction, psychotropic medication use, LGBTQ status, autistic spectrum disorder, history of abuse, eating disorders, intellectual disability, children in out of home placement, and physical health problems were inconsistently collected. There was also variability in the measures and methods used for collecting data on moderators and outcomes. Standardisation of effects within each study and by timepoint was employed where possible to address this issue. While the IPD method allowed us to include subsets of participants from larger studies, these larger studies often focused more on adults, making them less likely to include child and adolescent specific characteristics.

Our choice of a binary primary outcome of any repetition of self-harm could be seen as a limitation but was chosen considering the serious difficulties in accuracy and consistency of alternative measurements such as the number/ frequency of self-harm across studies, particularly for those who self-harm repeatedly. However, relevant outcomes varied considerably across studies, as well as the measures used to collect outcomes, and included NSSI, hospital attendance for self-harm/poisoning, suicide attempt, parasuicidal behaviour NSSI, unspecified self-report. Therefore, our primary outcome described whether young people had self-harmed or not at 12 months, not a reduction in the number of self-harm attempts and our conclusions need to be understood in this light. Evidence from those with lived experience of NSSI suggest that recovery needs to be understood in more complex ways than just cessation or reduced frequency of self-harming behaviour[75].

Finally, it should be noted that when fewer than 10 trials are included in the meta-analysis, or when trials are small, or the outcome is rare, no currently available method can reliably estimate the heterogeneity.[29] More sophisticated analysis approaches were considered to account for some of the further data complexities, but it was felt that while none would alter the conclusions materially, they would add to the complexity of conducting and reporting the analyses. These are left for future research.

In summary, we had hoped that using IPD methods would provide more accurate information about effective interventions and identify subgroups of young people who might benefit from specific types of intervention. Unfortunately, we were unable to fully achieve these aims. Although IPD allowed us to include substantially more studies and participants, we still pooled a relatively small number of studies once different interventions, endpoints and follow-up times were considered. While IPD meta-analysis remains a potentially powerful tool, its effectiveness is dependent on the quantity and quality of the studies it draws upon.

Nevertheless, we argue that this robust reinforcement of the broad message from other reviews is clinically important. It serves as a counterweight to some assessments of the long-term effectiveness of psychological therapies for treating self-harm in adolescents, which are often based on relatively small (underpowered) studies. We did not reach the same conclusion as Kothgassner[11] that, overall, any intervention is more effective than active controls. Our findings also suggest caution regarding the effectiveness of DBT. The most recent NICE guideline recommended 'consideration' of DBT-A for children and young people but could not make a 'strong recommendation' due to limited evidence on repeat self-harm by 12-month follow-up. This recommendation was based on a recent Cochrane review[10], which found

positive effects of DBT-A on repetition of self-harm post-intervention, as well as improved depression, hopelessness, and suicidal ideation outcomes in the short term. These short term findings align with our 6-month timepoint and were observed immediately after the intensive treatment ended. However, these effects were not present at our 12-month analysis. Therefore, despite the public health importance of self-harm, its many adverse outcomes, and a rigorous IPD meta-analysis design, we cannot recommend a specific, safe intervention for the prevention of self-harm repetition among those who present with self-harm. Why might this be?[76]

Ethical concerns have meant that except for very low intensity 'postcard' type interventions, all the studies we examined looked at specific intervention types compared with a control treatment, not a no treatment control. Control treatments varied widely from 'treatment as usual' to active manualised specific treatments, including antidepressant pharmacotherapy, family-enhanced nondirective supportive therapy, supportive relationship treatment, hospitalisation, and individual and group supportive therapy. There is therefore likely to be overlap of treatment components not just between different interventions but between intervention and control treatments. Well trained clinicians are likely to be well versed in the existing evidence base, and so some or many TAU control interventions may also be effective. Thus, both arms in some RCTs may have received comprehensive care, thereby reducing treatment effect sizes. It is possible practitioners delivering control treatments used some of the same techniques specified in intervention arms, but even if this did not occur, staff seeing control arm participants were able to conduct assessments and tailor their interventions to individual needs, perhaps in a more flexible way than in the manualised intervention arm where only the target intervention could be delivered.

Our findings might also reflect the nature of the standardised interventions evaluated, which focus to a large degree on risk assessment and management. This nearly always includes encouragement to report self-harm, which may inflate incidence of reported repeat self-harm in the follow up period. Some support for this argument comes from our findings that there were generally more positive (but not significant) effects across studies for general psychopathology, depression, and suicidal ideation, compared to self-harm outcomes.

Finally, it is important to stress that, in line with the conclusions of the most recent Cochrane review[10], we found no evidence that interventions were more, or less, effective than control treatments. We did not find that treatments were ineffective. Clinicians should not interpret these findings as meaning that interventions do not work, rather that we do not yet know which interventions are effective for which young people. It is important to emphasise that young people who self-harm are at elevated risk for many adverse outcomes and should undoubtedly receive help and support. Our findings indicate that those who have engaged in recurrent self-harm (defined here as more than two previous episodes) might respond more positively to treatment. Given their higher risk, it is imperative that they are thoroughly assessed and offered an intervention deemed most appropriate by a trained and qualified clinician.

Future recommendations include the need for innovative ideas to optimise standard care and develop alternative interventions. These should be more deeply grounded in theoretical considerations or mechanisms that drive self-harm. Collaborative research programmes are essential to rapidly execute large-scale, well-designed studies and to establish core outcome sets of trials of self-harm interventions. Given how difficult it is to alter the developmental trajectories of those who self-harm, it is also crucial to explore prevention strategies.

Future evaluations of therapeutic interventions for adolescents following self-harm should employ novel and efficient clinical trial designs to (i) optimise existing therapeutic approaches or treatment as usual, for example, through a SMART design[77], and (ii) ensure larger comparative head to head comparisons of interventions and tailor intervention strategies, such as through platform[78] or factorial[79] trial designs that focus on optimising fixed or adaptive interventions. These approaches are better suited to address multiple research questions and determine the most effective interventions for different individuals, while simultaneously ensuring consistency in research methods.

It seems unlikely that future IPD meta-analyses in this area will be fruitful until we have more well designed and conducted studies that utilise agreed outcome measures. Future research to develop and agree on core outcome sets for self-harm trials is vital. More detailed reporting of control arm interventions would increase the homogeneity of outcome reporting, allowing future studies to be pooled more efficiently and inclusively.[76] Establishing a clear and agreed definition of the primary outcome with standardised follow-up times is crucial, given the complex nature of self-harm and its related outcomes. There should be consensus on the moderators to be included in future studies. It is important to consider participant-level treatment effect moderators during trial design to ensure that the study is tailored to detect important differences and trends. Many important moderators are currently inconsistently collected. Ensuring that these are consistently included in future research will greatly enhance the quality and applicability of the findings. Additionally, interventions delivered in control groups should be better measured and accounted for in analyses.

Research funders play a crucial role in developing a comprehensive plan for future trials and data collection. They can facilitate this by commissioning the organisation of international

conferences or a series of meetings for the scientific and service user communities, and subsequently requiring changes in funded projects. Funders should also ensure that future trials include appropriate consent to allow data sharing for meta-analysis.

References

1. Department of Health. *Preventing suicide in England: A cross-government outcomes strategy to save lives*. 2012. Accessed 26/09/2024.
2. Gillies D, Christou MA, Dixon AC, Christou E, Kabir NA, Christou PA. Prevalence and Characteristics of Self-Harm in Adolescents: Meta-Analyses of Community-Based Studies 1990–2015. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2018;57(10):733-741. doi:10.1016/j.jaac.2018.06.018
3. Hawton K, Bale L, Brand F, et al. Mortality in children and adolescents following presentation to hospital after non-fatal self-harm in the Multicentre Study of Self-harm: a prospective observational cohort study. *Lancet Child & Adolescent Health*. 2020;4(2):111-120. doi:10.1016/S2352-4642(19)30373-6
4. Hawton K, Saunders KEA, O'Connor RC. Self-harm and suicide in adolescents. *Lancet*. 2012;379(9834):2373-2382. doi:10.1016/S0140-6736(12)60322-5
5. Hawton K, Harriss L. Deliberate self-harm in adolescent: Characterisics and subsequent mortality in a 20-year old cohort of patients presenting to hospital. *Journal of Clinical Psychiatry*. 2007;68:1574-1583. doi:10.4088/JCP.v68n1017
6. Hawton K, Bergen H, Waters K, et al. Epidemiology and nature of self-harm in children and adolescents: findings from the multicentre study of self-harm in England. *European child & adolescent psychiatry*. 2012;21(7):369-377. doi:10.1007/s00787-012-0269-6
7. Harris LM, Huang X, Funsch KM, Fox KR, Ribeiro JD. Efficacy of interventions for suicide and self-injury in children and adolescents: a meta-analysis. *Sci Rep*. Jul 19 2022;12(1):12313. doi:10.1038/s41598-022-16567-8
8. Owens D, Kelley R, Munyombwe T, et al. Switching methods of self-harm at repeat episodes: Findings from a multicentre cohort study. *Journal of affective disorders*. 2015;180:44-51. doi:10.1016/j.jad.2015.03.051
9. National Institute for H, Care E. *Self-harm in over 8s: long-term management. Clinical Guideline 133*. 2011. Accessed 26/09/2024.
10. Witt KG, Hetrick SE, Rajaram G, et al. Interventions for self-harm in children and adolescents. *Cochrane Database of Systematic Reviews*. 2021;(1469-493X (Electronic))doi:10.1002/14651858.CD013667.pub2
11. Kothgassner OD, Robinson K, Goreis A, Ougrin D, Plener PL. Does treatment method matter? A meta-analysis of the past 20 years of research on therapeutic interventions for self-harm and suicidal ideation in adolescents. *Borderline personality disorder and emotion dysregulation*. 2020;7:1-16. doi:10.1186/s40479-020-00123-9
12. Cottrell DJ, Wright-Hughes A, Eisler I, et al. Longer-term effectiveness of systemic family therapy compared with treatment as usual for young people after self-harm: An extended follow up of pragmatic randomised controlled trial. *EClinicalMedicine*. 2020;18:100246. doi:10.1016/j.eclinm.2019.10024610.1016/j.eclinm.2019.100246. eCollection 2020 Jan.
13. National_Institute_for_Health_and_Care_Excellence. *Self-harm: assessment, management and preventing recurrence*. 2022. Accessed 26/09/2024.
14. Hawton K, Witt KG, Taylor-Salisbury TL, et al. Psychosocial interventions for self-harm in adults. *The Cochrane Library*. 2016;Issue 5: CD012189. doi:10.1002/14651858.CD012189.pub2.

15. Fox KR, Huang X, Guzmán EM, et al. Interventions for suicide and self-injury: A meta-analysis of randomized controlled trials across nearly 50 years of research. *Psychological bulletin*. 2020;146(12):1117. doi:10.1037/bul0000305
16. Riley RD. Meta-analysis of individual participant data: rationale, conduct, and reporting. *British Medical Journal*. 2010;340:c221doi:10.1136/bmj.c221
17. Bahji A, Pierce M, Wong J, Roberge JN, Ortega I, Patten S. Comparative Efficacy and Acceptability of Psychotherapies for Self-harm and Suicidal Behavior Among Children and Adolescents: A Systematic Review and Network Meta-analysis. *JAMA Netw Open*. 2021;4(2574-3805 (Electronic))doi:10.1001/jamanetworkopen.2021.6614
18. Wright-Hughes A, Walwyn R, Wright JB, et al. Reducing Self-harm in Adolescents. An individual participant data meta-analysis (RISA-IPD): systematic review protocol. *BMJ OPen*. 2021;11(5): e049255(2044-6055 (Electronic))doi:10.1136/bmjopen-2021-049255
19. Cottrell D W-HA, Farrin A, Walwyn R, Mughal F, Truscott A, Diggins E, Irving D, Fonagy P, Ougrin D, Stahl D & Wright J. . Reducing Self-harm in Adolescents: the RISA-IPD individual patient data meta-analysis and systematic review. Methods, Studies and Risk of Bias. *Health Technology Assessment*. 2024;doi:https://doi.org/10.3310/GTNT6331
20. Cottrell D, Irving D, Farrin AA-O, et al. Reducing self-harm in adolescents: an individual participant data meta-analysis. 2022. https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=152119
21. Stewart LA, Clarke M, Rovers M, et al. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA* 2015;(1538-3598 (Electronic))doi:10.1001/jama.2015.3656
22. Higgins JPT, Savovic J, Page M, Elbers R, Sterne J. Chapter 8. Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, et al, eds. *Cochrane Handbook for Systematic Reviews of Interventions version 63 (updated February 2022)*. Cochrane; 2022.
23. SAS_Institute_Inc. *SAS/Access 9.4 Interface to ADABAS*. SAS Institute; 2013.
24. StataCorp. *Statistical Software: Release 17.* StataCorp LLC; 2021.
25. Chen H, Cohen P, Chen S. How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Communications in Statistics - Simulation and Computation*. 2010/03/31 2010;39(4):860-864. doi:10.1080/03610911003650383
26. Hughes JL, Asarnow JR. Enhanced Mental Health Interventions in the Emergency Department: Suicide and Suicide Attempt Prevention in the ED. *Clinical pediatric emergency medicine*. 2013;14(1):28-34. doi:10.1016/j.cpem.2013.01.002
27. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Statistics in Medicine*. 2017;36(855-875)doi:10.1002/sim.7141
28. Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993;80(1):27-38. doi:10.1093/biomet/80.1.27
29. Langan D, Higgins JPT, Jackson D, et al. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Res Synth Methods*. Mar 2019;10(1):83-98. doi:10.1002/jrsm.1316
30. IntHout J, Ioannidis JP, Borm GFJBmrm. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. 2014;14(1):25. doi:10.1186/1471-2288-14-25
31. Riley RD, Simmonds MC, Look MPJJoce. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. 2007;60(5):431. e1-431. e12. doi:10.1016/j.jclinepi.2006.09.009

32. Riley RD, Debray TP, Fisher D, et al. Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: statistical recommendations for conduct and planning. *Statistics in medicine*. 2020;39(15):2115-2137. doi:10.1002/sim.8516
33. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*. 1997;315:629. doi:10.1136/bmj.315.7109.629
34. Asarnow JR, Baraff LJ, Berk M, et al. An emergency department intervention for linking pediatric suicidal patients to follow-up mental health treatment. Empirical Study; Interview; Quantitative Study; Treatment Outcome/Clinical Trial. *Psychiatric Services*. Nov 2011;62(11):1303-1309. doi:http://dx.doi.org/10.1176/appi.ps.62.11.1303
35. Asarnow JR, Hughes JL, Babeva KN, Sugar CA. Cognitive-Behavioral Family Treatment for Suicide Attempt Prevention: A Randomized Controlled Trial. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2017;56(6):506-514. doi:10.1016/j.jaac.2017.03.015
36. Brent DA, Greenhill LL, Compton S, et al. The Treatment of Adolescent Suicide Attempters Study (TASA): Predictors of Suicidal Events in an Open Treatment Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2009;48(10):987-996. doi:10.1097/CHI.0b013e3181b5dbe4
37. Carter GL, Clover K, Whyte IM, Dawson AH, D'Este C. Postcards from the EDge: 24-month outcomes of a randomised controlled trial for hospital-treated self-poisoning. *British Journal of Psychiatry*. 2007;191:548-553. doi:10.1192/bjp.bp.107.038406
38. Chanen AM, Jackson HJ, McCutcheon LK, et al. Early intervention for adolescents with borderline personality disorder using cognitive analytic therapy: randomised controlled trial. *British Journal of Psychiatry*. Dec 2008;193(6):477-84. doi:10.1192/bjp.bp.107.048934
39. Cooney E, New Z, Ministry of H, Wise G, Te Pou o te Whakaaro N. *Feasibility of Evaluating DBT for self-harming adolescents : a small randomised controlled trial*. 2010. Accessed 26/09/2024.
40. Cotgrove A, Zirinsky L, Black D, Weston D. Secondary Prevention of Attempted-Suicide in Adolescence. *J Adolescence*. Oct 1995;18(5):569-577. doi:10.1006/jado.1995.1039
41. Cottrell DJ, Wright-Hughes A, Collinson M, 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. *The Lancet Psychiatry*. 2018;5(3):203-216. doi:10.1016/S2215-0366(18)30058-0
42. Diamond GS, Wintersteen MB, Brown GK, et al. Attachment-Based Family Therapy for Adolescents with Suicidal Ideation: A Randomized Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2010;49(2):122-131. doi:10.1016/j.jaac.2009.11.002
43. Diamond GS, Kobak RR, Krauthamer Ewing ES, et al. A Randomized Controlled Trial: Attachment-Based Family and Nondirective Supportive Treatments for Youth Who Are Suicidal. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2019;58(7):721-731. doi:10.1016/j.jaac.2018.10.006
44. Donaldson D, Spirito A, Esposito-Smythers C. Treatment for adolescents following a suicide attempt: Results of a pilot trial. *Journal of the American Academy of Child and Adolescent Psychiatry*. Feb 2005;44(2):113-120. doi:10.1097/00004583-200502000-00003
45. Esposito-Smythers C, Spirito A, Kahler CW, Hunt J, Monti P. Treatment of Co-Occurring Substance Abuse and Suicidality Among Adolescents: A Randomized Trial. *Journal of Consulting and Clinical Psychology*. Dec 2011;79(6):728-739. doi:10.1037/a0026074

46. Esposito-Smythers C, Hadley W, Curby TW, Brown LK. Randomized pilot trial of a cognitive-behavioral alcohol, self-harm, and HIV prevention program for teens in mental health treatment. *Behaviour research and therapy*. 2017;89:49-56. doi:10.1016/j.brat.2016.11.005
47. Esposito-Smythers C, Wolff JC, Liu RT, et al. Family-focused cognitive behavioral treatment for depressed adolescents in suicidal crisis with co-occurring risk factors: a randomized trial. *J Child Psychol Psychiatry*. 2019;60(10):1133-1141. doi:10.1111/jcpp.13095
48. Green JM, Wood AJ, Kerfoot MJ, et al. Group therapy for adolescents with repeated self harm: randomised controlled trial with economic evaluation. *British Medical Journal*. 2011;342doi:10.1136/bmj.d682
49. Griffiths H, Duffy F, Duffy L, et al. Efficacy of Mentalization-based group therapy for adolescents: the results of a pilot randomised controlled trial. *BMC Psychiatry*. 2019;19(1):167. doi:10.1186/s12888-019-2158-8
50. Hassanian-Moghaddam H, Sarjami S, Kolahi AA, Lewin T, Carter G. Postcards in Persia: a Twelve to Twenty-four Month Follow-up of a Randomized Controlled Trial for Hospital-Treated Deliberate Self-Poisoning. *Archives of suicide research*. 2017;21(1):138-154. doi:10.1080/13811118.2015.1004473
51. Harrington R, Kerfoot M, Dyer E, et al. Randomized trial of a home-based family intervention for children who have deliberately poisoned themselves. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1998;37(5):512-518. doi:10.1016/S0890-8567(14)60001-0
52. Hatcher S, Sharon C, Parag V, Collins N. Problem-solving therapy for people who present to hospital with self-harm: Zelen randomised controlled trial. *British Journal of Psychiatry*. 2011;199(4):310-316. doi:10.1192/bjp.bp.110.090126
53. Hatcher S, Sharon C, House A, Collins N, Collings S, Pillai A. The ACCESS study: Zelen randomised controlled trial of a package of care for people presenting to hospital after self-harm. *British Journal of Psychiatry*. 2015;206(3):229-236. doi:10.1192/bjp.bp.113.135780
54. Hazell PL, Martin G, McGill K, et al. Group therapy for repeated deliberate self-harm in adolescents: failure of replication of a randomized trial. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2009;48(6):662-670. doi:10.1097/CHI.0b013e3181a0acec
55. Huey SJ, Henggeler SW, Rowland MD, et al. Multisystemic Therapy Effects on Attempted Suicide by Youths Presenting Psychiatric Emergencies. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2004;43(2):183-190. doi:10.1097/00004583-200402000-00014
56. Husain N, Afsar S, Ara J, et al. Brief psychological intervention after self-harm: randomised controlled trial from Pakistan. *British Journal of Psychiatry*. 2014;204(6):462-470. doi:10.1192/bjp.bp.113.138370
57. Kaess M, Edinger A, Fischer-Waldschmidt G, Parzer P, Brunner R, Resch F. Effectiveness of a brief psychotherapeutic intervention compared with treatment as usual for adolescent nonsuicidal self-injury: a single-centre, randomised controlled trial. *European Child and Adolescent Psychiatry*. 2019;doi:http://dx.doi.org/10.1007/s00787-019-01399-1
58. King CA, Kramer A, Preuss L, Kerr DC, Weisse L, Venkataraman S. Youth-Nominated Support Team for Suicidal Adolescents (Version 1): a randomized controlled trial. *Journal of consulting and clinical psychology*. 2006;74(1):199-206. doi:10.1037/0022-006x.74.1.199 10.1037/0022-006X.74.1.199.
59. King CA, Klaus N, Kramer A, Venkataraman S, Quinlan P, Gillespie B. The Youth-Nominated Support Team-Version II for Suicidal Adolescents: A Randomized Controlled

Intervention Trial. *Journal of Consulting and Clinical Psychology*. Oct 2009;77(5):880-893. doi:10.1037/a0016552

60. McCauley E, Berk MS, Asarnow JR, et al. Efficacy of Dialectical Behavior Therapy for Adolescents at High Risk for Suicide: A Randomized Clinical Trial. *JAMA psychiatry*. 2018;75(8):777-785. doi:10.1001/jamapsychiatry.2018.1109

61. Mehlum L, Tormoen AJ, Ramberg M, et al. Dialectical Behavior Therapy for Adolescents With Repeated Suicidal and Self-harming Behavior: A Randomized Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2014;53(10):1082-1091. doi:10.1016/j.jaac.2014.07.003

62. Morthorst B, Krogh J, Erlangsen A, Alberdi F, Nordentoft M. Effect of assertive outreach after suicide attempt in the AID (assertive intervention for deliberate self harm) trial: randomised controlled trial. *British Medical Journal*. 2012;345doi:10.1136/bmj.e4972

63. O'Connor RC, Ferguson E, Scott F, et al. A brief psychological intervention to reduce repetition of self-harm in patients admitted to hospital following a suicide attempt: a randomised controlled trial. *The lancet psychiatry*. 2017;4(6):451-460. doi:10.1016/S2215-0366(17)30129-3

64. Ougrin D, Boege I, Stahl D, Banarsee R, Taylor E. Randomised controlled trial of therapeutic assessment versus usual assessment in adolescents with self-harm: 2-year follow-up. *Archives of Disease in Childhood*. Oct 2013;98(10):772-6. doi:10.1136/archdischild-2012-303200

65. Pineda J, Dadds MR. Family intervention for adolescents with suicidal behavior: a randomized controlled trial and mediation analysis. *J Am Acad Child Adolesc Psychiatry*. Aug 2013;52(8):851-62. doi:10.1016/j.jaac.2013.05.015

66. Robinson J, Yuen HP, Gook S, et al. Can receipt of a regular postcard reduce suicide-related behaviour in young help seekers? A randomized controlled trial. *Early Intervention in Psychiatry*. 2012;6:145-152. doi:10.1111/j.1751-7893.2011.00334.x

67. Rossouw TI, Fonagy P. Mentalization-based treatment for self-harm in adolescents: a randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2012;51(12):1304-1313. e3. doi:10.1016/j.jaac.2012.09.018

68. Santamarina Perez P, Romero Cela S, Mendez Blanco I, et al. Efficacy of dialectical behavior therapy compared to supportive therapy in adolescents with suicidal behavior. *European neuropsychopharmacology*. 2017;27:S853-S854. doi:10.1016/j.euroneuro.2016.09.012

69. Santamarina-Perez P, Mendez I, Singh MK, et al. Adapted Dialectical Behavior Therapy for Adolescents with a High Risk of Suicide in a Community Clinic: A Pragmatic Randomized Controlled Trial. *Suicide Life Threat Behav*. 2020;50(3):652-667. doi:10.1111/sltb.1261210.1111/sltb.12612. Epub 2020 Jan 16.

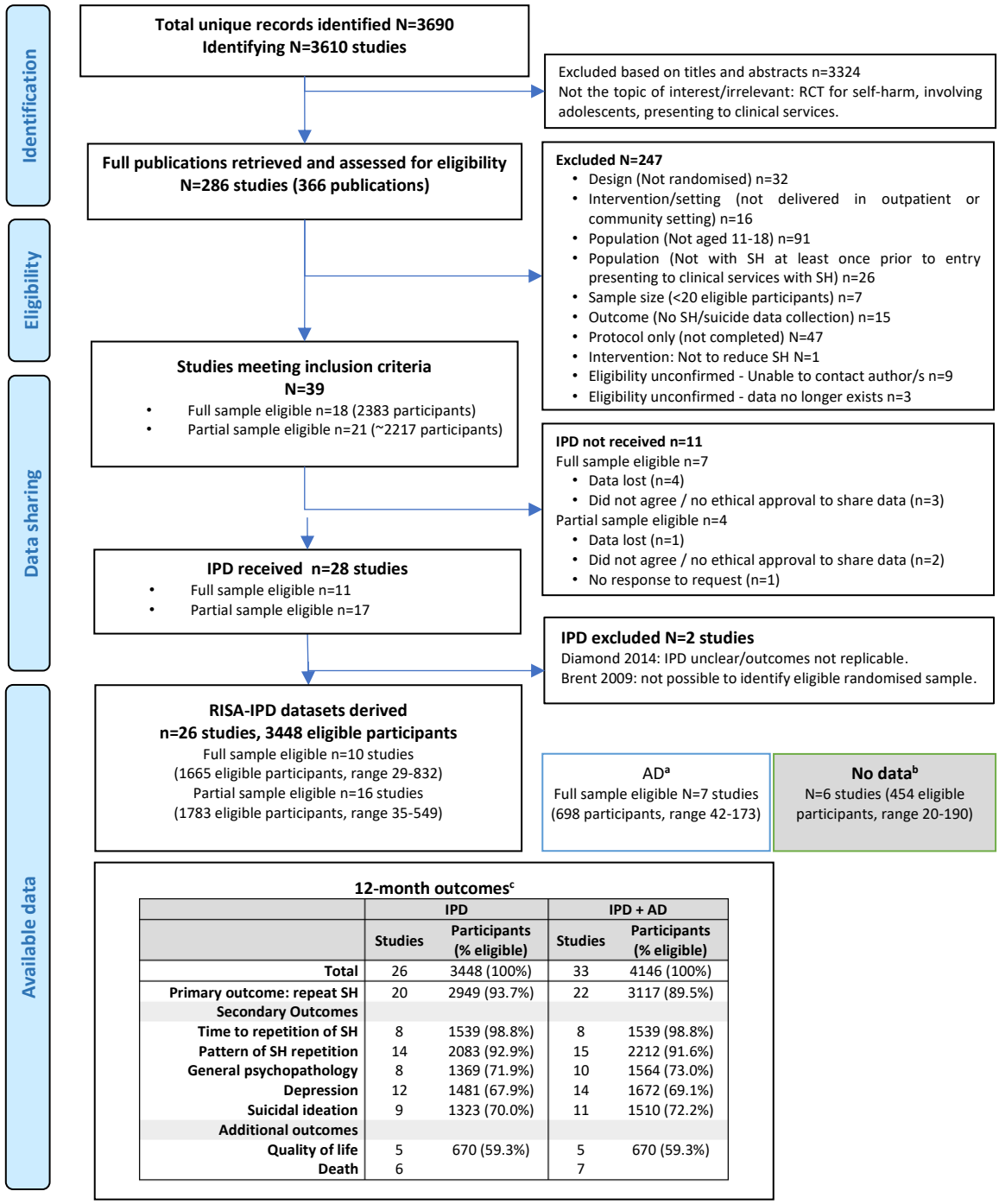
70. Spirito A, Boergers J, Donaldson D, Bishop D, Lewander W. An intervention trial to improve adherence to community treatment by adolescents after a suicide attempt. *Journal of the American Academy of Child and Adolescent Psychiatry*. Apr 2002;41(4):435-442. doi:10.1097/00004583-200204000-00016

71. Tyrer P, Thompson S, Schmidt U, et al. Randomized controlled trial of brief cognitive behaviour therapy versus treatment as usual in recurrent deliberate self-harm: the POPMACT study. *Psychological Medicine*. 2003;33(6):969-976. doi:10.1017/S0033291703008171

72. Wood A, Trainor G, Rothwell J, Moore A, Harrington R. Randomized trial of group therapy for repeated deliberate self-harm in adolescents. *Journal of the American Academy*

of Child & Adolescent Psychiatry. 2001;40(11):1246-1253. doi:10.1097/00004583-200111000-00003
73. Hawton K, Witt KG, Taylor-Salisbury TL, et al. Interventions for self-harm in children and adolescents. . *Cochrane Database of Systematic Reviews* 2015;(12):Art. No.: CD012013. doi:DOI: 10.1002/14651858.CD012013
74. Mehlum L, Ramleth RK, Tormoen AJ, et al. Long term effectiveness of dialectical behavior therapy versus enhanced usual care for adolescents with self-harming and suicidal behavior. *J Child Psychol Psychiatry*. 2019;60(10):1112-1122. doi:10.1111/jcpp.13077
75. Lewis SP, Hasking PA. Self-injury recovery: A person-centered framework. *J Clin Psychol*. Apr 2021;77(4):884-895. doi:10.1002/jclp.23094
76. Owens C, Fox F, Redwood S, et al. Measuring outcomes in trials of interventions for people who self-harm: qualitative study of service users' views. *BJPsych Open*. 2020;6(e22):1-8. doi:10.1192/bjo.2019.93
77. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A "SMART" Design for Building Individualized Treatment Sequences. *Annual Review of Clinical Psychology*. 2012;8(1):21-48. doi:10.1146/annurev-clinpsy-032511-143152
78. Woodcock J, LaVange LM. Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both. *New England Journal of Medicine*. Jul 6 2017;377(1):62-70. doi:10.1056/NEJMra1510062
79. Collins LM, Kugler KC. Optimization of behavioral, biobehavioral, and biomedical interventions. *Cham: Springer International Publishing*. 2018;10(1007):978-3. doi:10.1007/978-3-319-72206-1

Figure 1: PRISMA Flow Chart



12-month outcomes[c]

	IPD		IPD + AD	
	Studies	Participants (% eligible)	Studies	Participants (% eligible)
Total	26	3448 (100%)	33	4146 (100%)
Primary outcome: repeat SH	20	2949 (93.7%)	22	3117 (89.5%)
Secondary Outcomes				
Time to repetition of SH	8	1539 (98.8%)	8	1539 (98.8%)
Pattern of SH repetition	14	2083 (92.9%)	15	2212 (91.6%)
General psychopathology	8	1369 (71.9%)	10	1564 (73.0%)
Depression	12	1481 (67.9%)	14	1672 (69.1%)
Suicidal ideation	9	1323 (70.0%)	11	1510 (72.2%)
Additional outcomes				
Quality of life	5	670 (59.3%)	5	670 (59.3%)
Death	6		7	

Note: AD = aggregate data; IPD = individual participant data.
[a] AD indicate studies where published summary statistics were available for inclusion in the meta-analysis (i.e. for studies in which the full sample were eligible for RISA-IPD).
[b] No data indicate studies where no data were available for inclusion in the meta-analysis as reliable published statistics were not available for the RISA-IPD eligible sample of participants.
[c] The IPD and AD available, for 26 and 7 studies respectively, included data on a range of outcomes and timepoints therefore the number of studies with data available at the primary 12-month time-point represents a subset of all studies with data available.

Figure 2: Forest Plot of the Effect of Intervention versus Control on the Primary Outcome Self-Harm Repetition at 12 Months

See Upload Fig2_panel 1	See Upload Fig2_panel 2	See Upload Fig2_panel 3

Note: Panel 1 = Adjusted IPD meta-analysis. Panel 2 = Unadjusted IPD + AD meta-analysis. Panel 3 = Unadjusted IPD + AD meta regression. Colour coding is to highlight studies where estimates are from aggregate data in purple. AD = aggregate data. CBT = cognitive–behavioural therapy; DBT = dialectical behaviour therapy; IPD = Individual participant data; MBT/CAT = mentalisation based, psychodynamic, cognitive analytic therapy; MST = multi-systemic therapy; PST = problem solving, psychoeducation, support.

Figure 3: Forest Plot of the Effect of Intervention versus Control nn General Psychopathology at 12 Months

See Upload Fig3_Panel 1	See Upload Fig3_Panel 2

Note: Panel 1 = Adjusted IPD meta-analysis. Panel 2 = Unadjusted IPD + AD meta-analysis. Colour coding is to highlight studies where estimates are from aggregate data in purple. AD = aggregate data. CBT = cognitive–behavioural therapy; DBT = dialectical behaviour therapy; IPD = Individual participant data; MBT/CAT = mentalisation based, psychodynamic, cognitive analytic therapy; MST = multi-systemic therapy; PST = problem solving, psychoeducation, support.

Figure 4: Forest Plot of the Differential Effects for Intervention versus Control According to Participants Number of Self-Harm Episodes

See Uploads Fig4_Panel 1, Fig4_Panel 2, Fig4_Panel 3, Fig4_Panel 4, Fig4_Panel 5, Fig4_Panel 6

Note: Panel 1 = Repetition of self-harm within 6 months. Panel 2 = Repetition of self-harm within 12 months. Panel 3 = Repetition of self-harm between 6 to 12 months. Panel 4 = General Psychopathology. Panel 5 = Depression. Panel 6 = Suicidal ideation. The x-axis presents the odds ratio (OR). See Supplement 5, available online, for associated subgroup effects. CBT = cognitive–behavioural therapy; DBT = dialectical behavior therapy; MBT/CAT = mentalisation based, psychodynamic, cognitive analytic therapy; MST = multi-systemic therapy; PST = problem solving, psychoeducation, support.

Table 2: Study Characteristics by Data Availability

	IPD (n=26) N (%)	AD (n=7) N (%)	No Data (n=6) N (%)	Total (n=39) N (%)
Eligibility				
Full sample eligible	10 (38.5)	7 (100.0)	1 (16.7)	18 (46.2)
Partial eligible	16 (61.5)	0 (0.0)	5 (83.3)	21 (53.8)
N Eligible participants				
Mean (SD)	132.6 (187.54)	99.7 (50.06)	75.7 (63.06)	117.9 (156.68)
Median (Range)	69.0 (29, 832)	77.0 (42, 173)	63.0 (20, 190)	70.0 (20, 832)
Total	3448	698	454	4600
Age (years)[a]				
Weighted mean	15.7	14.8	15.4	15.5
Range (of study means)	14.5, 18.3	14.2, 15.6	12.9, 18.6	12.9, 18.6
SD	1.6	--	--	--
Female participants (%)[b]				
Weighted mean	82.0	88.9	63.7	79.6
Range (of study means)	65.7, 95.9	78.0, 94.8	35.0, 80.0	35.0, 95.9
Years since primary publication				
Mean (SD)	9.3 (4.79)	15.3 (9.89)	12.7 (3.78)	10.9 (6.18)
Median (Range)	10.0 (2.0, 19.0)	20.0 (4.0, 27.0)	12.0 (8.0, 18.0)	11.0 (2.0, 27.0)
Country[c]				
US	7 (26.9)	3 (42.9)	5 (83.3)	15 (38.5)
Rest of world	19 (73.1)	4 (57.1)	1 (16.7)	24 (61.5)
Pilot/Feasibility or Effectiveness trial				
Pilot/Feasibility	5 (19.2)	2 (28.6)	2 (33.3)	9 (23.1)
Effectiveness	21 (80.8)	5 (71.4)	4 (66.7)	30 (76.9)
Study powered				
No	9 (34.6)	4 (57.1)	4 (66.7)	17 (43.6)
Yes	17 (65.4)	3 (42.9)	2 (33.3)	22 (56.4)
RISA Intervention				
CBT	7 (26.9)	1 (14.3)	2 (33.3)	10 (25.6)
DBT	2 (7.7)	2 (28.6)	0 (0.0)	4 (10.3)
Family Therapy	3 (11.5)	1 (14.3)	1 (16.7)	5 (12.8)
Group Therapy	2 (7.7)	1 (14.3)	0 (0.0)	3 (7.7)
CAT/MBT	3 (11.5)	0 (0.0)	0 (0.0)	3 (7.7)
MST	0 (0.0)	0 (0.0)	1 (16.7)	1 (2.6)
PST	5 (19.2)	0 (0.0)	1 (16.7)	6 (15.4)
Postcards/tokens	3 (11.5)	1 (14.3)	1 (16.7)	5 (12.8)
Brief-intervention	1 (3.8)	1 (14.3)	0 (0.0)	2 (5.1)
RISA Control Group[d]				
TAU	18 (69.2)	4 (57.1)	3 (50.0)	25 (64.1)
E-TAU	5 (19.2)	2 (28.6)	1 (16.7)	8 (20.5)
Active	3 (11.5)	1 (14.3)	2 (33.3)	6 (15.4)
Treatment Intensity				
Low	4 (15.4)	2 (28.6)	2 (33.3)	8 (20.5)
Medium	16 (61.5)	3 (42.9)	1 (16.7)	20 (51.3)
High	6 (23.1)	2 (28.6)	2 (33.3)	10 (25.6)
Unknown	0 (0.0)	0 (0.0)	1 (16.7%)	1 (2.6)
Treatment Duration (weeks)				
Mean (SD)	21.3 (16.77)	15.2 (10.40)	23.3 (16.23)	20.6 (15.70)
Median (Range)	14.0 (1, 52)	15.5 (0, 26)	21.0 (4, 52)	16.0 (0, 52)
Number of treatment sessions				
Mean (SD)	17.1 (18.95)	19.4 (19.71)	14.2 (11.45)	17.2 (17.95)
Median (Range)	10.0 (0, 64)	10.0 (0, 52)	16.0 (0, 26)	10.0 (0, 64)
Group element in treatment (Yes)	6 (23.1)	3 (42.9)	0 (0.0)	9 (23.1)
Family element in treatment (Yes)	14 (53.8)	5 (71.4)	5 (83.3)	24 (61.5)
Method of data collection: self-harm				
Self-report /Researcher interview	15 (57.7)	5 (71.4)	6 (100.0)	26 (66.7)
Hospital / medical records	8 (30.8)	2 (28.6)	0 (0.0)	10 (25.6)
Combined approach	3 (11.5)	0 (0.0)	0 (0.0)	3 (7.7)
Overall Risk of Bias (on the primary outcome)				
Low	6 (23.1)	0 (0.0)	0 (0.0)	6 (15.4)
Some concerns	17[d] (65.4)	5 (71.4)	3 (50.0)	25 (64.1)
High	3 (11.5)	2 (28.6)	3 (50.0)	8 (20.5)

Note: AD = aggregate data; CBT = cognitive–behavioural therapy; DBT = dialectical behavior therapy; IPD = individual participant data; MBT/CAT = mentalisation based, psychodynamic, cognitive analytic therapy; MST=multi-systemic therapy; PST = problem solving, psychoeducation, support; SD = standard deviation; TAU = treatment as usual.

[a] For partially eligible studies with IPD the mean age and % female participants is for the eligible sample, otherwise data are for the full recruited sample. Participants with IPD range in age from 11-18.9 years.

[b] Meta regression based on: Primary outcome data collection self-report (inc. combined) vs medical records.

[c] Seven studies (26.9%) were conducted in the US, 7 (26.9%) in Australasia, 7 (26.9%) in the UK, 3 (11.5%) elsewhere in Europe, and 2 (7.7%) in Asia. Additional AD for 7 studies included 3 US studies, 3 UK studies and 1 study in Norway.

[d] Active controls included: Antidepressant pharmacotherapy, Family-enhanced nondirective supportive therapy, Supportive relationship treatment, Hospitalisation, individual and group supportive therapy, TAU + group session. Enhanced TAU was described as such in original papers and also included standardised good clinical care, and intensive TAU.

Table 3: Summary of Participant Moderators in studies with Individual Participant Data

Potential Moderator	n (%) Studies with moderator available	Participants	Values	n (%) participants[a]
Key moderators				
Age	26 (100)	3437	Mean (SD)	15.7 (1.6)
Gender	26 (100)	3448	Female participant	2823 (81.9)
			Male participant	619 (18.0)
Depression (Clinically indicated)	15 (57.7)	1989	Clinical diagnosis	1138 (57.2)
			No clinical diagnosis	771 (38.8)
Presenting self-harm method	14 (53.8)	2779	Self-injury	1052 (37.9)
			Self-poisoning	1471 (52.9)
			Combined	245 (8.8)
Borderline personality disorder (BPD)	7 (26.9)	921	Clinical diagnosis	163 (17.7)
			No clinical diagnosis	682 (74.0)
Additional and emerging moderators explored in meta-analyses				
Number of previous self-harm episodes[b]	21 (80.8)	3213	<=2	1271 (39.6)
			Multiple	1874 (59.3)
Anxiety disorder (Clinical or Questionnaire indicated)	14 (53.8)	1682	Yes	917 (54.5)
			No	632 (37.6)
Family dysfunction	5 (19.2)[d]	966	Indicated	747 (77.3)
			Not indicated	205 (21.2)
Ethnicity	18 (69.2)	2482	Black, Asian, Other ethnicity	505 (20.3)
			White	1947 (78.4)
Suicidal ideation	15 (57.7)	2221	Indicated	1620 (72.9)
			Not indicated	546 (24.6)
Psychotropic medication use	12 (46.2)[c]	2198	Yes	325 (14.8)
			No	1534 (69.8)
Additional and emerging moderators available in <~50% of studies or not included in further analysis				
Baseline self-harm Outcome/severity	18 (69.2)	2643	NA not possible to harmonise[d]	
Identify as LGBTQ	2 (7.7)	223	Yes	74 (33.2)
			No	121 (54.3)
Autistic Spectrum Disorder	1 (3.8)	35	Indicated	0
			Not indicated	35 (100)
History of Abuse (physical, sexual)	9 (34.6)	2148	Yes	645 (30.0)
			No	1379 (64.2)
Eating disorder	7 (26.9)	762	Yes	52 (6.8)
			No	508 (66.7)
Intellectual disability	3 (11.5)	985	Yes	55 (5.6)
			No	912 (92.6)
Out of home placement	9 (34.6)	2394	Parents/guardians	2056 (85.9)
			Foster care/other	236 (9.9)
Physical health problem	4 (15.4)	1488	Yes	387 (26.0)
			No	1098 (73.8)

Note: Additional continuous moderators not presented due to use of different scales, standardised prior to analysis: depression in 17 (65.4%) studies, family dysfunction in 5 (19.2%) studies, suicidal ideation score in 14 (53.8%) studies, unemotional/ callous traits in one (3.8%) study. Study level summary in Supplement 3, available online.

[a] Unless otherwise indicated. Where % does not add to 100% this is due to missing participant level data.

[b] Due to variability in data collection methods within each study, we categorised the number of previous self-harm episodes as ≤two or multiple for analysis (rather than one, two, multiple episodes as had been planned, Supplement 3 Table S3.16)

[c] Included in analysis as present in 100% family therapy studies.

[d] Due to with considerable variability in data collection and definitions further synthesis of these data was not carried out.

Table 4: Summary of 12-Month Meta-Analysis of Treatment Effect Estimates

	Adjusted IPD random effects meta-analysis					Unadjusted IPD+AD random effects meta-analysis				
	Studies	N	Pooled effect (95% CI)	I^2, p-value	T^2 (95% CI)	Studies	N	Pooled effect (95% CI)	I^2, p-value	T^2 (95% CI)
Repeat Self-harm (randomization to 12-months)			Odds Ratio					Odds Ratio		
CBT	7	483	1.45 (0.77, 2.73)	20.1%, 0.277	0 (0, 1.52)	7	483	1.40 (0.79, 2.47)	3.3%, 0.401	0.00 (0, 1.07)
DBT	1	28	1.31 (0.15, 11.26)	NA	NA	1	28	1.54 (0.16, 14.54)	NA	NA
Family therapy	1	826	1.13 (0.81, 1.58)	NA	NA	1	826	1.13 (0.81, 1.58)	NA	NA
Group therapy	2	437	1.28 (0.00, 6545.76)	72.2%, 0.058	0.67 (0, 117.02)	3	500	0.70 (0.02, 23.84)	76.2%, 0.015	1.49 (0, 35.28)
CAT/MBT	3	190	0.71 (0.10, 4.84)	25.2%, 0.263	0.08 (0, 11.27)	3	190	0.78 (0.16, 3.93)	8.3%, 0.336	0.00 (0, 8.27)
PST	3	339	0.94 (0.29, 3.03)	0.0%, 0.742	0 (0, 3.37)	3	339	0.97 (0.30, 3.14)	0.0%, 0.670	0.00 (0, 3.63)
Postcards/tokens	2	577	1.03 (0.02, 65.50)	0.0%, 0.810	0 (0, 16.94)	3	682	0.91 (0.25, 3.27)	0.0%, 0.623	0.00 (0, 3.45)
Brief intervention	1	69	0.47 (0.12, 1.89)	NA	NA	1	69	0.47 (0.11, 1.94)	NA	NA
Overall	**20**	**2949**	**1.06 (0.86, 1.31)**	**1.1%, 0.443**	**0 (0, 0.20)**	**22**	**3117**	**1.02 (0.82, 1.27)**	**12.1%, 0.299**	**0 (0, 0.21)**
	Heterogeneity between groups: p=0.779					Heterogeneity between groups: p=0.759				
Time to SH			Hazard Ratio					Hazard Ratio		
CBT	1	50	1.85 (0.78, 4.42)	NA	NA	1	50	1.73 (0.73, 4.10)	NA	NA
Family therapy	1	828	1.03 (0.83, 1.28)	NA	NA	1	828	1.04 (0.84, 1.28)	NA	NA
Group therapy	2	439	1.14 (0.14, 9.55)	41.1%, 0.193	0.03 (0, 8.27)	2	439	1.17 (0.10, 14.01)	53.9%, 0.141	0.05 (0, 10.54)
PST	1	46	0.84 (0.27, 2.60)	NA	NA	1	46	0.92 (0.30, 2.86)	NA	NA
Postcards/tokens	2	107	1.25 (0.00, 497.19)	0.0%, 0.594	0 (0, 26.93)	2	107	1.18 (0.00, 410.28)	0.0%, 0.537	0 (0, 28.41)
Brief intervention	1	69	0.68 (0.25, 1.86)	NA	NA	1	69	0.72 (0.27, 1.93)	NA	NA
Overall	**8**	**1539**	**1.07 (0.90, 1.26)**	**0.0%, 0.693**	**0 (0, 0.15)**	**8**	**1539**	**1.07 (0.90, 1.27)**	**0.0%, 0.706**	**0 (0, 0.15)**
	Heterogeneity between groups: p = 0.722					Heterogeneity between groups: p = 0.813				
Pattern of Self-harm (self-harm between 6-12 months)			Odds Ratio					Odds Ratio		
CBT	4	216	0.90 (0.18, 4.67)	30.3%, 0.230	0.29 (0, 10.17)	4	216	0.92 (0.21, 4.11)	21.3%, 0.282	0.21 (0, 8.68)
DBT	1	24	1.69 (0.30, 9.52)	NA	NA	2	153	0.66 (0.01, 42.97)	0.0%, 0.333	0.00 (0, 47.00)
Family therapy	1	825	0.93 (0.59, 1.45)	NA	NA	1	825	0.93 (0.59, 1.46)	NA	NA
Group therapy	2	436	1.28 (0, 463.54)	69.0%, 0.072	0.31 (0, 56.24)	2	436	1.25 (0, 319.24)	67.1%, 0.081	0.27 (0, 49.89)
CAT/MBT	3	188	0.81 (0.11, 5.99)	29.0%, 0.244	0.22 (0, 10.15)	3	188	0.85 (0.08, 8.72)	51.4%, 0.128	0.45 (0, 14.23)
PST	1	258	0.51 (0.19, 1.35)	NA	NA	1	258	0.49 (0.18, 1.33)	NA	NA
Postcards/tokens	1	67	0.40 (0.05, 2.87)	NA	NA	1	67	0.50 (0.07, 3.72)	NA	NA
Brief intervention	1	69	1.27 (0.24, 6.89)	NA	NA	1	69	1.40 (0.25, 7.81)	NA	NA
Overall	**14**	**2083**	**0.93 (0.71, 1.23)**	**4.2%, 0.405**	**0 (0, 0.43)**	**15**	**2212**	**0.90 (0.69, 1.16)**	**9.0%, 0.352**	**0 (0, 0.45)**
	Heterogeneity between groups: p = 0.855					Heterogeneity between groups: p = 0.854				
General psychopathology			SMD					SMD		
CBT	3	187	-0.16 (-0.79, 0.48)	0.0%, 0.492	0 (0, 0.96)	3	187	-0.16 (-0.80, 0.48)	4.1%, 0.352	0 (0, 1.34)
DBT	0	0				1	133	-0.07 (-0.41, 0.27)	NA	NA
Family therapy	1	465	-0.11 (-0.27, 0.04)	NA	NA	1	465	-0.19 (-0.37, -0.00)	NA	NA
Group therapy	2	397	-0.35 (-3.47, 2.78)	67.0%, 0.082	0.09 (0, 16.18)	3	459	-0.26 (-0.86, 0.35)	33.4%, 0.223	0.03 (0, 1.16)
CAT/MBT	1	64	-0.20 (-0.65, 0.24)	NA	NA	1	64	-0.08 (-0.57, 0.41)	NA	NA
PST	1	256	0.04 (-0.20, 0.27)	NA	NA	1	256	0.07 (-0.17, 0.32)	NA	NA
Overall	**8**	**1369**	**-0.13 (-0.25, -0.01)**	**5.7%, 0.386**	**0 (0, 0.08)**	**10**	**1564**	**-0.13 (-0.25, -0.02)**	**0.0%, 0.458**	**0 (0, 0.05)**
	Heterogeneity between groups: p = 0.609					Heterogeneity between groups: p = 0.558				
Depression			SMD					SMD		
CBT	5	286	-0.08 (-0.41, 0.25)	0.0%, 0.830	0 (0, 0.20)	5	286	-0.05 (-0.38, 0.28)	0.0%, 0.847	0 (0, 0.19)
DBT	0					1	133	-0.14 (-0.48, 0.20)	NA	NA
Family therapy	1	431	-0.04 (-0.22, 0.14)	NA	NA	1	431	-0.05 (-0.24, 0.14)	NA	NA
Group therapy	2	395	-0.19 (-2.33, 1.94)	44.6%, 0.179	0.03 (0, 8.50)	3	453	-0.13 (-0.74, 0.48)	34.2%, 0.219	0.02 (0, 1.24)
CAT/MBT	2	89	-0.40 (-3.28, 2.48)	0.0%, 0.735	0 (0, 16.16)	2	89	-0.34 (-3.05, 2.37)	0.0%, 0.688	0 (0, 10.63)
PST	2	280	0.05 (-1.42, 1.53)	0.0%, 0.807	0 (0, 4.17)	2	280	0.03 (-1.49, 1.54)	0.0%, 0.399	0 (0, 10.75)
Overall	**12**	**1481**	**-0.07 (-0.18, 0.04)**	**0.0%, 0.784**	**0 (0, 0.04)**	**14**	**1672**	**-0.07 (-0.17, 0.04)**	**0.0%, 0.856**	**0 (0, 0.03)**
	Heterogeneity between groups: p = 0.432					Heterogeneity between groups: p = 0.746				
Suicide Ideation			SMD					SMD		
CBT	3	166	-0.18 (-1.10, 0.73)	27.2%, 0.253	0.05 (0, 2.35)	3	166	-0.20 (-1.08, 0.67)	25.0%, 0.264	0.04 (0, 2.14)
DBT	1	23	-0.15 (-1.01, 0.70)	NA	NA	2	153	-0.05 (-2.11, 2.01)	0.0%, 0.516	0.00 (0, 7.0)
Family therapy	1	459	-0.15 (-0.33, 0.02)	NA	NA	1	459	-0.14 (-0.33, 0.04)	NA	NA
Group therapy	2	395	-0.07 (-1.24, 1.10)	0.0%, 0.376	0 (0, 3.86)	3	452	-0.06 (-0.47, 0.34)	0.0%, 0.608	0 (0,0.47)
PST	2	280	-0.01 (-1.48, 1.46)	0.0%, 0.452	0 (0, 8.41)	2	280	0.02 (-1.50, 1.53)	0.0%, 0.505	0 (0, 8.21)
Overall	**9**	**1323**	**-0.09 (-0.21, 0.03)**	**0.0%, 0.734**	**0 (0, 0.05)**	**11**	**1510**	**-0.08 (-0.20, 0.03)**	**0.0%, 0.823**	**0 (0, 0.04)**
	Heterogeneity between groups: p=0.868					Heterogeneity between groups: p=0.822				

Note: Full results in Supplement 6, available online. AD = aggregate data; CBT = cognitive–behavioural therapy; DBT = dialectical behavior therapy; IPD = individual participant data; MBT/CAT = mentalisation based, psychodynamic, cognitive analytic therapy; MST=multi-systemic therapy; NA = not applicable; PST = problem solving, psychoeducation, support; SMD = standardised mean difference.