EFFICACY AND MECHANISM EVALUATION

VOLUME 5 ISSUE 7 DECEMBER 2018
ISSN 2050-4365

Radically open dialectical behaviour therapy for refractory depression: the RefraMED RCT

Thomas R Lynch, Roelie J Hempel, Ben Whalley, Sarah Byford, Rampaul Chamba, Paul Clarke, Susan Clarke, David Kingdon, Heather O'Mahen, Bob Remington, Sophie C Rushbrook, James Shearer, Maggie Stanton, Michaela Swales, Alan Watkins and Ian T Russell

NHS
National Institute for Health Research

DOI 10.3310/eme05070

Radically open dialectical behaviour therapy for refractory depression: the RefraMED RCT

Thomas R Lynch,[1]* Roelie J Hempel,[1] Ben Whalley,[2] Sarah Byford,[3] Rampaul Chamba,[4] Paul Clarke,[5] Susan Clarke,[6] David Kingdon,[7] Heather O'Mahen,[8] Bob Remington,[1] Sophie C Rushbrook,[6] James Shearer,[3] Maggie Stanton,[9] Michaela Swales,[10] Alan Watkins[11] and Ian T Russell[11]

[1]Department of Psychology, University of Southampton, Southampton, UK
[2]School of Psychology/Cognition Institute, Plymouth University, Plymouth, UK
[3]King's Health Economics, King's College London, London, UK
[4]Member of Trial Management Committee responsible for Public and Patient Involvement, Bilston, UK
[5]Institute for Social and Economic Research, University of Essex, Colchester, UK
[6]Intensive Psychological Therapies Service, Dorset HealthCare University NHS Foundation Trust, Poole, UK
[7]School of Medicine, University of Southampton, Southampton, UK
[8]Psychology, University of Exeter, Exeter, UK
[9]Psychological Services, Southern Health NHS Foundation Trust, Winchester, UK
[10]School of Psychology, Bangor University, Bangor, UK
[11]Medical School, Swansea University, Swansea, UK

*Corresponding author

Declared competing interests of authors: Thomas R Lynch reports grants from the University of Southampton during the conduct of the study and personal fees from New Harbinger Publications and R.O.B.T Ltd outside the submitted work. Roelie J Hempel is co-owner and director of Radically Open Ltd. No money was exchanged between Roelie J Hempel and the company before, during or immediately after the study. David Kingdon reports grants from the National Institute for Health Research Efficacy and Mechanism Evaluation programme during the conduct of the study. Heather O'Mahen reports personal fees from the Charlie Waller Institute and personal fees from Increasing Access to Psychological Therapies outside the submitted work and reports being a member of the Guidelines Development Group National Institute for Health and Care Excellence (NICE) 2014 Antenatal and Postnatal Mental Health, an executive member of the British Psychological Society Perinatal Faculty, a NICE expert advisor for the Centre for Clinical Practice and a member of the Clinical Reference Group for Perinatal Mental Health. Maggie Stanton reports personal fees from British Isles DBT Training, personal fees from Stanton Psychological Services Ltd and personal fees from the Routledge, Taylor and Francis Group outside the submitted work. Michaela Swales reports personal fees from British Isles DBT Training and personal fees from Guilford Press, Taylor and Francis and Oxford University Press outside the submitted work. Ian T Russell reports grants from Health and Care Research Wales during the conduct of the study.

Published December 2018
DOI: 10.3310/eme05070

This report should be referenced as follows:

Lynch TR, Hempel RJ, Whalley B, Byford S, Chamba R, Clarke P, *et al*. Radically open dialectical behaviour therapy for refractory depression: the RefraMED RCT. *Efficacy Mech Eval* 2018;**5**(7).

Efficacy and Mechanism Evaluation

ISSN 2050-4365 (Print)

ISSN 2050-4373 (Online)

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

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

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

Criteria for inclusion in the *Efficacy and Mechanism Evaluation* journal

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

EME programme

The Efficacy and Mechanism Evaluation (EME) programme was set up in 2008 as part of the National Institute for Health Research (NIHR) and the Medical Research Council (MRC) coordinated strategy for clinical trials. The EME programme is broadly aimed at supporting 'science driven' studies with an expectation of substantial health gain and aims to support excellent clinical science with an ultimate view to improving health or patient care.

Its remit includes evaluations of new treatments, including therapeutics (small molecule and biologic), psychological interventions, public health, diagnostics and medical devices. Treatments or interventions intended to prevent disease are also included.

The EME programme supports laboratory based or similar studies that are embedded within the main study if relevant to the remit of the EME programme. Studies that use validated surrogate markers as indicators of health outcome are also considered.

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

This report

The research reported in this issue of the journal was funded by the EME programme as project number 09/150/12. The contractual start date was in September 2011. The final report began editorial review in January 2017 and was accepted for publication in October 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The EME editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, NETSCC, the EME programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the EME programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

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

NIHR Journals Library Editor-in-Chief

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals)

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

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

Abstract

Radically open dialectical behaviour therapy for refractory depression: the RefraMED RCT

Thomas R Lynch,[1]* Roelie J Hempel,[1] Ben Whalley,[2] Sarah Byford,[3] Rampaul Chamba,[4] Paul Clarke,[5] Susan Clarke,[6] David Kingdon,[7] Heather O'Mahen,[8] Bob Remington,[1] Sophie C Rushbrook,[6] James Shearer,[3] Maggie Stanton,[9] Michaela Swales,[10] Alan Watkins[11] and Ian T Russell[11]

[1]Department of Psychology, University of Southampton, Southampton, UK
[2]School of Psychology/Cognition Institute, Plymouth University, Plymouth, UK
[3]King's Health Economics, King's College London, London, UK
[4]Member of Trial Management Committee responsible for Public and Patient Involvement, Bilston, UK
[5]Institute for Social and Economic Research, University of Essex, Colchester, UK
[6]Intensive Psychological Therapies Service, Dorset HealthCare University NHS Foundation Trust, Poole, UK
[7]School of Medicine, University of Southampton, Southampton, UK
[8]Psychology, University of Exeter, Exeter, UK
[9]Psychological Services, Southern Health NHS Foundation Trust, Winchester, UK
[10]School of Psychology, Bangor University, Bangor, UK
[11]Medical School, Swansea University, Swansea, UK

*Corresponding author t.lynch@soton.ac.uk

Background: About one-third of patients who are depressed do not respond to antidepressant medication (ADM) and traditional psychotherapy shows limited benefits. However, most randomised trials have excluded the most sick patients, especially with comorbid personality disorder. Radically open dialectical behaviour therapy (RO DBT) is a new treatment targeting emotionally overcontrolled personality, which is common in refractory depression.

Objective: To evaluate the efficacy, cost-effectiveness and therapeutic mechanisms of RO DBT for patients with refractory depression.

Design: The Refractory depression: Mechanisms and Efficacy of RO DBT (RefraMED) trial was a multicentre, parallel-group, randomised trial in which participants were randomised to receive either RO DBT plus treatment as usual (TAU) or TAU alone. Participants were assessed at 7, 12 and 18 months after randomisation. Therapeutic mechanisms were explored in causal analyses.

Setting: Participants were recruited from three secondary care NHS organisations in the UK: Dorset, Hampshire and North Wales.

Participants: Patients were eligible if they were aged $\geq$ 18 years, had a Hamilton Rating Scale for Depression (HRSD) score of at least 15, had a current diagnosis of major depressive disorder in the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition Axis I, were suffering either refractory or chronic depression and, in their current episode, had taken an adequate dose of ADM for at

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

least 6 weeks without relief. Patients who met criteria for dramatic-erratic personality disorder (Cluster B), bipolar disorder or psychosis or who had a primary diagnosis of substance dependence or abuse were excluded.

Interventions: The intervention, RO DBT, comprised 29 weekly individual sessions lasting 1 hour and 27 weekly skills classes lasting 2.5 hours. Participants allocated to TAU could access any treatment offered by the NHS or privately.

Main outcome measures: The reported HRSD score 12 months after randomisation and cost per quality-adjusted life-year (QALY).

Results: A total of 250 participants were allocated at random: 162 (65%) participants were randomised to RO DBT plus TAU and 88 (35%) were randomised to TAU. The difference between RO DBT and TAU in the primary outcome at 12 months was not statistically significant. Nevertheless, after 7 months, participants randomised to RO DBT had substantially, and significantly, reduced depressive symptoms, relative to TAU, by 5.40 HRSD points [standardised mean difference 1.03 points, 95% confidence interval (CI) 0.94 to 9.85 points; $p = 0.02$]. Thereafter, RO DBT remained the better treatment with net, but non-significant, reductions of 2.15 HRSD points (standardised mean difference 0.41 points, 95% CI –2.28 to 6.59 points; $p = 0.29$) after 12 months and 1.69 points (standardised mean difference 0.32 points, 95% CI –2.84 to 6.22 points; $p = 0.42$) after 18 months. Participants allocated to TAU could access any treatment offered by the NHS or privately. Participants randomised to RO DBT reported significantly better psychological flexibility and emotional coping than TAU participants; these differences increased over time. From the perspective of the NHS and personal social services, RO DBT was not cost-effective; the incremental cost-effectiveness ratio was £220,000 per QALY, which is considerably above the willingness-to-pay threshold of £30,000 set by the National Institute for Health and Care Excellence (NICE). RO DBT participants reported eight possible serious adverse reactions compared with none by TAU participants; however, we believe that this imbalance was a result of major differences in reporting opportunities. The Data Monitoring and Ethics Committee agreed that there was no reason to suspect that RO DBT was harmful.

Conclusions: Although RO DBT achieved the target effect size (Cohen's d) with a d of 0.4 at the primary end point of 12 months, this was not significant. RO DBT was not cost-effective by NICE criteria. Nevertheless, RO DBT enhanced psychological flexibility and emotional coping, which are potential mechanisms of change.

Limitations: Analysing only 190 participants, instead of the target of 245, reduced the statistical power of the trial.

Future work: Future studies should aim to refine RO DBT, by maintaining clinical effectiveness while reducing costs, and to evaluate it for a wider range of overcontrolled disorders over a longer period.

Trial registration: Current controlled trials ISRCTN85784627.

Funding: This project was funded by the Efficacy and Mechanism Evaluation (EME) programme, a MRC and NIHR partnership, and will be published in full in *Efficacy and Mechanism Evaluation*; Vol. 5, No. 7. See the NIHR Journals Library website for further project information.

Contents

List of tables **xiii**

List of figures **xv**

List of boxes **xvii**

List of abbreviations **xix**

Plain English summary **xxi**

Scientific summary **xxiii**

Chapter 1 Introduction **1**
Background and rationale 1
Treatment resistance 1
Defining refractory depression 1
The costs of refractory depression 1
Psychological intervention research for refractory depression 1
Major paradigm shift: refractory depression reconceptualised as personality dysfunction 3
Radically open dialectical behaviour therapy 3
Objectives 4
Efficacy 4
Cost-effectiveness 4
Mechanisms 5
Temporal patterns and precedents of change 6

Chapter 2 Methods **7**
Trial design 7
Sample size estimation 7
Adjusted sample size targets 7
Randomisation 8
Primary randomisation and allocation ratio 8
Secondary randomisation 8
Participants 8
Inclusion criteria 8
Exclusion criteria 8
Discontinuations 9
Procedures 9
Recruitment and eligibility screening 9
Eligibility interview 9
Assessment training and inter-rater reliability 9
Blinding 11
Reimbursement 11
Settings and locations of data collection 11
Electronic data capture system 11
Changes to methods and procedures after trial commencement 11
Ethics considerations 12

© Queen’s Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Study monitoring 12
Reporting of adverse events and study termination 12
Patient and public involvement 13
Interventions 13
Treatment as usual 13
Radically open dialectical behaviour therapy 14
Radically open dialectical behaviour therapy therapist training, supervision and adherence 14
Measures 15
Primary outcomes: depression symptoms 15
Secondary outcomes 15
Economic measures 16
Moderating variables 17
Mediating variables 18
Temporal sequencing measures 18
Therapeutic alliance 19
Therapist characteristics 19
Client Satisfaction Questionnaire 19
Statistical methods 19
Data processing and reduction 19
Statistical analysis 19
Primary and secondary outcome analyses: continuous variables 19
Outcome analyses: binary or categorical outcomes 20
Economic evaluation 21
Mechanisms analysis using instrumental variables 25
Process variables 26
Direct effects of allocation 30

Chapter 3 Results **31**
Recruitment 31
Recruitment conversion rates 31
Participant flow 31
Study withdrawals, treatment withdrawals and losses to follow-up 31
Inter-rater reliability 33
Demographic and baseline data 33
Demographics 33
Clinical backgrounds 33
Outcome of adaptive randomisation procedures 37
Provision of therapy 37
Patients' treatment attendance 37
Adverse events 38
Continuous outcomes 39
Number of participants analysed 39
Primary outcome: depression symptoms measured by the Hamilton Rating Scale for Depression 39
Other continuous outcomes 41
Other outcomes 45
Remission rates and reliable change 45
Clinical predictions for new patients 46
Health economics analyses 47
Resource use 47
Costs 49
Outcomes 50
Cost-effectiveness 50
Sensitivity analyses 52

Mechanisms analyses using instrumental variables 53
  Single-mediator model 53
  All-mediator model 54
Participant perspectives 54
  Radically open dialectical behaviour therapy was considered helpful 54
  Suggestion 1: increase treatment duration or provide follow-up sessions 55
  Suggestion 2: increase skills group duration or change format slightly 55
  Suggestion 3: simplify language and homework 55
  Suggestion 4: develop application or other supporting materials 55
  Open group structure 56

Chapter 4 Discussion **57**
Principal findings 57
  Depressive symptoms 57
  Personality matters 58
  Remission rates 58
  Suicide ideation 59
Mechanisms of change 59
  Instrumental variables analysis 59
  Psychological flexibility 59
  Emotional approach coping and expressiveness 59
Health economics 60
Strengths and limitations 61
  Generalisability of findings 61
  Therapist variation 61
  Sample size and design considerations 61
  Weak instrumental variables 62
Patient and public involvement 62
  Recruitment and membership 62
  Effect of patient and public involvement 62

Chapter 5 Conclusions **63**
Implications for health care 63
Recommendations for future research 64
  Improving cost-effectiveness 64
  Skill consolidation after end of treatment 64
  Longer follow-up periods 64
  Transdiagnostic studies 64

Acknowledgements **65**

References **67**

Appendix 1 Additional tables **79**

Appendix 2 Trial procedure for reporting serious adverse events **85**

Appendix 3 Radically open dialectical behaviour therapy skills group lesson plan **97**

Appendix 4 Additional figures **99**

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of tables

TABLE 1 Assessment schedule and overview of measures used at each time point during the study (in months) 10

TABLE 2 Unit costs of health and social services 22

TABLE 3 Valuation of the RO DBT therapists' time 23

TABLE 4 Valuation of RO DBT group attendance 24

TABLE 5 Demographic data per treatment group allocation 34

TABLE 6 Basic clinical data per treatment group allocation 35

TABLE 7 Adverse events reported to trial manager 38

TABLE 8 Planned contrasts between TAU and RO DBT groups at 7, 12 and 18 months for depression scores (as measured using the HRSD), psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ) 41

TABLE 9 Descriptive statistics of depression scores (as measured using the HRSD), psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ) presented per group per month 42

TABLE 10 Model-adjusted means by group at each time point for depression scores (as measured using the HRSD), psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ) 44

TABLE 11 Remission rates (%) per group and follow-up month using model-adjusted and raw data 45

TABLE 12 The prognosis for new patients choosing RO DBT vs. TAU 47

TABLE 13 Resource use at baseline and at 7, 12 and 18 months' follow-up by group 48

TABLE 14 Disaggregated mean costs by group at baseline and each follow-up 49

TABLE 15 The EQ-5D-3L scores and QALYs over the 18-month follow-up period 51

TABLE 16 Summary of sensitivity analyses 52

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 17 Estimates of the total, direct and indirect effects using single-mediator models 53

TABLE 18 Estimates of the direct and indirect effects mediated through exposure using the all-mediators *Equations 4–7* 54

TABLE 19 The effect sizes for RO DBT for TRD from pilot studies (for group differences in change on the HRSD score from baseline to the end of treatment) 79

TABLE 20 A breakdown of study and treatment withdrawals 79

TABLE 21 Age (years) by site and group 79

TABLE 22 Basic clinical data per site 80

TABLE 23 The number of treatment attendances by site 82

TABLE 24 The RO DBT group session attendance across treatment sites 82

TABLE 25 The variance explained at each level of the primary outcome model with 95% highest probability density interval 82

TABLE 26 The range of expected prognoses for patients receiving TAU or RO DBT based on final HRSD score 82

TABLE 27 The expected additional change in symptoms for the 5th to 95th centiles of patients when treated with RO DBT rather than TAU, based on final HRSD score 83

TABLE 28 Types of AEs 86

TABLE 29 Event types, actions and deadlines 95

TABLE 30 Radically open dialectical behaviour therapy skills group lesson plan 97

List of figures

FIGURE 1 The causal model of primary interest showing the hypothesised relationships between treatment allocation (RO DBT vs. TAU), exposure, alliance, skills and treatment outcome (Y), allowing for unknown variables (U) 26

FIGURE 2 The Consolidated Standards of Reporting Trials (CONSORT) flow diagram showing the flow of study participants throughout the study and the numbers analysed per assessment point 32

FIGURE 3 Model-adjusted predictions for each treatment group based on HRSD score 40

FIGURE 4 The expected benefits for new patients accepting RO DBT rather than TAU at 7, 12 and 18 months 46

FIGURE 5 A scatterplot of differences in costs vs. differences in QALYs for RO DBT vs. TAU at 12 months' follow-up, taking the NHS/personal social services perspective 51

FIGURE 6 The cost-effectiveness acceptability curve for QALYs showing the probability that RO DBT is cost-effective compared with TAU at 12 months' follow-up, taking the NHS/personal social services perspective 52

FIGURE 7 Flow chart of safety reporting 89

FIGURE 8 Histograms showing the distribution of RefraMED patients across deciles for geographical indices of deprivation, by group 99

FIGURE 9 The non-monotonic pattern of non-attendance of individual therapy sessions by site 101

FIGURE 10 The percentage of available continuous data (text labels indicate the number of data points available for analysis) 102

FIGURE 11 The relative differences in therapist effectiveness 104

FIGURE 12 The AAQ-II 104

FIGURE 13 The EAC 105

FIGURE 14 The LIFE-RIFT 106

FIGURE 15 The MSSI 107

FIGURE 16 The SSQ 108

FIGURE 17 Adjusted (a–f) and raw (g–l) data remission rates by group 109

FIGURE 18 The expected outcome for patients choosing RO DBT vs. TAU 111

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

FIGURE 19 A scatterplot of differences in costs vs. differences in HRSD scores for RO DBT vs. TAU at 12 months' follow-up taking the NHS/personal social services perspective **111**

FIGURE 20 Cost-effectiveness acceptability curve showing the probability that RO DBT is cost-effective compared with TAU for different values that a decision-maker might be willing to pay for a point of improvement in HRSD scores **111**

List of boxes

BOX 1 Details of analyses 20

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of abbreviations

A&E	accident and emergency
AAQ-II	Acceptance and Action Questionnaire-II
ADM	antidepressant medication
AD-SUS	Adult Service Use Schedule
AE	adverse event
app	application
BDI	Beck Depression Inventory®
BDI-II	Beck Depression Inventory®-II
BPD	borderline personality disorder
CAT	cognitive–analytic therapy
CBASP	Cognitive Behavioural Analysis System of Psychotherapy
CBT	cognitive–behavioural therapy
CD	compact disc
CI	confidence interval
CMHT	Community Mental Health Team
CoBalT	Cognitive Behavioural Therapy as an adjunct to pharmacotherapy for primary care patients with treatment-resistant depression: a randomised controlled trial
DBT	dialectical behavior therapy
DBT-WCCL	Dialectical Behavior Therapy Ways of Coping Checklist
DMEC	Data Monitoring and Ethics Committee
DNA	did not attend
DSM-IV	*Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition
DVD	digital versatile disc
EAC	Emotional Approach Coping scale
EQ-5D-3L	EuroQoL Questionnaire-5 Dimensions, three-level version
GMM	generalised method of moments
GP	general practitioner
HRSD	Hamilton Rating Scale for Depression
IAPT	Increasing Access to Psychological Therapies
ICER	incremental cost-effectiveness ratio
IPT	interpersonal psychotherapy
IV	instrumental variable
LIFE-RIFT	The Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool
LTPP	long-term psychoanalytic psychotherapy
MREC	Medical Research and Ethics Committee
MSSI	Modified Scale for Suicidal Ideation
NICE	National Institute for Health and Care Excellence
PANAS	Positive and Negative Affect Scale
PD	personality disorder
PSSRU	Personal Social Services Research Unit
QALY	quality-adjusted life-year
RCT	randomised controlled trial
Re-ChORD	Relief of Chronic Or Resistant Depression
RefraMED	Refractory depression: Mechanisms and Efficacy of RO DBT
RO	radical openness
RO DBT	radically open dialectical behaviour therapy
RO DBT-E	earlier versions of radically open dialectical behaviour therapy
SAE	serious adverse event
SAR	serious adverse reaction

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

SCID	Structured Clinical Interviews for DSM-IV	SUSAR	suspected unexpected serious adverse reaction
SCID-I	Structured Clinical Interview for DSM-IV Axis I Disorders	TAU	treatment as usual
		TMG	Trial Management Group
SCID-II	Structured Clinical Interview for DSM-IV Axis II	TRD	treatment-resistant depression
SD	standard deviation	TSC	Trial Steering Committee
SSQ	Social Support Questionnaire		

Plain English summary

We evaluated the clinical effectiveness and cost-effectiveness of radically open dialectical behaviour therapy (RO DBT), which is a new treatment for depression and problems of emotional overcontrol, such as misguided perfectionism and emotional loneliness, by comparing its effects with treatment as usual (TAU), the health care normally available from the NHS or private providers. From three dissimilar UK locations, we recruited adults who were diagnosed with major depressive disorder who had not responded to treatment and who scored above the threshold on a standard measure of current depression. We randomly assigned them to either 7 months of RO DBT ($n = 162$ patients) or TAU ($n = 88$ patients). Recently trained therapists delivered RO DBT over 29 weekly 1-hour individual therapy sessions and 27 weekly 2.5-hour skills group sessions. We reassessed patients at the end of the treatment and at 12 and 18 months after entering the study. After 12 months, the main time at which we assessed clinical effectiveness, both groups reported worthwhile reductions in depression scores; however, the difference could have occurred by chance. At the end of treatment, in contrast, RO DBT patients had reported a reduction in depressive symptoms that was much larger by both statistical and clinical criteria than that reported by TAU patients. Furthermore, more of participants reported that they no longer felt depressed. The economic analyses concluded that, although RO DBT had reduced the extent and cost of other health care, it was too costly in its current form to achieve cost-effectiveness by published criteria. Hence, the next challenge is to find ways of delivering RO DBT more efficiently and more sustainably.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Scientific summary

Background

Major depressive disorder is a disabling condition causing substantial impairment in psychosocial functioning and quality of life – in the workplace, at home and with friends. Only two-thirds of patients respond to antidepressant medication (ADM) and only half respond to psychological treatment. Thus, treatment resistance is a common outcome in major depression. The personal, social and economic burden of depression is substantial, especially for treatment-resistant depression.

Research on psychological interventions for refractory depression, defined as chronic depression or recurrent depression with two or more previous episodes, is sparse; most studies focus on acute or episodic depression and pharmacological or somatic treatments. Historically, psychotherapy trials avoided patients with refractory depression because these treatments were not designed to meet their needs. More recently, treatments that were developed specifically for refractory depression have achieved small to moderate effect sizes.

One potential reason why treatments work for only some patients with refractory depression is comorbidity with other mental disorders, especially personality disorders (PDs). Current treatments neglect the role of emotional overcontrol in refractory depression. However, many adults with chronic depression show overcontrolled traits, including greater self-criticism, impaired autonomy, rigid internalised expectations, excessive control of spontaneous emotion and inordinate fear of making mistakes. Radically open dialectical behaviour therapy (RO DBT), a novel transdiagnostic psychotherapy, aims to address this rigid coping style and the associated emotional loneliness. Earlier versions of RO DBT showed promise in two small randomised trials of patients with refractory depression and comorbid PDs.

Objectives

The primary objective of this trial was to estimate the efficacy of RO DBT plus treatment as usual (TAU) for refractory depression compared with TAU alone. The relative cost-effectiveness of RO DBT plus TAU compared with TAU alone was also estimated. Furthermore, the study aimed to explore the mechanisms of RO DBT treatment and moderators of treatment efficacy using novel statistical methods that exploit instrumental variables. These novel statistical methods aim to ameliorate bias caused by unobserved confounding, which can distort conventional analyses.

Methods

The Refractory depression: Mechanisms and Efficacy of RO DBT (RefraMED) trial was a multicentre, parallel-group, randomised trial in which participants were randomised to receive either 7 months of RO DBT plus TAU or TAU alone. RO DBT comprised 29 weekly individual sessions lasting 1 hour and 27 weekly skills classes lasting 2.5 hours. Patients allocated to TAU could access any treatment offered by the NHS or privately.

Participants were recruited from three secondary care NHS organisations in the UK: in Dorset, Hampshire and North Wales. Patients were eligible if they were aged ≥ 18 years, had a Hamilton Rating Scale for Depression (HRSD) score of at least 15 points, had a current diagnosis of major depressive disorder in the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV) Axis I, were suffering either refractory or chronic depression and, in their current episode, had taken an adequate dose of ADM for at least 6 weeks without relief. Patients were excluded who met

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

criteria for dramatic–erratic PD (cluster B), bipolar disorder or psychosis or who had a primary diagnosis of substance dependence or abuse.

Once eligibility was confirmed and consent received, the Swansea Trials Unit used an adaptive algorithm to allocate participants at random to treatments and therapists. Three stratifying variables were used to ensure that there was balance between groups: (1) early or late onset of depression, (2) a HRSD score of < or > 25 points and, presence or absence of PDs. Trained assessors, who were blind to these allocations, assessed participants at 7, 12 and 18 months after randomisation. The primary outcome measure was the severity of depressive symptoms, as measured by the HRSD, at 12 months.

To aid interpretation of the primary outcome, the number of patients achieving a 'meaningful' reduction in symptoms was estimated, and full and partial depression remission rates were computed for each group. These remission criteria accounted for both depressive symptoms (evaluated using the HRSD) and quality of life [evaluated using the Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool (LIFE-RIFT)]. Other secondary outcome measures included suicidal ideation and behaviour. The cost-effectiveness of RO DBT was estimated by cost per quality-adjusted life-year (QALY). Causal analyses were conducted to explore the mechanisms by which RO DBT changes prognosis.

Results

In the trial a total of 250 participants were allocated by adaptive randomisation: 162 participants to RO DBT plus TAU and 88 participants to TAU. Patients reported substantial comorbidity: 96% met the criteria for at least one comorbid disorder, 86% for at least one DSM-IV Axis-I disorder and 78% for at least one DSM-IV Axis-II disorder. A total of 183 patients (73%) after 7 months, 190 (76%) patients after 12 months and 167 (67%) patients after 18 months provided data on the primary outcome. There was no significant difference between treatment groups in the proportion of participants who contributed data to the analysis.

The trial did not show a statistically significant difference between RO DBT and TAU for the primary outcome at 12 months. By 7 months, RO DBT had substantially and significantly reduced depressive symptoms relative to TAU by 5.40 HRSD points [standardised mean difference 1.03, 95% confidence interval (CI) 0.94 to 9.85 points; $p = 0.02$]. Although this advantage for RO DBT continued, the reduction was no longer statistically significant at 12 months (2.15 HRSD points, standardised mean difference 0.41 points; 95% CI –2.28 to 6.58 points; $p = 0.29$) or at 18 months (1.69 HRSD points, standardised mean difference 0.32 points, 95% CI –2.84 to 6.22 points; $p = 0.42$). However, the TAU group achieved surprising improvements from month 12.

Full remission rates were low in both groups: 1%, 8% and 7% for the RO DBT group and 0%, 0% and 1% in for TAU group at 7, 12 and 18 months, respectively. Partial remission rates were higher in the RO DBT group participants (23%, 26% and 33% at successive assessments) than in the TAU participants (6%, 22% and 24% at successive assessments). Other indicators of improvement also showed superiority for RO DBT. The proportion of participants achieving a 'reliable improvement' or a 'worthwhile' change from the participants' perspective – at least a 17.5% reduction in symptoms – was consistently higher for RO DBT group participants than for TAU group participants. To aid patients and clinicians in future treatment choices, and to set realistic expectations for probable outcomes, data were simulated from the primary outcome model. It was found that for every 100 new patients, 32 would experience 'worthwhile' improvements in symptoms after 12 months by choosing RO DBT rather than TAU.

We found significant advantages for RO DBT participants in both emotional approach coping [as assessed by the Emotional Approach Coping scale (EAC)] and psychological inflexibility [as assessed by the Acceptance and Action Questionnaire (AAQ-II)] throughout the trial. In the RO DBT group participants, the EAC mean score and the standardised mean difference between groups increased over time at month 7 by 1.50 EAC points (standardised mean difference 0.32 points, 95% CI 0.09 to 2.90 points; $p < 0.05$); at month 12 by 3.55 EAC points (standardised mean difference 0.76 points, 95% CI 2.14 to 4.95 points; $p < 0.001$); and at month 18 by 2.98 EAC points (standardised mean difference 0.64 points, 95% CI 1.14 to 4.82 points; $p < 0.01$). The mean AAQ-II scores decreased over time and all standardised mean differences were medium to large (month 7: –3.37 AAQ-II points, standardised mean difference 0.49 points, 95% CI –5.92 to –0.82 points; $p = 0.01$; month 12: –4.94 AAQ-II points, standardised mean difference 0.72 points; 95% CI –7.53 to –2.36 points; $p < 0.001$; and month 18: –5.48 AAQ-II points, standardised mean difference 0.79 points, 95% CI from –8.74 to 2.22 points; $p = 0.001$). However, we found no significant advantage for RO DBT group participants in impaired functioning (as assessed by the LIFE-RIFT), suicide ideation scores [as assessed by the Modified Scale for Suicidal Ideation (MSSI)] or perceived social support (as assessed by the Social Support Questionnaire). Both trial groups showed substantial reductions in LIFE-RIFT scores from baseline; mean MSSI scores remained below 8 points throughout the trial for both groups, indicating low levels of suicidal ideation, and, although perceived social support increased in the RO DBT group, the difference between groups was not significant.

The instrumental variables that we incorporated into the design of the study yielded estimates that were large, implausible and imprecise. Even conventional analyses, which assume no unobserved confounding, were unable to divide the total effect of RO DBT into direct and mediated components. Single-mediator models suggested that the pathway, either through learning skills or through therapeutic alliance, had improved participants' HRSD scores and the all-mediators model suggested that skills had played the greater role.

Primary cost-effectiveness analysis showed that the additional cost for RO DBT of £7050 (95% CI £5822 to £8274; $p < 0.001$) was associated with an increase of 0.032 QALYs (95% CI –0.028 to 0.093 QALYs; $p = 0.30$), yielding an incremental cost-effectiveness ratio (ICER) of £220,000 per QALY. This ICER was considerably above the willingness-to-pay threshold of £30,000 set by the National Institute for Health and Care Excellence (NICE) in the UK, but comparable with ICERs reported for standard dialectical behaviour therapy. Hence, from the perspective of the NHS and personal social services, RO DBT was not cost-effective relative to TAU alone in treating patients with refractory depression. RO DBT did not achieve sufficient gains in QALYs and savings in other health and social services to justify the cost of the resource-intensive therapy delivered in this trial. This was true when productivity losses were added.

In total, 33 serious adverse events (SAEs) were reported to the trial manager during the study: four for TAU participants, 28 for RO DBT participants and one for a participant before randomisation. It was judged that 17 SAEs (including all four in the TAU group) were definitely not related to RO DBT, eight were unlikely to be related, five were possibly related and three were probably related to RO DBT. Although none of the resulting eight possible serious adverse reactions (SARs) was classed as 'unexpected', all eight occurred in the RO DBT group, which was a statistically significant finding. However, TAU participants were seen by trial assessors at only the three follow-up interviews; in contrast, RO DBT participants were seen twice a week by trial therapists. We believe that the imbalance in SARs was because of this gross difference in reporting opportunities. Indeed, the Data Monitoring and Ethics Committee agreed that there was no reason to suspect that RO DBT was harmful.

Conclusions

The study found that RO DBT is much more efficacious than TAU in reducing participants' depressive symptoms over the course of treatment. After 7 months of treatment, participants in the RO DBT group show significantly fewer depressive symptoms by having an unusually large effect size (Cohen's d) with a d of 1.03. Thereafter, TAU patients also showed reductions in symptoms, reducing the effect size to a d of 0.41 at 12 months and a d of 0.32 at 18 months. Hence, differences between trial groups were not statistically significant at 12 or 18 months. However, the mean effect of RO DBT at 12 months reached the target effect size with a d of 0.4 and by exceeding 2 HRSD points, which is generally regarded as clinically relevant. Rates of full and partial remission were also consistently higher in RO DBT than TAU participants, especially after 7 months. Nevertheless, RO DBT is not cost-effective according to the NICE criteria.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Although the trial participants reported moderate to severe initial depression scores and considerable comorbidity, they reported low to moderate suicidal ideation throughout the study, with no significant differences between groups. Participants in the RO DBT group achieved significantly better psychological flexibility than those in the TAU group at 7, 12 and 18 months, and significantly better emotional approach coping and expressiveness after 7 months, with the difference increasing after 12 and 18 months. This suggests that patients continue to use and improve their RO DBT skills.

However, we could not draw firm conclusions from the analysis using instrumental variables. Standard diagnostics suggested that these variables were weak; therefore, the resulting estimates were biased and unreliable. However, as this was, to our knowledge, the first trial of psychotherapy to allocate instrumental variables at random, the experience will inform others.

Implications for health care

This study has demonstrated that depressive symptoms of patients receiving RO DBT decreased significantly during the 7 months of treatment and these reductions in symptoms were maintained at 12 and 18 months. However, the difference between the RO DBT and TAU groups did not achieve statistical significance after 12 and 18 months. This lack of significance was because of, in part, a smaller analysable sample at 12 months than was sought and an improvement in depressive symptoms after 7 months in the TAU group, who received more treatment between months 7 and 12 than the RO DBT group. Despite the higher cost of RO DBT than treatments typically offered within the NHS, it can improve depressive symptoms in a highly symptomatic population who suffer from many mental health problems and PDs and who may require specialised treatment for their problems.

Unlike other treatments for depression, RO DBT does not consider depression to be the primary problem. Instead, RO DBT targets emotional overcontrol – a maladaptive personality style shown to predict the development of chronic internalising disorders, such as resistant depression. Although overcontrolled PDs, including obsessive–compulsive PD, are more common than undercontrolled PDs, patients' innate capacity to tolerate distress, delay gratification and avoid public displays of emotion make their problems less noticeable. Overcontrolled individuals are likely to play down personal distress when queried and are, therefore, less likely to seek mental health treatment.

Implications for future research

The RefraMED trial was the first multisite trial of RO DBT. RO DBT improved depressive symptoms over time, and mediational analyses provided preliminary support for hypothesised mechanisms of change. Future studies should investigate variations on RO DBT that have the potential to be clinically effective but are less costly, for example by reducing the length of treatment or using stepped care that starts with group skills training and reserves one-to-one treatment for non-responsive patients. However, patients such as those who were studied by the RefraMED trial are difficult to treat and unlikely to benefit from short-term solutions. Although the RefraMED trial has investigated some mechanisms underpinning RO DBT, a more thorough investigation of which RO DBT skills are most clinically effective will be critical for developing a shorter version of the treatment, suitable for health-care systems with limited funding.

It is important to develop and test programmes that incorporate feedback from users and, thus, address tapering, online support, 'graduate groups' and other forms of continuing support. Given the recurring nature of depression, and the emphasis in RO DBT on changing maladaptive personality, future studies should extend assessment to investigate long-term differences between RO DBT and other treatments. Finally, the transdiagnostic approach of RO DBT supports testing RO DBT across a wider range of conditions, including eating disorders, overcontrolled PDs, anxiety disorders and autism spectrum disorders.

Trial registration

This trial was registered as ISRCTN85784627.

Funding

This project was funded by the Efficacy and Mechanism Evaluation (EME) programme, a MRC and NIHR partnership.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 1 Introduction

Background and rationale

Treatment resistance

Major depressive disorder is a disabling condition that causes substantial impairment in psychosocial functioning and quality of life, especially in the workplace, at home and with friends.[1] By 2030, depression will be the leading cause of disability in developed countries and the second leading cause of disability worldwide.[2]

Only 30–40% of individuals treated with antidepressant medication (ADM) achieve full remission and only half of individuals respond to psychological treatment.[3] About one-third of patients do not respond to ADM[4] and, of those non-responding patients, fewer than one-third benefit from adding, or switching to, cognitive therapy.[5,6] Approximately 50% of patients who are diagnosed with major depressive disorder experience recurrent or chronic illness needing long-term treatment.[7] Those with chronic depression are least likely to respond to currently available treatments.[8,9] Thus, treatment resistance is a common outcome for individuals with major depression and reported prevalence rates range from 15% to 60%, depending on how treatment-resistant depression (TRD) is defined.[7,10–12]

Defining refractory depression

Definitions of TRD and refractory depression vary across studies; for example, Berlim and Turecki[3] found > 10 different definitions of TRD. In practice, chronic depression and TRD often overlap, with many patients meeting both definitions. In the present study, we define 'refractory depression' as comprising both TRD, that is, depression that does not respond to adequate intervention, and chronic depression, that is, depression lasting > 2 years.

In a recent systematic review of TRD, Negt *et al.*[9] found that, compared with acute or episodic depression, chronic depression is associated with earlier age at onset, higher rates of adverse childhood experiences, more comorbid disorders and poorer social adjustment.

The costs of refractory depression

The economic burden of depression is substantial, especially for refractory depression. The total cost of adult depression across England in 2007 was > £7.5B, including £1.6B for patient care (treatment and social care) and £5.8B for loss of earnings.[13] Many of these costs are a result of refractory depression; for example, Crown *et al.*[7] found that depression-related costs for treatment-resistant inpatients were 19 times greater than those for patients not meeting criteria for treatment resistance and the costs for treatment-resistant outpatients were 2.5 times greater than for those not meeting the criteria. Similarly, in a US-based study, Ivanova *et al.*[14] reported that, after adjusting for differences in baseline comorbidities, risk-adjusted costs for employees with TRD were US$11,600 more than for depressed control employees who did not have TRD.

Psychological intervention research for refractory depression

Most studies investigating psychological interventions for depression have focused on acute or episodic depression rather than refractory depression.

For acute or episodic depression, cognitive–behavioural therapy (CBT), behavioural activation, interpersonal psychotherapy (IPT) and short-term psychodynamic psychotherapy have the strongest empirical support.[9] Meta-analyses of psychotherapies for non-chronic depression typically find effects (Cohen's d) relative to waiting list and minimal treatment control groups ranging from a d of 0.70 to a d of 0.90, whereas placebo control groups result in smaller effect sizes, averaging a d of 0.36.[15] A recent systematic review found weak evidence that behavioural therapies and other psychological therapies for acute depression are equally effective.[16]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The evidence for these treatments for chronic depression is much more limited.[9] Cuijpers *et al.*[15] conducted a meta-analysis of studies investigating psychological treatments for chronic depression or dysthymia. It included 16 studies published between 1966 and 2009, of which only four studied chronic major depression; the remainder studied dysthymia, double depression (both major depressive disorder and dysthymia), or other categories of chronic depression or dysthymia. Of the psychological treatments investigated, seven were CBT and six were IPT, in addition to eight other treatments, including the Cognitive Behavioural Analysis System of Psychotherapy (CBASP), problem-solving therapy, cognitive-interpersonal group psychotherapy for chronic depression and supportive therapy. The authors reported a small but significant effect size (Cohen's *d*) with a *d* of 0.23 for psychotherapy on depression symptom scores when compared with control groups. McPherson *et al.*[17] identified 12 studies investigating psychotherapeutic treatment for TRD, but only four were controlled trials and none recruited > 25 participants;[17] the effect sizes ranged from a *d* of 1.23 to a *d* of 3.10.

One of the reasons that there are fewer trials that investigate psychotherapy for refractory depression is because most established treatments have been developed for less severe forms of depression. However, recently, treatments and studies have been developed specifically for refractory depression. The most investigated treatment to date is CBASP, a behavioural analytic therapy focusing on behaviour and its consequences.[18] McCullough[18] hypothesises that chronically depressed individuals have deficits in social problem-solving and interpersonal communication, arising from early adverse events (AEs). Although the treatment was originally developed for individual therapy, CBASP has been modified for group formats.[19,20] A recent meta-analysis of six randomised trials,[19,21–25] studying 1510 patients, found a small effect size for CBASP, with a *d* of 0.34.[9] Remission rates ranged from 19% to 57% in the CBASP groups (median 35%), compared with 6–50% in the control groups (median 25%).

Another recently developed treatment for chronic depression is the Relief of Chronic Or Resistant Depression (Re-ChORD) programme, an intensive, time-limited outpatient programme lasting 4 months. The major components of the Re-ChORD programme are medication management, group-based IPT and group occupational therapy.[26] The Re-ChORD study was a parallel-group, open randomised trial comparing the Re-ChORD programme with treatment as usual (TAU). This study was somewhat underpowered because (1) the authors chose a binary primary outcome [clinical remission, defined as a Hamilton Rating Scale for Depression (HRSD) score of ≤ 7 points after 4 months] and (2) they recruited only 64 patients. As a result, the difference in the proportion of patients remitting after 4 months (36% in the Re-ChORD programme group vs. 13% in the TAU group) was only on the borderline of statistical significance.

The Tavistock Adult Depression Study[27] was a small pragmatic randomised trial evaluating long-term psychoanalytic psychotherapy (LTPP) as an adjunct to TAU for 129 patients with long-standing major depression who had failed to respond to at least two different treatments. Patients received either TAU alone or LTPP plus TAU and were assessed at 6-monthly intervals until 42 months, including the 18 months of treatment. Complete remission (i.e. a HRSD score of ≤ 8 points) was infrequent and never significantly different between groups. Partial remission (i.e. a HRSD score of ≤ 12 points) was significantly more likely in the LTPP plus TAU group than in the TAU-alone group (at 24 months, 39% vs. 19%; at 30 months, 35% vs. 12%; and at 42 months, 30% vs. 44%). Similarly, the difference between the group means became significant only at 24 months, when the LTPP plus TAU group started to show significantly larger decreases in HRSD scores than those shown in the TAU-alone group, with moderate effect sizes.

The CoBalT (Cognitive Behavioural Therapy as an adjunct to pharmacotherapy for primary care patients with treatment-resistant depression: a randomised controlled trial)[28] recruited 213 primary care patients with TRD to evaluate the clinical effectiveness of CBT as an adjunct to TAU, compared with TAU alone. The trial found that 28% of patients in the CBT group and 15% of those in the TAU group achieved remission [i.e. a Beck Depression Inventory®-II (BDI®-II) score of < 10 points]. The difference between groups in improving BDI-II scores had an effect size with a *d* of 0.53. A recently published follow-up study[29] reported that this effect size still had a *d* of 0.45 at around 40 months after the end of treatment (BDI-II scores at 6 months: CBT, 18.9 points and TAU, 24.5 points; BDI-II scores at 40 months: CBT, 19.2 points and TAU, 23.4 points). The CoBalT team concluded that CBT as an adjunct to usual care that includes ADM is clinically effective and cost-effective over the long term for individuals whose depression has not responded to pharmacotherapy.

In summary, immediately after treatment, CBASP had a small beneficial effect on depression,[9] whereas CBT and the Re-ChORD programme had moderate beneficial effects on depression.[26,28] Moreover, emerging evidence suggests that longer and more complex treatments may be needed for treatment-resistant and chronic forms of depression than for more acute forms of depression.[30]

Major paradigm shift: refractory depression reconceptualised as personality dysfunction

A major premise of the Refractory depression: Mechanisms and Efficacy of RO DBT (RefraMED) trial is that when psychopathology is long-standing and does not respond to efficacious first-line treatments, then broad-based personality dimensions and overlearned perceptual and regulatory biases are preventing change. Biotemperamental deficits, combined with damaging family environmental influences, are hypothesised to severely handicap openness and flexible responding, resulting in habitual overcontrol or undercontrol of socioemotional behaviour. This distinction shares features with the well-established division between internalising and externalising disorders.[31] Major depressive disorder exemplifies this perspective of limited openness and flexible responding, resulting in an overcontrolled coping style. This is further supported by studies reporting that 40–60% of unipolar depressed patients meet the criteria for comorbid personality disorder (PD),[32–34] the most common being paranoid, avoidant and obsessive–compulsive PDs, all of which are overcontrolled PDs that do not respond well to treatment.[8,32,35] However, most trials exclude the most severely unwell individuals, for example those with PDs who are known to respond less favourably to existing treatments for acute depression, notably CBT.[8]

Previous studies have reported that adults with chronic depression are characterised by overcontrolled traits, including rigid internalised expectations, excessive control of spontaneous emotion, greater self-criticism, inordinate fears of making mistakes and impaired autonomy.[34] Excessive self-control has been linked to social isolation, aloof interpersonal functioning, maladaptive perfectionism, disingenuous emotional expression and mental health problems such as anorexia nervosa, obsessive–compulsive PD and chronic depression.[36–40] However, current treatments neglect the potential mediating role of personality and maladaptive emotional overcontrol in refractory depression.

Radically open dialectical behaviour therapy

A novel transdiagnostic psychotherapy,[41,42] RO DBT was designed specifically to address maladaptive emotional overcontrol and associated loneliness. The neurobiosocial theory underlying RO DBT proposes that individuals presenting with problems of overcontrol are biotemperamentally predisposed to exhibit heightened threat sensitivity, diminished reward sensitivity and heightened capacities for self-control and detailed-focused processing.[41,42] These biotemperamental biases can be strengthened by family, cultural and environmental histories that value performance and self-control and by avoiding risk and masking emotions. A person with heightened threat sensitivity is more likely to scan new or unfamiliar circumstances for potential harm rather than reward: their sympathetic nervous system – responsible for defensive-arousal and flight–fight responses – is activated while the ventral vagal complex of their parasympathetic nervous system withdraws.[43] Unfortunately, as a result of withdrawal of the parasympathetic nervous system ventral vagal complex, facial expressions become frozen and the ability to express oneself flexibly is lost; in other words, prosocial co-operative social signalling becomes impaired.[44] People who are inexpressive or who show expressions that do not match their inner experiences are often perceived as inauthentic or untrustworthy by others, which may lead to social ostracism and emotional loneliness.[41,45,46] Consequently, RO DBT emphasises the importance of targeting social-signalling deficits and uniquely links the communicative and facilitative functions of emotional expressions to the formation of close social bonds and uses micromimicry as a means to understand and increase emphatic behaviours.[41,42,47] The skills that are taught in RO DBT focus on relaxing rigid inhibitory self-control, activating the social safety system, expressing context-appropriate emotions and vulnerable self-disclosure, practising self-enquiry to learn from new experiences and critical feedback, and reducing maladaptive social comparisons related to envy and bitterness.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Earlier versions of RO DBT (RO DBT-E) have shown promise in two small randomised trials[48,49] for patients with refractory depression and comorbid PD. The first study[49] explored the feasibility of a group intervention for TRD in chronically depressed older adults aged > 60 years. The treatment comprised 28 weeks attending a skills training group and weekly 30-minute telephone contact with an individual therapist, followed by 3 months in which telephone contact was made every 2 weeks and 3 months in which it was made every 3 weeks. The 17 participants who received RO DBT-E showed significantly greater improvements than the 17 control participants in self-rated and interviewer-rated depression ($d = 0.71$). Post-treatment interviewer ratings showed that 71% of RO DBT-E recipients met criteria for remission [i.e. a Beck Depression Inventory® (BDI) score of ≤ 9 points or a HRSD score of ≤ 7 points], but only 47% of control participants did so. After 6 months, the corresponding percentages were 75% and 31%, respectively. In addition, RO DBT-E recipients improved significantly in adaptive coping and dependency, whereas control participants did not.

The second trial[48] compared 24 weeks of both individual and group RO DBT-E plus ADM in 21 adults aged > 55 years who had PD and depression with ADM alone in 14 similar adults. To be included, participants had to demonstrate TRD prospectively by having a poor response to an 8-week course of researcher-controlled ADM. The RO DBT-E recipients showed significantly greater decreases in interpersonal sensitivity and aggression than control participants. At the end of the skills group, 71% of RO DBT-E recipients were in remission (i.e. had a HRSD score of < 10 points), compared with only 50% of control participants. Both groups of participants showed significant reductions in clinician-rated depression, with a between-group d of 0.85.

Objectives

The RefraMED trial aimed to conduct an appropriately powered randomised trial to evaluate the efficacy and investigate the mechanisms of RO DBT in patients with TRD who were, thus, difficult to treat. The study had five main objectives, as detailed in the following sections.

Efficacy

The primary objective was to estimate the efficacy of RO DBT for refractory depression in addition to TAU, compared with TAU alone, over the course of 18 months. This objective can be achieved in two ways: first, as the effect of being randomly allocated to RO DBT plus TAU, rather than TAU alone; and, second, as the effect of exposure to specific 'doses' of RO DBT, in which exposure was measured by adherence to RO DBT treatment protocols and zero exposure corresponded to TAU. For both methods, the primary outcome was depressive symptoms measured using HRSD score at 12 months (i.e. 5 months after the end of treatment). To interpret this outcome, the resulting rates of remission and measures of other symptoms were used, including suicidal ideation or behaviour and global functioning.

Conventional analysis 'by treatment allocated' was used to fulfil the first of these approaches. Although this analysis provided an unbiased pragmatic estimate of the effectiveness of RO DBT in clinical practice, this is not identical to 'efficacy' under ideal conditions because participants varied in their adherence to the recommended course of treatment, and, therefore, there was heterogeneity in exposure to RO DBT. To account for this, estimations of the causal effect of exposure to RO DBT were sought. In doing so, this allowed for the fact that attendance at therapy sessions occurred after randomisation and, thus, could be subject to confounding.

Cost-effectiveness

Analysis 'by treatment allocated' was used to estimate the relative cost–utility of RO DBT plus TAU in comparison with TAU alone over 12 months. We calculated cost per quality-adjusted life-year (QALY) using the EuroQoL Questionnaire-5 Dimensions, three-level version (EQ-5D-3L).[50] We also explored cost-effectiveness using HRSD score. The economic perspective was that of the NHS and personal social services, as preferred by the National Institute for Health and Care Excellence (NICE).[51] The addition of productivity losses was explored, resulting from time off work because of illness, in the sensitivity analysis.

Mechanisms

We extended the efficacy analysis, which addressed questions relating to the effect of allocation and exposure to treatment, to ask questions about how RO DBT might have been efficacious. The approach included both RO DBT-specific and transtheoretical concepts, and recognised that important elements of psychotherapy might be common to many treatments.

The primary aim was to estimate the causal effects of the pathways through which RO DBT is hypothesised to reduce depressive symptoms (as measured via the HRSD score) for patients in the trial. The chief challenge in doing this was the potential presence of unobserved confounding. Bias from unobserved confounding potentially arises because, unlike the treatment itself, patients are not randomly allocated to the number of sessions that they attend, the strength of the therapeutic alliances that they develop with their therapists or the degree to which they acquire new skills through engaging with the treatment. Patients make decisions to engage with treatment in a non-random manner that may be related to their underlying mental health. For example, suppose that patients with certain characteristics choose to attend as many sessions as possible, form strong alliances with their therapists and develop the skills needed for RO DBT; if these same characteristics are also associated with a milder form of refractory depression, then the statistical estimates of these pathways obtained using standard pathway models[52] will overstate the causal effect of each pathway because of uncontrolled confounding bias.

There are two main objectives of this analysis:

1. to specify a series of pathway models to decompose the overall effect of RO DBT into that resulting from the pathways through which the treatment is hypothesised to work
2. to estimate the decomposition of this treatment effect using:
    - standard adjustments for confounding variables
    - instrumental variables (IVs) to adjust for possible unobserved confounding.

The analyses focused on four specific pathways between allocation and outcome: treatment exposure, therapeutic alliance, skill acquisition and expectancy. To estimate the causal effect of mediators, the trial manipulated selected mediators in a random fashion. We also measured variables that represented sources of variation in mediators that were unlikely to be contaminated by selection effects. In other words, IVs were identified, both experimental and observational, to facilitate causal analyses.

In addition to these primary pathways, several potential modifiers of treatment outcomes were measured. Based on this study and other research showing links between refractory depression and PD,[32] temperamental risk aversion and reward insensitivity[53–55] and childhood adversity,[56] we assessed potential moderators of treatment response by measuring the following at baseline:

1. PD diagnosis
2. invalidating childhood experiences
3. reward sensitivity or risk aversion.

These variables will be used to conduct a complementary analysis of repeated measurements of outcomes and mediators using longitudinal models. This analysis will address theoretically driven questions about the ordering of changes in key variables.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Temporal patterns and precedents of change

In conjunction with the causal analyses of the therapeutic alliance, the study wanted to examine patterns of change and cross-lagged effects among depressive symptoms, alliance ratings and skills learned in RO DBT. In addition, whether or not there are differential rates of change in positive versus negative affect and whether or not different patterns of temporal ordering exist for positive affect and negative affect will be investigated. Previous research has highlighted the distinct and important role of positive affect in adaptive coping.[57] Positive affect is thought to broaden an individual's attentional focus and behavioural repertoire and, as a consequence, build social, intellectual and physical resources.[58] A change in positive affect is expected to be more rapid and more closely associated with factors common to psychosocial interventions (e.g. expectancy and alliance) than for change in negative affect and to precede improvements in coping strategy. Furthermore, ecological momentary assessment will enable questions to be answered in relation to the variability of affect in treated versus untreated patients. Daily variability in affect is expected to rise early in treatment for RO DBT participants (as a consequence of the difficult work clients undertake with therapists), but to decline relative to the TAU group patients by 7 months. Although these longitudinal analyses estimate temporal ordering or so-called 'Granger causality' rather than true causality, they complement the primary mediation analyses by providing a richer picture of patterns of change in response to treatment.

Chapter 2 Methods

Please note that for the sake of consistency, several paragraphs from the methods section of the published protocol paper in *BMJ Open*[59] have been reproduced. Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to http://group.bmj.com/group/rights-licensing/permissions. This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

Trial design

The RefraMED trial is a multicentre randomised controlled trial (RCT) conducted at three NHS sites in the UK: Dorset, Hampshire and North Wales. All trial participants received TAU, but those allocated, at random, to the experimental arm also received approximately 29 weeks of RO DBT. Participants were further allocated, at random, between non-therapeutic alternatives designed to facilitate mediational analyses. Clinical outcomes were assessed at four time points (baseline and at 7, 12 and 18 months after randomisation) using assessors blind to participants' allocated treatment.

Sample size estimation

Two pilot studies of RO DBT-E for refractory depression report effect size for group differences in change on the HDRS score (from baseline to the end of treatment) with a d of 0.71[49] and with a d of 0.85.[48] In addition, one trial of standard DBT for TRD[61] reports an effect size with a d of 1.45 (see *Table 19* in *Appendix 1* for more details). Therefore, it was judged feasible, and desirable, to recruit enough analysable participants to yield an 80% power to detect, at a statistically significant level of 5%, a standardised difference of 0.4 between groups (RO DBT and TAU). This equates to a between-group difference on the HRSD of > 2 points. Furthermore, it was judged that NICE would consider this difference clinically important.

Simulations of the random-effects models described in the following sections suggest that, if there were no intraclass correlation, a sample of 200 analysable participants would detect a mean difference of 2 points on the HRSD (equivalent to a standardised difference of 0.4). We expected to collect analysable data from at least 83% of participants and, therefore, increased the target to 240. To increase the power of the analysis of the mechanisms of RO DBT, it was planned to randomise in the ratio of 3 : 2, seeking to allocate 144 'unclustered' patients to RO DBT and 96 to TAU.

However, although participants in the RO DBT arm were clustered by therapist, the 96 control participants were not. To allow for an intracluster correlation coefficient of 0.025 between the HRSD scores and an average cluster size of 11 participants for each of the expected 16 therapists, we increased the RO DBT sample size to 180, yielding the same statistical power as 144 unclustered participants. Thus, initially, it was aimed to randomise 276 patients, namely 180 to RO DBT and 96 to TAU, giving an allocation ratio of 15 : 8.

Adjusted sample size targets

Recruitment rates in all centres were limited by available assessor time. The consequent difficulty in meeting target recruitment rates and a reduced therapist capacity in one centre forced there to be a re-evaluation of the target sample size. In consultation with the Data Monitoring and Ethics Committee (DMEC), we agreed to recalculate the sample size using less conservative assumptions and, consequently, reduced the target number of patients randomised to 240.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Randomisation

Primary randomisation and allocation ratio

Participants were allocated between treatments through an adaptive randomisation algorithm administered by the Swansea Trials Unit. This system of randomisation maintained the balance across groups of the three stratifying variables stochastically, rather than deterministically, chosen as potential outcome moderators to minimise the risk of subversion:[60]

1. early onset of depression (before or after 21 years of age)
2. depression severity at baseline (a HRSD score of $<$ or $\geq$ 25 points)
3. PD [meets the Structured Clinical Interview for DSM-IV Axis II (SCID-II) criteria for cluster A or cluster C or not].

Participants were allocated to RO DBT or TAU using a final allocation ratio of 15 : 8 in favour of RO DBT.

Secondary randomisation

Within the RO DBT group, participants were further randomly allocated to their therapist and the three IVs listed below. An adaptive randomisation method was used to allocate patients between therapists so as to use as many as feasible of the treatment slots at each centre. We also randomly allocated all RO DBT participants to one of eight combinations of contextual non-therapeutic features (i.e. IVs) within a 2 × 2 × 2 factorial design:

(a) the participants received individual RO DBT in either a standard therapy room or an enhanced therapy room that potentially improves both the strength of the alliance with the therapist and attendance at treatment sessions
(b) after every session the participants either did or did not have the opportunity to provide written feedback to their therapist, which potentially improves the strength of the alliance with the therapist[61]
(c) the participants received compensation for questionnaire completion either personally (via their therapists) or impersonally (via mail), potentially increasing attendance at treatment sessions.

The allocation sequence was determined dynamically using a database independently administered at the Swansea Trials Unit. The database incorporated a study-specific version of a published dynamic randomisation allocation method[62] and the resulting allocations were then sent by e-mail to the trial manager for further dissemination to participants and study therapists.

Participants

Inclusion criteria

Participants were considered eligible if they were aged $\geq$ 18 years, had a HRSD[63] score of at least 15 points (note that a HRSD score of 14 points is regularly used as the upper limit for partial remission), had a current diagnosis of major depressive disorder [as assessed with the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)[64]] and had refractory depression, which was defined as two or more previous episodes of depression or having experienced their current episode for $\geq$ 2 years. Finally, participants had to have taken an adequate dose of ADM for at least 6 weeks without symptom relief during their current episode.

Exclusion criteria

We excluded those participants who met criteria for dramatic–erratic PD (cluster B), bipolar depression, psychosis or a primary diagnosis of substance dependence or substance abuse disorder. Patients had to have an intelligence quotient of $>$ 70 and speak English well enough to participate in the treatment and the study. Finally, patients who were receiving standard DBT at the time of recruitment were not eligible to take part in the current study because standard DBT aims to treat borderline PD, which was one of the exclusion criteria.

Discontinuations

Participants were allowed to withdraw from the treatment or the study at any time without their regular care being affected. Treatment withdrawal was defined as participants being unable or unwilling to attend RO DBT sessions; on such occasions, we encouraged participants to provide follow-up data. Study withdrawals were defined as participants being unable or unwilling to attend follow-up sessions; on such occasions, we permitted participants to continue with RO DBT if they had been allocated to that treatment arm. However, if participants withdrew consent to participate in the study, we did not seek additional data.

Procedures

Recruitment and eligibility screening

Participants were recruited through mental health professionals, such as psychologists, psychiatrists and mental health nurses, through general practitioners (GPs) and via advertisements in secondary and primary care clinics (e.g. waiting rooms or GP surgeries) and in the community (e.g. libraries). We carried out database searches for potential participants in the primary and secondary care services, although GPs and secondary care clinicians could also refer patients directly to the study. In addition, patients could self-refer after seeing posters, leaflets or the website.

In the primary care settings, database searches were conducted by the GP practice staff; medical records were checked for eligibility using the inclusion and exclusion criteria and the resulting lists were screened by GPs. The GPs and practice staff then signed pre-prepared letters describing the study and inviting patients to contact their local RefraMED clinical studies officer if they were considering participating in the trial. Patients who called to express an interest provided oral consent to be screened for eligibility.

In secondary care, database searches were carried out by the local RefraMED clinical studies officers. The resulting lists were screened by the responsible clinicians before letters were sent out. Patients received letters describing the study and inviting them to opt out or consider participating in the trial. Unless patients opted out, the clinical studies officer tried to contact them by telephone to discuss the study and, with their oral consent, screened them for eligibility.

Recruitment started at the Dorset site in March 2012 and at the other two sites in September 2012, and ended at all sites in March 2015. The last follow-up interviews were completed in May 2016.

Eligibility interview

In-person appointments were made with trained assessors for the potential participants who were eligible at the telephone screening stage and willing to attend the study assessment. The participants signed a formal consent form on arrival. The team of trained assessors screened them for eligibility and conducted baseline assessments of the outcome measures set out in *Table 1* for those who were found to be eligible for study inclusion. After the interview, participants received further questionnaires for home completion.

Assessment training and inter-rater reliability

Assessors were trained to competence on the Structured Clinical Interviews for DSM-IV (SCID) and HRSD by a clinical psychologist who was experienced in administering the SCID and HRSD in clinical trials. Ongoing queries regarding the SCID and HRSD were discussed at weekly consensus meetings and inter-rates reliability on the HRSD was re-assessed at regular (9-month) intervals to ensure ongoing adherence with the HRSD.

Following the recommended training sequence of SCID, all assessors read through the SCID and HRSD and watched the SCID training digital versatile discs (DVDs). Assessors watched key sections of the SCID training tapes of interest for this study (i.e. mood episodes, anxiety disorders and DSM-IV Axis II) together

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 1 Assessment schedule and overview of measures used at each time point during the study (in months)

	Time point (months)													
	Baseline	Treatment						Follow-up						
Outcome	0	1	2	3	4	5	6	7	8	9	10	11	12	18
Primary														
Depression (assessed via the HRSD and LIFE-RIFT)	✓							✓					✓	✓
Health-related quality of life (assessed via the EQ-5D-3L)	✓			✓				✓					✓	✓
Health services use/costs (assessed via the AD-SUS)	✓							✓					✓	✓
Secondary														
Suicide (assessed via the MSSI and SBQ)	✓							✓					✓	✓
Depression and affect (assessed via the PHQ-9 and PANAS)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mechanisms and mediators														
Psychosocial function (assessed via the AAQ-II, WBSI, the 3-item SSQ, IIP-PD-25 and DBT-WCCL)	✓			✓				✓					✓	✓
Emotional approach and expectancy (assessed via the EAC and CEQ)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Alliance and delivery of treatment (assessed via the CALPAS, CSQ-8 and number of sessions attended)	✓	✓	✓	✓	✓	✓	✓	✓	✓					
Moderators of RO DBT effectiveness														
Personality and PDS (assessed via the CID-I and II, NEO FFI-C, applied conscientiousness task and SVS)	✓													
Temperament and emotional control (assessed via the UPPS, PNS, Ego-Undercontrol, Ego-Resiliency, BIDR-16 and FMPS)	✓													
Childhood experience and invalidation (assessed via the ICES and MOPS)	✓													

AAQ-II, Acceptance and Action Questionnaire-II; AD-SUS, Adult Service Use Schedule; BIDR-16, Balanced Inventory of Desirable Responding-Short Form; CALPAS, California Psychotherapy Alliance Scales; CEQ, Credibility/Expectancy Questionnaire; CSQ-8, Client Satisfaction Questionnaire-8; DBT-WCCL, Dialectical Behavior Therapy Ways of Coping Checklist; EAC, Emotional Approach Coping Scale; FMPS, Frost Multidimensional Perfectionism Scale; ICES, Invalidating Childhood Experiences Scale; IIP-PD-25, Inventory of Interpersonal Problems–Personality Disorders-25 items; LIFE-RIFT, The Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool; MOPS, Measure of Parenting Style; MSSI, Modified Scale for Suicidal Ideation; NEO FFI-c, NEO (neuroticism, extraversion and openness) Five-Factor Inventory, Conscientiousness Subscale; PANAS, Positive and Negative Affect Scale; PHQ-9, Patient Health Questionnaire-9; PNS, Personal Need for Structure; SBQ, Suicidal Behaviors Questionnaire; SSQ, Social Support Questionnaire; SVS, Schwartz Values Scale; UPPS, Urgency Premeditation Perseverance Sensation Seeking Scale; WBSI, White Bear Suppression Inventory.
This table has been amended from Lynch *et al.*[59] Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to http://group.bmj.com/group/rights-licensing/permissions. This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

with the trainer, in order to ensure depth of diagnostic and assessment learning. In-group learning included role-plays of both the SCID and HRSD with assessors and with the trainer, followed by questions and discussion. Assessors subsequently conducted role-plays of the SCID and HRSD with a colleague and compared their ratings with the expert ratings from the training DVDs. Assessors then either audiotaped an assessment and had this assessment reviewed by the trainer or shadowed an experienced assessor and completed an initial joint assessment.

We included an embedded 'reliability study' in which all assessors coded audiotaped interviews that had been taped and coded by another member of the team. Krippendorff's alpha[65] was used, as it can assess reliability between multiple raters and it handles nominal data and missing data.

Blinding

The assessors were blind to participant allocation. All participants were informed of allocation by the trial manager or administrator. Follow-up assessments were carried out at different locations from treatment locations to ensure that the research assistants carrying out the follow-ups remained blind to treatment allocations. Precautionary strategies included consideration of locations for follow-up assessments, participants being reminded by the assessors not to talk about allocation before the assessment and, after the initial assessment, assessors not looking at clinical notes. If an allocation was revealed, then reblinding occurred by using another rater for the subsequent follow-up. If the blinding was broken during an assessment session, then these ratings were used. Unblinded ratings were used for 17 7-month follow-ups, 12 12-month follow-ups and zero 18-month follow-ups.

Reimbursement

Participants in both groups were reimbursed by cheque for their time spent completing research questionnaires and attending the follow-up interviews. The total amount of reimbursement ranged from £0 to £150, depending on whether or not the participants were still in the study and the number of follow-up assessments they attended. Participants received £30 after they had been in the study for 3 months, £40 for attending the first follow-up interview after 7 months, another £50 for attending the follow-up interview after 12 months and a final £30 for attending the follow-up interview after 18 months.

Settings and locations of data collection

Participants were recruited, assessed and treated at three different NHS trusts across the UK: Dorset HealthCare University NHS Foundation Trust, with locations in Bournemouth and Poole (Dorset); Southern Health NHS Foundation Trust, with locations mainly in Winchester, Eastleigh and Southampton (Hampshire); and Betsi Cadwaladr University Health Board, with locations in Bangor and Rhyl (North Wales). All three sites routinely offer outpatient psychological treatment.

To ensure accuracy, completeness and reliability, trained assessors collected outcome data during interviews in person or by telephone following standard operating procedures for data collection and transferred the data to the trial office in Southampton. In addition, monthly paper questionnaires were posted directly to participants with a postage paid envelope, and text or e-mail reminders were sent after 10 days if the questionnaires had not been received by the trial office.

Electronic data capture system

All paper-based data were entered twice onto Signalbox (version 2.3.7; B Whalley, Plymouth, UK),[66] a validated electronic data capture system. All duplicate replies were then checked by an experienced data manager and, in the case of discrepancies between the two replies, the correct reply was selected for further processing. This system also collected data from automated telephone calls used to collect certain measures (described in *Measures*) and prompted therapists to enter their online treatment notes through weekly electronic reminders.

Changes to methods and procedures after trial commencement

After the trial commenced, we added the following exclusion criteria: patients were excluded if they were on a standard DBT waiting list or were attending standard DBT at the time of assessment.

The original trial protocol states that outcome measures would be assessed at baseline and at 6, 12 and 18 months after randomisation. However, the 6-month assessment was pushed back to 7 months once it became apparent that those participants in the RO DBT group would still be receiving active treatment

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

6 months after randomisation, whereas we aimed to assess participants immediately after treatment, which was, on average, 7 months after randomisation. This change was made before any patient was due their 6-month follow-up and did not differentially affect follow-up times in the two groups.

In addition, as a result of decreased therapist availability and limited research assessor time, recruitment did not progress at the expected rate. Three changes were made to ensure that the trial finished within the allocated time and budget:

1. we extended the recruitment period by 9 months
2. the target sample size was reduced from 276 to 240 participants
3. participants who were recruited after 1 September 2014 were followed up for 12 months instead of 18.

These changes affected a total of 27 participants who were still in the study after 12 months.

Ethics considerations

We asked trial participants to give consent on three occasions:

1. oral consent before answering questions via the telephone screening interview
2. signed consent before the baseline assessment, in which they confirmed that they had read and understood the information sheet and gave permission for the interview to be audio-recorded
3. signed consent to participate in the trial for eligible participants, stating that they had understood all information provided to them and were willing to participate in the trial.

Personal information was kept in a password-protected database, which was accessible only to members of staff of each site who needed these details for making assessment appointments and sending out letters. These were the only files in which both personal information and the participant identification numbers were stored.

We conducted the RefraMED trial in accordance with the Declaration of Helsinki.[67] Prior to the start of the study, we received approval from the Hampshire Research Ethics Committee (National Research Ethics Service reference number 11/SC/0146) and the ethics and research governance department of the sponsor of this study, the University of Southampton.

Study monitoring

Two committees, which were independent from the funder and the sponsor, monitored the trial to ensure that it complied with the rigorous standards defined in the National Institute for Health Research Clinical Research Network's guidelines for good clinical practice:[68] the Trial Steering Committee (TSC) and DMEC. The TSC met, via a conference call, twice a year during the study and the DMEC met twice in the first year and annually after that.

Reporting of adverse events and study termination

Site principal investigators were responsible for monitoring and reporting serious adverse events (SAEs). These were immediately reported to the chief investigator. The trial office reported these events annually to the DMEC or immediately in the case of a suspected unexpected serious adverse reaction (SUSAR). The trial procedure for evaluating and reporting AEs and SAEs can be found in *Appendix 2*.

Patient and public involvement

Service users were actively involved with the development of patient information leaflets, the management of the study and the dissemination of the study information and results. The Mental Health Research Network (London, UK) and INVOLVE (Southampton, UK) helped with recruitment of service users. Four mental health service users were recruited: two for the TSC and two for the Trial Management Group (TMG).

Interventions

Treatment as usual

All participants received TAU, consisting (primarily) of prescribed ADM or psychotherapy (with the exception of standard DBT). At each follow-up assessment, participants were asked to report the type of ADM, their adherence to it and the type and amount of psychotherapy that they had accessed in the months since their previous assessment (or in the 6 months preceding their baseline assessment). In addition, we did not restrict access to appropriate mental health care during follow-up. 'Service contact' may have served different functions, namely for the purpose of (1) assessment, (2) psychological treatment or (3) routine clinical monitoring of needs and risk. TAU differed across sites.

In Dorset, the practice of many services was to assess patients (with services typically offering three assessments) to formulate their difficulties and determine their needs, before putting their name on a waiting list for appropriate treatment. During the waiting period, the patient may have received input from a Community Mental Health Team (CMHT) to review their needs, risk and medication. Patients who had been referred to CMHTs received variable treatment, which may have included CBT, schema-focused cognitive therapy, standard DBT-based skills groups or cognitive–analytic therapy (CAT), all of which delivered was by psychologists, occupational therapists or nurses. Alternatively, the patients were case managed by CMHT staff, which did not include a therapeutic model as such. In many services the wait to receive treatment was long, typically up to 1 year. Some patients were seen only by their GP and TAU consisted of ADM and whatever support their GP offered in the administration and monitoring of ADM.

In North Wales, patients referred to mental health services were assessed via a single-point-of-access system that occurred within hours for urgent cases and within 28 days for routine cases. Typically, all of these patients were already in receipt of ADM. Patients who were assessed as having mild depression were seen in primary care mental health services, in which the interventions offered included basic behavioural activation, support and monitoring of ADM and occasionally CBT. Others may also have been referred to third-sector counselling services and, in some areas, access to mindfulness-based cognitive therapy groups was available. Patients with moderate to severe problems were seen in secondary care by the CMHTs and would typically have a generic case manager and may also have been seen by the psychiatrist or a clinical psychologist. In most areas, the clinical psychologist would offer a formulation-based psychological intervention; the models offered would include CBT, dialectical behavior therapy (DBT)-informed, schema-focused cognitive therapy, acceptance and commitment therapy and CAT. Depending on their geographical area, patients may also have received CBT from a CBT therapist. Waiting lists for therapy were long, ranging from 3 months to 2 years, depending on the area, and the approximate average wait was 12 months.

In Hampshire, some GP practices and all universities provide a counselling service. There is also an active voluntary sector, which provides support, information, counselling and group work. Patients could self-refer or be referred to Increasing Access to Psychological Therapies (IAPT) services. IAPT offers a stepped approach to treatment for anxiety and depression with telephone sessions, group sessions, couple work or up to 20 individual sessions of CBT or IPT. Patients with more severe and enduring difficulties were referred to their CMHT by their GP. The patients were typically seen within a few weeks for an assessment of their

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

needs. They may also have received regular monitoring and case-management by CMHT staff, including problem-solving, reviewing of risk and medication advice or being offered a range of treatment options, for example, group work, such as mindfulness or standard DBT-based skills groups. If the patient required more intensive input, then a home treatment team or admission to an inpatient acute unit was available. In addition to these options, patients could have been referred for specific psychological therapy. The waiting time would typically be up to 4 months from referral, and therapy options included group and individual work using CBT, DBT-informed, schema-focused cognitive therapy, CAT or acceptance and commitment therapy.

Radically open dialectical behaviour therapy

Radically open dialectical behaviour therapy is a transdiagnostic treatment designed to address a spectrum of difficult-to-treat disorders, including chronic depression. RO DBT significantly differs from other treatment approaches, most notably by linking the communicative functions of emotional expression to the formation of close social bonds and by skills targeting social signalling and changing neurophysiological arousal.

The experimental intervention was fully manualised and comprised 29 weekly individual therapy sessions lasting 50–60 minutes and 27 weekly skills training classes lasting 2.5 hours, including 15-minute breaks.[41,42,69] In addition, RO DBT therapists met weekly in person for a consultation team meeting that lasted 1.5–2.5 hours and by telephone when needed. The RO DBT treatment commenced as soon as possible after the participant had been notified of their treatment allocation. Although RO DBT participants received ADM as prescribed, we strongly discouraged them from seeking additional psychotherapy during RO DBT. The RO DBT treatment developer and study chief investigator (Thomas R Lynch) did not contribute to treatment delivery.

The RO DBT lesson plan used in this trial is presented in *Appendix 3*. The research manual comprised a combination of newly developed RO DBT lessons, which are now part of a fully comprehensive RO DBT skills manual,[69] and standard DBT.[70]

Radically open dialectical behaviour therapy therapist training, supervision and adherence

Twenty-three therapists (male, $n = 2$, female, $n = 21$; age range from 32 to 61 years), who were trained in RO DBT, delivered the treatment across the three sites (Dorset, $n = 8$ therapists; Hampshire, $n = 10$ therapists; and North Wales, $n = 5$ therapists). All therapists received a minimum of 10 days of specific training in RO DBT from the treatment developer and chief investigator of the study (Thomas R Lynch).

Therapist selection

In order to be recruited onto the trial, therapists were required to have three treatment tapes rated as adherent on the standard DBT Adherence Rating Scale (Marsha M Linehan and Kathryn Korslund, University of Washington, Seattle, WA, 2003, personal communication): a score of ≥ 4.0 points on the 5-point scale denoted adherence. This scale is the gold standard measure of adherence in standard DBT and is commonly used in research trials.[71,72]

Training, supervision and adherence

The RO DBT training and supervision were conducted by the treatment developer (TRL) and covered the biosocial theory of the development of overcontrolled behavioural patterns; the novel RO DBT mechanism of change, linking open expression of emotion to increased trust and social connectedness, which are new overcontrolled treatment targets; modifications to the relationship strategies; the importance of social signalling and mindfulness practices, including loving kindness, self-enquiry and the new radical openness (RO) skills. Some therapists elected to attend the training offered for the treatment on more than one occasion.

On commencing the trial, all therapists attended weekly 2.5-hour RO DBT consultation team meetings. The RO DBT consultation team meeting served several important functions; the meetings not only provided a platform for therapist practice of RO, they also helped to reduce therapist burnout and enhance empathy and adherence to the treatment. During consultation team meetings, therapists presented problems that they were experiencing within the treatment and sought consultation from each other. This involved showing video clips of therapy sessions and feedback about how best to apply the treatment to resolve clients' difficulties or address challenges.

Each team was led by a practitioner who had many years' experience of delivering standard DBT to individuals with borderline personality disorder (BPD) in NHS settings. Three of the four clinical team leads (the clinical lead for the Dorset trust changed part way through the trial) were experienced trainers in standard DBT and were well versed in developing and shaping a team towards adherent treatment delivery. To support the team leads in this vital role, the chief investigator supervised each team lead on a selection of their cases. Supervision involved a detailed microanalysis of therapy tapes in order to increase skill and expertise in the delivery of RO DBT and also to shape skills in how to supervise others in learning RO DBT. On average, team leaders were supervised once every 8 weeks in the early part of the study. The frequency of supervision reduced as competency developed. Therapists also received supervision from the chief investigator via site visits and individual supervision as required.

Finally, a random selection of RO DBT session tapes were coded for adherence using the standard DBT Adherence Rating Scale (Linehan and Korslund, personal communication). Feedback was provided to the site team lead and the individual therapist on the score obtained, with brief guidance on how to improve adherence if necessary. The four coders were trained to reliability with the scale developer and rerated for reliability at the mid-point of the trial.

Measures

Table 1 presents an overview of all included measures.

Primary outcomes: depression symptoms

The primary outcome was the difference in depressive symptoms at 12 months between patients receiving RO DBT and TAU, as measured using the 17-item HRSD.[63] The HRSD was assessed at four time points (baseline and at 7, 12 and 18 months after randomisation) by trained assessors.

Secondary outcomes

Secondary outcomes at the same four time points included remission status and suicidal ideation and behaviour. We also examined clinically significant change post hoc using reliable change indices.[73]

Remission status

In line with recommendations from Greer *et al.*,[1] remission status was based on the level of depression symptoms (as measured using the HRSD) and psychosocial functioning level [as measured with The Longitudinal Interval Follow-up Evaluation – Range of Impaired Functioning Tool (LIFE-RIFT)].[74]

For the primary definition of remission the HRSD and LIFE-RIFT were recoded in the following way: Patients with a HRSD score of < 15 points and a LIFE-RIFT score of < 13 points were defined as experiencing partial remission. Patients with a HRSD score of < 8 points and a LIFE-RIFT score of < 13 points were identified as in full remission. Other patients were defined as experiencing 'no remission' or, when data were missing, as having 'unknown' remission status. Because patient numbers differed in the two arms, we present the percentage of patients in each category, using the number (n) allocated to treatment as the denominator.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

In addition, because remission rates are defined in a number of ways in the literature, and because the selection of any single criterion is somewhat arbitrary,[75] we also computed a number of other thresholded scores based on the HRSD. These other thresholded scores included:

- Patients whose HRSD score dropped to < 8 points, and thus meet NICE's 2009 definition of 'not depressed'.
- Patients experiencing > 17.5% or > 50% change in symptoms from baseline. Externally anchored criteria, or those defined by distributions of 'sick' and 'well' patients,[73] may fail to capture patients' own views of their improvement. Recent work has mapped patients' 'global rating of change' to psychometric scale scores and found that a 17.5% change in symptoms corresponded to the minimal clinically important difference from the patient perspective.[76] This previous work has used the BDI, but because correlations between BDI and HRSD are high (in the region of $r = 0.7$[77]) we adopted the same approach using the HRSD.

Suicidal ideation and behaviour

In addition, we evaluated changes in suicidal thoughts and behaviours. Suicidal ideation was measured using the assessor-rated Modified Scale for Suicidal Ideation (MSSI).[78] All participants were asked to answer the first four questions of this scale. As per the instructions of the scale, only those who scored ≥ 2 points on questions 1 and 2 and ≥ 1 points on questions 3 and 4 were asked to answer questions 5–18. The questions are scored from 0 to 3 points, with higher scores indicating increased suicidal ideation. Scores of ≤ 8 points indicate low suicidal ideation, 9–20 points indicate mild to moderate suicidal ideation and 21 points or higher indicate severe suicidal ideation.

To measure suicidal and self-harm behaviour, we used an adapted version of the self-reported Suicidal Behaviors Questionnaire.[79] Suicide behaviour and intent was calculated using the sum score of the first 10 questions (we omitted questions 4 and 9 of the original scale to avoid repetitiveness). Self-harm behaviour was calculated summing the scores of all behaviours listed in questions 12a–k of the original scale.

Economic measures

We collected resource-use data using a version of the Adult Service Use Schedule (AD-SUS), which is designed for depression populations. The AD-SUS collects service use and associated data and has been successful in a range of adult mental health population trials.[80] Completed in interview with participants, the questionnaire records use of hospital- and community-based health and personal social care services for all causes. For medications, we asked participants to report prescribed antidepressants, antipsychotics, sleeping tablets and painkillers. To avoid unblinding researchers, participants reported their use of all talking therapies on a separate self-report questionnaire. We collected information about time off work as a result of illness alongside the AD-SUS, using the productivity questions from the World Health Organization (WHO) Health and Work Performance Questionnaire.[81] We asked participants to complete the AD-SUS and WHO's Health and Work Performance Questionnaire at baseline to cover the previous 6 months and at the 7-, 12- and 18-month interviews to cover the time since the previous interview, thus covering the full period from baseline to final follow-up. We extracted information on RO DBT-related service use, including the number of individual sessions and group sessions attended by each participant, from therapy records.

We estimated QALYs from EQ-5D-3L scores at baseline and at 7, 12 and 18 months. The EQ-5D-3L is a non-disease-specific measure for describing and valuing health-related quality of life.[50] The measure includes a rating of health in five domains (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a rating of own health using a visual analogue scale (a thermometer scaled from 0 to 100). The health states described in the EQ-5D-3L were assigned a utility weight or score using responses from a representative sample of adults in the UK.[82] These weights were applied to the time between interviews and QALYs calculated using the area under the curve approach.[83] The EQ-5D-3L has been validated in economic evaluations for common mental health disorders.[84]

Moderating variables

At baseline we collected the following data on potential moderating variables (see *Table 1*):

- Basic demographics, including age, sex and marital status.
- Balanced Inventory of Desirable Responding-Short Form[85] – this is a 16-item questionnaire that is answered using a seven-point Likert scale ranging from 1 = totally disagree to 7 = totally agree; the scale assesses two dimensions: (1) self-deceptive enhancement and (2) impression management.
- Ego-Undercontrol and Ego-Resiliency Scales[86] – the Ego-Undercontrol Scale has 37 questions and the Ego-Resiliency Scale has 14 questions; both questionnaires are answered using a four-point Likert scale ranging from 1 = disagree very strongly to 4 = agree very strongly.
- Invalidating Childhood Experiences Scale[87] – this scale asks 14 questions about negative childhood experiences (≤ 18 years) in relation to each parent. Each item is rated on a Likert-type scale (1 = never; 5 = all the time). The levels of perceived invalidation by mothers and fathers are given by the mean score for the 14 items for each parent. Higher scores reflect a greater perception of invalidation by that parent. Four additional questions include descriptions of three types of invalidating environment (perfect, chaotic or typical) and one description of a validating environment. The items were rated on a five-point scale ('not like my family' to 'like my family all of the time'). Higher mean scores indicate greater levels of either the validating environment or the three types of invalidating environment.
- Personal Need for Structure Scale[88] – this 11-item scale comprises two subscales: desire for structure and response to lack of structure. Questions are answered using a six-point Likert scale ranging from 1 = strongly disagree to 6 = strongly agree.
- Urgency Premeditation Perseverance Sensation Seeking Scale[89] – this is a 45-item questionnaire consisting of four subscales: (negative) urgency, (lack of) premeditation, (lack of) perseverance and sensation-seeking. Questions are answered using a four-point Likert scale ranging from 1 = agree strongly to 4 = disagree strongly.
- SCID-I and SCID-II[64,90] – these are semistructured diagnostic interviews for identifying DSM-IV Axes I and II mental health and PDs, respectively.

Both participants and therapists completed the following:

- Cambridge Prospective Memory Test[91]– this is a laboratory measure of prospective memory that comprises a total of six prospective memory tasks, three cued by time and three cued by events. Participants were asked to work on some distractor tasks, such as word-finder puzzles or a general knowledge quiz, for a 20-minute period while they also had to remember to perform the prospective memory tasks. For each task, between 0 and 6 points are awarded depending on whether or not the tasks were completed and to what degree they required prompting from the assessor, with higher scores reflecting better prospective memory performance.
- Conscientiousness Subscale [of the NEO (neuroticism, extraversion and openness) Five-Factor Inventory][92] – this subscale comprises 12 items scored using a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree.
- The Frost Multidimensional Perfectionism Scale[93] – 35 questions are answered using a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. We computed separate scores for adaptive and maladaptive perfectionism.[94–96]
- Schwartz Values Scale[97] – 57 statements are rated on a nine-point Likert scale indicating the degree to which each statement is 'a guiding principle' in the participant's life. The items reflect 10 universal value types: security, power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, conformity and tradition.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

In addition, therapists completed:

- Measure of Parenting Style Scale[98] – this scale measures perceived parenting style across three domains: indifference, abuse and overcontrol. Fifteen items ask about the behaviour of the participant's parents towards them during the first 16 years of their life for each parent separately. The items are rated from 0 = not true at all to 3 = extremely true.

Mediating variables

Data on potential mediators were collected at 0, 3, 7, 12 and 18 months using the following questionnaires (see *Table 1*):

- Acceptance and Action Questionnaire-II (AAQ-II)[99] – this questionnaire is a 7-item scale measuring psychological inflexibility. Questions are answered using a seven-point Likert scale ranging from 1 = never true to 7 = always true.
- Ambivalence over Emotional expression Questionnaire[100] – this is a 28-item questionnaire measuring conflict over emotional expression. It is answered using a five-point Likert scale ranging from 1 = never to 5 = very often.
- The Dialectical Behaviour Therapy Ways of Coping Checklist (DBT-WCCL)[101] – this checklist is based on the earlier Revised Ways of Coping Checklist and is an inventory of emotional coping skills taught in standard DBT. It comprises 59 items rated on a three-point Likert scale ranging from 0 = never used to 3 = regularly used. The scale has three subscales: DBT skills use, general dysfunctional coping and blaming others.
- Inventory of Interpersonal Problems–Personality Disorders[102] – this 25-item dimensional PD measure is answered using a four-point Likert scale ranging from 0 = not at all to 4 = extremely. Five subscales can be calculated: interpersonal sensitivity, interpersonal ambivalence, aggression, need for social approval and lack of sociability.
- The 3-item Social Support Questionnaire (SSQ)[103] – this questionnaire asks participants to list the initials of people who provide support to them in three different manners and participants are asked to rate their degree of satisfaction for each manner of support received using a Likert scale ranging from 1 = very dissatisfied to 6 = very satisfied.
- White Bear Suppression Inventory[104] – this 15-item questionnaire measures suppression and avoidance of unwanted thoughts using a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree.

Temporal sequencing measures

Participants were asked to complete the following measures at baseline, then once a month for the first 12 months and again at 18 months:

- The Patient Health Questionnaire-9[105] – this is a 9-item psychometrically valid assessment of depressive symptoms using a three-point Likert scale ranging from 0 = not at all to 3 = nearly every day.
- The Positive and Negative Affect Scale (PANAS)[106] – the 20-item PANAS provides independent estimates of positive and negative affect and is rated from 1 = very slightly/not at all to 5 = extremely.
- The Emotional Approach Coping scale (EAC)[107] – this scale comprises eight items and has two subscales: emotional processing and emotional expression. Items are rated from 1 = I usually don't do this at all to 4 = I usually do this a lot.
- The Credibility/Expectancy Questionnaire[108] – this questionnaire indexes positive expectancies for treatment and comprises six items with various scales. The first four items aim to measure credibility and the last two items aim to measure expectancy.

In addition to the listed questionnaires, we also used an automated telephone system to measure mood using the 10-item PANAS Short-Form[109] and coping skills using the six items from the DBT-WCCL[101] once a week over the first 6 months that participants were in the trial (ecological momentary assessments).

Therapeutic alliance

To measure therapeutic alliance, patients allocated to RO DBT were asked to complete the patient version of the California Psychotherapy Alliance Scales[110] during the first 4 weeks of treatment and then once a month until the end of treatment. Therapists were asked to complete the therapist version at the same time points.

Therapist characteristics

To investigate whether or not therapist characteristics, and their relationship with patient characteristics, influence outcome, we asked them to complete measures of conscientiousness (the conscientiousness scale from the NEO Five-Factor Inventory),[92] perfectionism (Frost Multidimensional Perfectionism Scale),[93] personal values (Schwartz Values Scale)[97] and attachment experiences (Measure of Parenting Style Scale).[98] More detailed descriptions of the scales are available in *Moderating variables*.

We also measured therapists' capacity to generate strong alliances before the trial by asking several patients from the caseload of each therapist to complete the California Therapeutic Alliance Scale.[110] This measure was used as an IV for strength of therapeutic alliance.

Client Satisfaction Questionnaire

To assess the acceptability of the treatment and the study as a whole, we asked the first 50 participants to complete the Client Satisfaction Questionnaire-8.[111]

Statistical methods

Data processing and reduction

We stored study data in several databases, which were subsequently linked for analysis using the five-digit participant code (in which the first digit indicates study site) provided with each randomisation request. In parallel to the randomisation database, the main study database stored questionnaire responses and notes, and further databases stored additional information on therapists and therapist sessions.

Data validation included confirmation of the consistency of information in various databases, and thorough checks on the fidelity of double-entry data, with inconsistencies checked against original paper forms. Following such checks and the resolution of all identified inconsistencies, documented audits of questionnaire responses across a range of questionnaires and time points were undertaken, in each case using a random selection of participants representing all study sites and both study arms.

Study data were exported from the main study database and subsequently processed by a series of scripts (primarily written in Stata®, version 14; StataCorp LP, College Station, TX, USA) that renamed variables, created labels for variables and variable values, created codes for reasons for missing data, identified canonical responses for double-entry data, resolved time points to specific months, defined pairings of participant and therapist and scored questionnaires, using agreed rules on the treatment of missing data. Each script generated an interim file for any further checking required. The final stage dropped all intermediate variables and individual responses before merging the reduced data set with randomisation and therapist information to create a final data set for use in the analyses.

Statistical analysis

Primary and secondary outcome analyses: continuous variables

The primary outcome, HRSD score, and all continuous secondary outcomes were analysed by treatment allocated, modelling differences between groups at baseline and then at 7, 12 and 18 months after randomisation. Covariates in the primary analyses included treatment site (Dorset, Hampshire, North Wales), baseline HRSD score, baseline PD status (presence or absence) and early-onset depression (yes or no).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

A three-level mixed-effects model was used to account for clustering of data by patient and therapist. Notably, this model does not assume that all therapists are equally effective.[112] These mixed models are efficient and unbiased in the case that data are missing at random, that is, missing at random conditional on the covariates included in the model. As suitable auxiliary data were not available, missing scale responses were not imputed (e.g. via multiple imputation). However, missing questionnaire items were imputed for rows when fewer than 10% of items were missing in a given scale. This item imputation was performed by linear regression using the other scale items as covariates.

For each outcome, the main effects of time and treatment allocation and the interaction of time and treatment were tested for. Planned contrasts were then made for each outcome at 12 months. Secondary analyses computed differences between treatment groups at months 7 and 18. Details of the analyses are available in *Box 1*.

Outcome analyses: binary or categorical outcomes

Remission status and thresholds for clinically reliable/significant change

As remission rates require the application of binary thresholds to an underlying continuous score, the primary and secondary outcome models (as described in *Primary and secondary outcome analyses: continuous variables*) were used to compute predictions for each patient and calculate remission or change rates based on these predictions, which are, thus, adjusted for baseline covariates. As remission rates are based on an underlying continuous score, more powerful models of these continuous variables were used for statistical inference and significance tests for the difference in rates between groups are not presented.

Clinical predictions for new patients

Understandably, most reports of clinical trials focus on mean differences between treatment groups and the precision of this difference. However, clinicians and patients may be less concerned with the average population effect of a treatment (and the associated *p*-value for this contrast) than with the likely range of clinical outcomes given different treatment choices. For example, patients might wish to know what benefit they may obtain if the treatment is successful, or is less successful than hoped, compared with continuing in TAU. To provide this information the primary outcome model was reran for HRSD using the popular Bayesian modelling package, *Stan*® (see http://mc-stan.org/). This model was identical to the

BOX 1 Details of analyses

Both primary and secondary outcomes were modelled using mixed-effects models using the *Lmer* package in R (version 3.3.3; The R Foundation for Statistical Computing, Vienna, Austria).[113] Models for primary and secondary outcomes were specified in the form:

```
[dv] ~ (grp * month * site) + (grp * month * pd.baseline) + (grp * month *
[dv].baseline) + (grp * month * age.depression.onset) + (1 | patient) +
(1 | therapist)
```

In these models *[dv].baseline* indicates baseline score for the dependent variable, and *pd.baseline* denotes a binary predictor for the presence of any PD at baseline. The *month*, *patient* and *therapist* variables are factors. Type III tests of fixed effects were computed using Satterthwaite correction[114] for degrees of freedom, and between-group contrasts at each time point were computed using the contrast function within the *emmeans* package. As the sample of therapists was small, null-hypothesis tests on the random effects in the models were not performed, but the proportion of variance attributable to therapists (see *Chapter 3*, *Variations in therapist performance*) was reported.

primary model in all respects, except that (1) the model was fit using Stan's Hamiltonian Monte Carlo sampling routines, (2) uninformative priors were placed on all model parameters and (3) for simplicity and to aid convergence, the non-significant interaction between PD status at baseline and treatment group was removed. Prior distributions were normal (0,10) for regression parameters and Cauchy (0,2) for variance parameters.

Once the model converged satisfactorily (judged by visual inspection of Markov Chain Monte Carlo traces), observations were simulated, representing new patients who did or did not accept RO DBT, by drawing from the posterior distribution.[115] Descriptive statistics based on these simulations are presented in *Chapter 3*, *Clinical predictions for new patients*.

Economic evaluation

Perspective

The economic evaluation took the perspective of NHS and personal social services preferred by NICE[51] and explored the impact that productivity losses attributable to time off work as a result of illness had in the sensitivity analysis.

Unit costs

With the exception of RO DBT, mean costs of health and social services were estimated for each group by multiplying patients' reported use by unit costs from national sources.[116–118] All unit costs, summarised in *Table 2*, were for the financial year 2014–15 and uprated, when necessary, using the Hospital and Community Health Services Index.[116] Medication costs were based on the median dose of the most common category of drug (e.g. ADM or antipsychotic medication) reported by participants. Assuming full compliance, participant-reported start and finish dates were used to estimate the duration of their time using that drug. Costs associated with depression-related absenteeism and presenteeism for patients in paid employment were estimated using the human capital approach based on the national gross average wage.[122,123]

The RO DBT costs were calculated using the microcosting approach developed by the Personal Social Services Research Unit (PSSRU) at the University of Kent.[124] Individual sessions were costed on the basis of the average therapist Agenda for Change salary bands, including employer's costs (national insurance and pension contributions) and overhead costs (buildings, utilities, management and administration), taken from national sources.[116,119] Salary costs were weighted to include non-face-to-face time using information collected from 16 RO DBT therapists on the time that they spent running RO DBT therapy sessions and the time that they spent on other activities, to calculate a direct-to-non-direct ratio of 1 : 0.91. Individual sessions lasted, on average, 60 minutes and the method of valuation of the therapists' time is summarised in *Table 3*.

The RO DBT group sessions were costed on the basis that they were closed to other participants and went ahead irrespective of how many participants attended.[125] This involved allocating the total cost of a group session across all of those participants invited to attend, irrespective of whether or not they did attend. Group sessions lasted, on average, 2.5 hours. The number of therapists running the groups varied by group size. Groups that were larger than three clients were typically run by two therapists. The valuation of the cost of these group sessions is summarised in *Table 4*.

The RO DBT-specific training costs were not included because equivalent costs for the TAU group could not be as easily identified and costed, making comparison difficult. In clinical practice, therapists undertake a wide range of training as part of usual professional development and it is reasonable to assume that RO DBT training, if rolled out, would form part of this professional development in the same way as therapists receive training in other therapies, such as CBT. Similarly, the cost of participants who failed to attend RO DBT individual sessions (who did not attend) was excluded from the analysis, given the absence of equivalent data for the TAU group. It is only when two specific therapies are compared that it is possible to accurately record and cost participants who did not attend from both groups; in this study, TAU varied and participants reported only attendances with various therapists and other health professionals and not did not attends (DNAs).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 2 Unit costs of health and social services

Item	Source	Unit cost
Talking therapy		
Individual		
Face to face	PSSRU's *Unit Costs of Health and Social Care 2014*;[119] 9.5 clinical psychologist	£139 per contact hour
Telephone	PSSRU's *Unit Costs of Health and Social Care 2014*;[119] 9.5 clinical psychologist	£139 per contact hour
Group		
Face to face	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 2.9 MBCT therapy – group-based interventions	£14 per person
Hospital services		
Mental health admission	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 2.1 mental health-care clusters (bed-day)	£223 per bed-day
Drug services admission	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 2.1 drug services – admitted (bed-day)	£433 per bed-day
Non-elective long stay (≥ 5 days)	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 7.1 NHS reference costs for hospital services	£2863 per episode
Non-elective inpatient short stay (< 5 days)	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 7.1 NHS reference costs for hospital services	£608 per episode
Outpatient appointments	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 7.1 NHS reference costs for hospital services	£112 per attendance
A&E department attendance	*NHS Reference Costs 2014 to 2015*[118]	£132 per attendance
Ambulance	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 7.1 NHS reference costs for hospital services	£231 per attendance
Community services		
GP		
Surgery	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 10.8b GP surgery visit of 11.7 minutes	£44 per visit
Home	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 10.8a GP home visit of 23.4 minutes	£89 per visit
Telephone	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 10.8b GP telephone consultation of 7.2 minutes	£27 per call
Practice nurse	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 10.6 nurse – GP practice	£56 per contact hour
Other community nurse	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 10.3 health visitor	£76 per contact hour
Mental health-care support workers	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 11.6 home care worker	£24 per contact hour
Community psychiatric nurse	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 10.2 nurse (mental health)	£75 per contact hour
Community psychiatrist	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 15.7 hospital-based consultant – psychiatric	£139 per contact hour
Occupational therapist	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 11.5 community occupational therapist	£44 per contact hour
Art therapy	Assumed equivalent to 11.5 community occupational therapist	£44 per contact hour
Social worker	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 11.2 social worker (adult services)	£57 per contact hour

TABLE 2 Unit costs of health and social services (*continued*)

Item	Source	Unit cost
Marriage counselling	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 11.8 family support worker	£51per contact hour
Advice service	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 11.4 social work assistant	£30 per contact hour
Helpline	Third sector[120]	£3.91 per call
Day care/drop-in centre	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 2.4 local authority day care for people with mental health problems	£32 per client attendance
Physiotherapist	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 13.1 hospital-based physiotherapist	£38 per contact hour
Audiologist	Assumed equivalent to 11.5 community occupational therapy	£44 per contact hour
CMHT	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 12.2 community mental health team for adults with mental health problems	£37 per contact hour
Dentist	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 10.11 dentist – providing-performer	£207 per contact hour
Hospital at home	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 12.11 re-enablement service	£43 per contact hour
Community pharmacist	PSSRU's *Unit Costs of Health and Social Care 2015*;[116] 9.6 community pharmacist	£144 per contact hour
Medications		
ADM	NHS Business Services Authority's *Drug Tariff Part VIII* Drug tariff (part VIIIA, category M);[121] median dose of 150 mg of venlafaxine taken in two 75-mg tablets once a day	£2.35 for a packet of 56 75-mg tablets
Antipsychotic medications	NHS Business Services Authority's *Drug Tariff Part VIII* Drug tariff (part VIIIA, category M);[121] median dose of 100 mg of quetiapine taken in one 100-mg tablet	£1.79 for a packet of 60 100-mg tablets

A&E, accident and emergency; PSSRU, Personal Social Services Research Unit.

TABLE 3 Valuation of the RO DBT therapists' time

Item	Cost	Source
A: clinical psychologist's salary	£55,243	Survey of RO DBT therapists
B: employers' NI and superannuation	£13,773	NI plus 14% pension (PSSRU 2015[116])
C: overheads	£40,078	PSSRU's *Unit Costs of Health and Social Care 2014*;[119] 9.5 clinical psychologist, uprated to 2015[116]
D: wages plus overheads (item A + item B + item C)	£109,093	Item A + item B + item C
E: working time	1538 hours	Hours per year, based on 37.5 hours/week for 41 weeks/year
F: cost per hour	£70.93	Item D/item E
G: cost per hour in direct client contact	£135.78	Item F × 1.91 ratio of face-to-face to indirect time estimated in survey of RO DBT therapists
H: cost per minute in direct client contact	£2.26	Item G/60 minutes

NI, national insurance.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 4 Valuation of RO DBT group attendance

Item	Cost	Source
A: clinical psychologist's cost per hour in direct client contact	£136	See *Table 3*
B: average number of therapists per group	1.71	Group therapy database
C: duration of group	2.5 hours	Study therapists
D: mean number of clients allocated to each group	5.9	Group therapy database
E: cost per client per group	£99	(Item A × item B × item C) ÷ item D

Furthermore, it could be reasonably assumed that therapists would undertake other work activities during the time made available by DNAs.

Analysis of costs and quality-adjusted life-years

Initially descriptive data on costs and outcomes will be presented, adjusted for baseline differences in costs and relevant outcomes and prespecified clinical predictors (HRSD score, PD, site, age, age at depression onset and sex). The cost-effectiveness analyses was adjusted for the same prespecified clinical predictors and baseline values of the variables of interest (costs or EQ-5D-3L). Cost differences were analysed by *t*-tests, with confidence intervals (CIs) around adjusted mean differences estimated using non-parametric bootstrapping to reflect non-normality of cost data. Missing cost and QALY data were imputed using multiple imputation using chained equations under the assumption that these data were missing at random.[126] Multiple imputation indicates the sensitivity of results to missing data and the assumption that the data are missing at random. It is generally accepted that multiple imputation provides less biased estimates of costs and effects than complete-case analysis, unless data are missing not at random.

Cost-effectiveness

The prespecified primary cost-effectiveness analysis was undertaken at the 12-month follow-up point, in line with the clinical analyses. Cost-effectiveness was assessed by estimating incremental cost-effectiveness ratios (ICERs).[127] The joint distribution of incremental mean costs and effects was generated for RO DBT relative to TAU using non-parametric bootstrapping to explore the probability that one of the groups is the better choice, given the 'willingness-to-pay' threshold of NICE of £20,000–30,000 per QALY. Uncertainty was characterised around the cost-effectiveness and clinical effectiveness estimates by cost-effectiveness acceptability curves.[128]

Sensitivity analysis

Sensitivity analysis was used to explore the impact that variations in methods and assumptions had on the relative cost-effectiveness of RO DBT and TAU. Five sensitivity analyses were conducted:

1. a complete-case analysis for comparison with the results using multiple imputation for missing data
2. a broader economic perspective, which also includes the cost of absenteeism, given that depression is known to have a substantial impact on employment[129]
3. adjustment of the cost of RO DBT group sessions, to take into consideration the fact that group session attendance was particularly low and, thus, costs were particularly high in comparison with the cost of groups reported in other similar studies; on the assumption that group attendance is unlikely to be as low in routine NHS services, sensitivity was tested by replacing the cost of the RO DBT groups with the national cost applied to TAU participants reporting group therapy attendances (£14 compared with £99)
4. analysis of cost-effectiveness using the primary clinical outcome, the HRSD score, as the measure of effect (cost-effectiveness analysis) as compared with QALYs (cost–utility analysis)
5. analysis of cost-effectiveness at the 18-month follow-up point to explore the impact that RO DBT had, compared with TAU, over a longer period of time.

Mechanisms analysis using instrumental variables

In randomised experiments, post-randomisation 'intermediate' outcomes (e.g. mediators such as treatment exposure and alliance scores) are influenced by the participant, the therapist and other factors and so are not under the control of the experimenter. Hence, there is potential for confounding variables, associated with both the intermediate and study outcomes, to bias estimates of the effect of mediators. If these confounding variables are known and measured in the study, then suitable adjustments may be possible. However, it is more likely that confounders are unknown or unmeasured or both and that the intermediate outcome is, therefore, confounded. Thus, conventional analyses often yield biased estimates of the effects of intermediate outcomes.

In econometrics, IV methods have long been used to estimate causal relationships from observational data. An instrument is a variable for which the effect on the study outcome is, by assumption, wholly mediated through the potentially confounded exposure(s) and possibly other measured variable(s). For example, economists have used changes in tobacco taxation as an instrument for health outcomes: if the tax change has an effect on health, it is assumed to be wholly mediated via increases or decreases in smoking behaviour and not by other direct or indirect pathways (often termed the 'exclusion restriction'). If these assumptions hold, the instrument may be used to estimate the causal effect of smoking on health, without fear of confounding. (Note that 'causal effect' was taken to mean what happens, on average, if the experimenter intervenes and changes the value of each patient's intermediate outcome while holding everything else constant; the precise definition of the causal effect is specific to the analytical approach and the strength of the assumptions that the analyst is prepared to make.)

Although conditions within a clinical trial are highly controlled, mediation analysis is still difficult when mediators are not selected by design. For example, unobserved variables (e.g. 'readiness for change') might be the cause of both high therapeutic alliance scores and positive outcomes; if 'readiness' were unmeasured, then researchers might incorrectly conclude that therapeutic alliance is a cause of good outcomes. Unless researchers are confident that they have measured all potential confounders, then standard analyses of mediators and outcomes cannot have a causal interpretation. Although finding IVs that satisfy the requisite conditions can be controversial in observational studies, the issue is less vexed in trials. Any randomly allocated exposure that influences a mediator is a promising candidate for an IV because, by design, it cannot be associated with unobserved confounding variables also affecting the mediator. The central issue that must be justified is that the IV cannot have a direct effect on the study outcome – the exclusion restriction. The effect of the randomised exposure must be wholly mediated by variables that are measured within the study. Estimators for causal effects based on IVs are now widely used;[130–132] for example, treatment assignment and treatment location have been used successfully as instruments in controlled studies,[133] but when multiple mediators are hypothesised, it is crucial to find additional strong instruments to facilitate more complex analyses.

The hypothesised mechanism through which RO DBT affects the main outcome (HRSD) is shown in *Figure 1*. The main outcome is represented by Y and the outcome of randomising patients to RO DBT or TAU by 'allocation'. The mediating variables are 'exposure', 'alliance' and 'skills'. The IV analyses attempt to mitigate the potential for confounding by unobserved variables (U). The measurements of expectancy and credibility investigate direct effects of allocation outcome. The mechanism is assumed to be mediated by exposure, alliance and skills, but the diagram is based on an approximating assumption that only exposure at 7 months and alliance and skills together at 7 months have distinct effects on the outcome. Alliance and skills are treated as contemporaneous intermediate outcomes to simplify the model. The figure also contains a direct effect of 'treatment allocation' on the main outcome and of 'exposure' on HRSD scores (Y). In the first case, this corresponds to allowing RO DBT to affect the outcome through pathways other than those involving exposure, alliance and skills, and, in the second case, exposure can affect the outcome through a pathway that does not work through alliance and skills.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

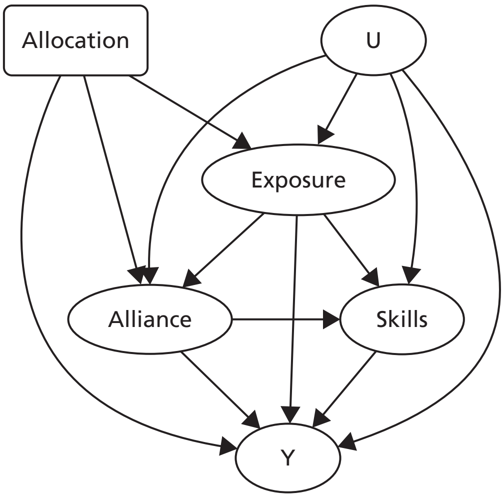

FIGURE 1 The causal model of primary interest showing the hypothesised relationships between treatment allocation (RO DBT vs. TAU), exposure, alliance, skills and treatment outcome (Y), allowing for unknown variables (U).

Process variables

The exposure, alliance and skills mediators will all be treated as 'process variables' in the model.[130] A process variable takes a fixed value (usually zero) for everyone receiving TAU. As such, in the model, there is no mediation pathway for TAU patients, only a direct effect of TAU on HRSD scores.

This is despite the fact that a measure, such as the DBT-WCCL score, is not truly a process variable because it is measured for patients in both the RO DBT and TAU groups. However, the skills that are relevant for TAU are different from those for RO DBT and, more generally, the other hypothesised pathways through which TAU affects HRSD scores are different from those for RO DBT. Hence, in line with the primary focus of this research on estimating the hypothesised pathways through which RO DBT works, no attempt to decompose the direct effect of TAU is made.

Decomposing the effect of radically open dialectical behaviour therapy into direct and indirect effects

In this section, the mediation models used to decompose the direct and indirect effects of RO DBT in the subsequent analysis are described. First, a single-mediator model is specified for analysing each of the three mediators in turn, followed by a joint model for the pathways involving all three of the hypothesised mediators.

Single-mediator model

For a given mediator M (which can be exposure, alliance or skills), the general form of the mediation model comprises the following model equations:

$$M = a_1 R + e_M, \tag{1}$$

and

$$Y = b_0 + b_1 R + b_2 M \times R + e_Y. \tag{2}$$

Equations 1 and *2* were fitted simultaneously to the data, as described in *Model fitting*. Baseline predictors can also be added to *Equations 1* and *2* to adjust for confounding bias, but the presence of such variables is suppressed herein to simplify notation.

Note that *Equation 1* for the mediator has no constant because M is a process variable and is exactly zero for TAU patients. As a result, the error term in *Equation 1* should be treated as heteroscedastic, because its variance is positive among the RO DBT patients but exactly zero for the TAU patients. Therefore, the joint model will be estimated using a heteroscedasticity-robust estimator.

Again, *Equation 2* for the HRSD outcome contains no mediator main effect because the mediator is a process variable that is always zero among those patients receiving TAU.

Effect decomposition for single-mediator models

The total effect of RO DBT is simply the sum of its direct (unmediated) effect and its indirect effect through the mediator. If it was assumed that there is no unobserved confounding, then the total effect will equal the intention to treat because treatment is randomly assigned in this study. The width of the CI for this estimate can be narrowed by including pre-randomisation variables that are associated with the outcome.

Standard results for mediator models[134] give:

$$\text{Total} = \text{direct} + \text{indirect} = b_1 + b_2 a_1, \tag{3}$$

where the regression coefficients a_1, b_1 and b_2 are those defined in *Equations 1* and *2*.

If both b_1 and $b_2 a_1$ are positive (or negative), then the proportion of the total effect mediated through M can be estimated; however, if the signs of the direct and indirect effects differ, then this summary becomes meaningless. Instead, the interesting question is whether or not focusing on the total effect could potentially be misleading if the direct effect is positive and the indirect effect is negative (or vice versa).

Joint all-mediators model

Recall that the mediators are denoted by X, A and S ('exposure', 'alliance' and 'skills', respectively). The joint model for all three mediators can be written as follows:

$$X = a_1 R + e_M, \tag{4}$$

$$A = (b_{11} + b_{12} X) \times R + e_A, \tag{5}$$

$$S = (b_{21} + b_{22} X) \times R + e_S, \tag{6}$$

$$Y = c_0 + c_1 R + c_2 X \times R + c_3 A \times R + c_4 S \times R + e_Y. \tag{7}$$

As with *Equations 1* and *2*, the error terms in *Equations 4–6* must be treated as heteroscedastic. The effects of the mediator variables enter the model through the interactions $X \times R$, $A \times R$ and $S \times R$, again, because the mediators are all treated as process variables.

Equations 4–7 are a model for the pathway diagram previously shown in *Figure 1*.

The purpose of this model is to decompose the mediated effect of RO DBT into (1) the unmediated effect of 'exposure' not passing through 'alliance' or 'skills', (2) the mediated effect of exposure working through 'alliance' and (3) the mediated effect of 'exposure' working through 'skills'.

Standard results for natural direct and indirect effects[135] give:

$$\text{Mediated} = \text{direct} + \text{indirect (alliance)} + \text{indirect (skills)} = a_1 c_2 + a_1 c_3 b_{12} + a_1 c_4 b_{22}, \tag{8}$$

where the coefficients are those in *Equations 4–7*.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Model fitting

The generalised method of moments (GMM) was used to fit these models.[136] The GMM is a flexible family of estimators for models of the form *Equations 1* and *2* and *Equations 4–7*. GMM is robust in the sense of producing valid estimates without having to assume that the residuals are normally distributed. All models are fitted using the command *gmm* in Stata.

Estimation if there is no unobserved confounding

In this section, a modification to the usual assumptions for pathway analysis is described.

To estimate a single-mediator model (*Equations 1* and *2*), two 'moment conditions' must be specified and a set of 'instruments' for each moment conditions:

1. The moment conditions for the single-mediator model are the residuals for *Equations 1* and *2*. For the all-mediators model, the moment conditions are the residuals for *Equations 4–7*.
2. The instruments are simply the independent variables used as predictors in each model equation. These are then multiplied together with the residuals to create an estimation equation to be solved. (Note that the GMM literature uses the term 'instrument' in a more general sense than the specialised 'IVs' to be introduced in the *Choosing instrumental variables*.)

If any model includes baseline predictors, then each baseline variable included in the model should be centred around its (complete-case) sample mean, even if it is a dummy variable, to ensure that the decompositions in *Equations 3* and *8* are correct. This is because the natural indirect effects under models that include baseline predictor variables will generally depend on the population average of these variables but mean-centring the variables removes this dependency.

Estimation if there is unobserved confounding

The main motivation for this analysis is the possibility that all three mediators are affected by unobserved confounding. The unobserved confounding is indicated in *Figure 1* by the latent variable *U*. This variable represents the combined effect of every factor influencing how patients select exposure, alliance and skills; these factors are also associated with the HRSD outcome, but have been omitted from the model.

Emsley *et al.*[130] (see also Have *et al.*[132] and Small[135]) described a framework for (single-mediator) mediation models using IVs that can be utilised to handle this problem. The definition and choice of IV(s) for the mediator variables in the analysis will be discussed. To implement the estimation approach of Emsley *et al.*[130] for the analysis, the *gmm* command was used as before but with the mediator variables in the set of GMM instruments replaced by the IVs.

This can be extended to the all-mediators model by replacing the mediators in *Equations 5–7* by IVs.

Choosing instrumental variables

An IV must be associated (preferably strongly) with the confounded mediators (in this case, alliance, exposure and skills) but not with the main outcome, either indirectly (by being associated with the omitted confounding variables *U*) or directly.

In classical IV studies, the randomised treatment variable is a good candidate for an IV primarily because randomisation ensures that treatment cannot be associated with the unobserved confounders. However, in this study, the explicit aim is to estimate the direct effect of treatment on the main outcome, which instantly precludes it from being an IV.

Instead, three putative IVs were incorporated into the design of the RefraMED trial. Specifically, as part of the trial, patients were randomised to receive a combination of feedback from the therapist, reimbursement for attendance at clinics and therapy in an enhanced room. Randomisation ensures that patients select these factors independently of the unobserved confounding variables, but the key issue is whether or not each IV is associated with at least one mediator.

Instruments for exposure to therapy, the therapeutic alliance and skills acquisition

Feedback

Research shows that when participants provided regular structured feedback to therapists regarding their progress, their perceptions of the alliance outcomes were improved.[137] Consequently, participants were randomised between providing and not providing feedback to therapists in each treatment session (Session Rating Scale) and their perceived progress (Outcome Rating Scale).

Mode of reimbursement

Because financial contingencies positively reinforce attendance at therapy sessions (thereby increasing exposure to the treatment), participants were randomised between receiving reimbursement cheques for completing research questionnaires within therapy sessions and receiving cheques to the same value by post from the trial manager. However, all participants received the same total reimbursement across the study.

Treatment setting

Because treatment setting is known to affect both adherence and expectancy for complementary therapies (e.g. acupuncture), and may also have the potential to influence impression formation (and thus the alliance), participants were allocated at random to the setting in which treatment is delivered: half the participants received treatment in standard NHS consulting rooms and the rest received treatment in enhanced rooms, simulating private consultation rooms used for many complementary therapies.

Therapists

Therapists are known to vary in their tendency to generate high or low alliance ratings from patients[138] and this may also influence the skills learned during treatment. Therefore, participants were randomised between therapists within centres, thereby adopting 'therapist' as an IV.

Other potential instrumental variables for alliance

Several other variables were measured at baseline to provide potential IVs for mediation analyses of exposure, alliance and skills:

- Therapist characteristics and client interactions – several characteristics of therapists, known to predict alliance within sessions, were measured to improve the prediction of therapeutic alliance from the random assignment. In therapists, a warm versus cold interpersonal style has been found to influence alliance scores.[139] Congruence of client and therapist characteristics (e.g. personal values[140] or cognitive style[141]) also predict alliance and outcomes. These variables were measured at baseline in both participants and therapists to provide additional potential IVs for the causal models.
- Pre-trial measures of alliance for trial therapists – to enter the study, therapists had to demonstrate adherence with non-trial patients. Mean alliance ratings from these sessions will be used to predict in-trial alliance scores.
- Travel distances from home to treatment centre – because travel times may predict drop out, this is a potential instrument for exposure. First-stage regressions will be conditioned on indicators of socioeconomic status of the home postcode to minimise the possibility that travel times are directly correlated with outcomes.

Instruments for skill acquisition

As the ability to learn and consistently apply new coping skills is hypothesised to play a crucial role in RO DBT treatment, it was important to identify variables with the potential to act as instruments for analyses examining mediation of treatment via skill learning.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Pre-treatment measure of behavioural compliance

Before randomisation, all patients were asked to complete a simple homework assignment, consistent with interventions from the positive psychological literature, at a set time each day. Adherence to this task was confirmed via automated telephone calls each day. The personality trait of Conscientiousness was also measured at baseline as a supplementary instrument for skill application.

Prospective memory

It was planned to measure prospective memory capacity at baseline as a potential IV for skill acquisition.

Weak instrumental variables

If a putative IV is not associated with any of the mediators, in the sense of having a zero correlation, then it cannot be an IV. However, in practice, the situation is less clear cut, as IVs will have small but non-zero correlations with the mediators. Such scenarios lead to a 'weak IV' problem, which can lead to badly biased estimates of the causal effect and its standard error.

With a large sample, this would not be a problem (unless the IV is truly independent of the mediator) but weak IVs can be a major problem for small samples, such as that in this study. Checks were carried out to establish the impact that weak IVs[142] have using the Stata function *ivregress*.

Direct effects of allocation

One important assumption of 'by treatment allocated' analyses is that of zero treatment value for participants allocated to a control condition. Specifically, inferences made about treatments in RCTs assume that the outcomes of participants who are allocated to a control condition do not differ as a function of their allocation, that is, these participants do no worse than if they had not been recruited to any trial and had been followed up covertly. Nonetheless, both research participation alone and reactions to the experimental procedures (e.g. to allocation itself, rather than the intervention under study) may bias treatment estimates.[143] Even when a trial offers bona fide alternative treatments, participants often express strong preferences for arms within RCTs and these preferences can, in and of themselves, influence trial outcomes.[144] When trials, such as the RefraMED trial, use TAU as a control then it is possible – perhaps likely – that participants will be disappointed if they are not allocated to the active intervention.

Cook and Campell[145] describe both 'resentful demoralisation' and 'compensatory rivalry' as potential sources of post-randomisation bias in which the allocation to treatment group cannot be fully concealed. Resentful demoralisation is the deleterious effect that randomisation to a control condition may have on disappointed participants, for example if they give up other attempts to change behaviour on hearing of their allocation. Compensatory rivalry denotes the reverse effect, for example if participants redouble efforts to find alternative sources of help if they are not allocated to their desired treatment. However, despite both the likelihood that participants may actively respond to treatment allocation and the concomitant potential for bias in unblinded intervention studies, few trials have engaged with these potential problems.[143]

In the RefraMED trial, it was explicitly sought to measure changes in patients' attitudes caused by allocation to treatment. This direct effect was operationalised as the change between participants' hypothetical expectancies before allocation (e.g. 'what will happen if you are assigned to RO DBT?') and actual expectancies after allocation (e.g. 'what will happen now you have been assigned to RO DBT [TAU]?'), when this change is different for participants assigned to RO DBT and TAU. These analyses are exploratory but are important because they seek to check the validity of the exclusion restriction for treatment allocation, upon which intention-to-treat and causal analyses depend.

Chapter 3 Results

Recruitment

Recruitment started at the Dorset site in March 2012, with the first participant randomised in April 2012. The Hampshire and North Wales sites started recruitment in September 2012 and their first participants were randomised in November 2012. Recruitment at all three sites continued until the end of April 2015. A total of 117 patients were recruited in Dorset, 95 in Hampshire and 38 in North Wales.

As can be seen in *Figure 2*, out of the 250 participants recruited into the trial, 170 (68%) were referred from secondary care services (either directly via clinicians or through database searches), 55 (22%) were referred from primary care database searches, 19 (8%) were self-referrals and 6 (2%) participants came from other sources (e.g. Family Intervention Project, private practitioners or family members).

Recruitment conversion rates

Of the 2678 patients referred to the study, 9% ($n = 250$) were eventually randomised. Conversion rates for referral sources into randomisations were 3% for primary care (55/1703), 20% for secondary care (170/849), 19% for self-referrals (19/101) and 24% for other sources (6/25).

Participant flow

Out of the 250 participants, 162 were allocated to RO DBT and 88 were allocated to TAU. *Figure 2* gives a detailed overview of participant flow and number of completed follow-up assessments at each time point. The numbers analysed differ slightly from the number of participants who attended assessments because of incomplete or lost data.

Study withdrawals, treatment withdrawals and losses to follow-up

Table 20 in *Appendix 1* presents a breakdown of reasons for study withdrawals (participants who indicated that they no longer wanted to provide outcome data) or losses to follow-up (participants who were uncontactable despite repeated attempts) per group for each assessment time point, as well as the number of participants allocated to RO DBT who withdrew from treatment (defined as having attended fewer than three sessions).

A total of 56 participants (22%) withdrew or were otherwise lost to follow-up during the study: 34 in the RO DBT group (21%) and 22 in the TAU group (25%). Unfortunately, only one of the participants withdrawing from treatment agreed to provide outcome data at follow-up interview and, thus, stayed in the study. Hence, there was no significant difference between treatment groups in the proportion of participants who did not contribute to analysis ($\chi^2 = 0.712$, degrees of freedom = 1; $p = 0.399$).

Of those participants allocated to RO DBT who withdrew from the study ($n = 34$), 10 participants did not attend any sessions and four withdrew from treatment after having attended fewer than three sessions and, thus, were considered treatment withdrawals.

The most common reasons for withdrawal for participants allocated to TAU was their dissatisfaction with having been allocated to TAU ($n = 9$), whereas for those allocated to RO DBT the most common reason for withdrawal was their being unable to commit to the study because of life, work or family commitments ($n = 10$).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

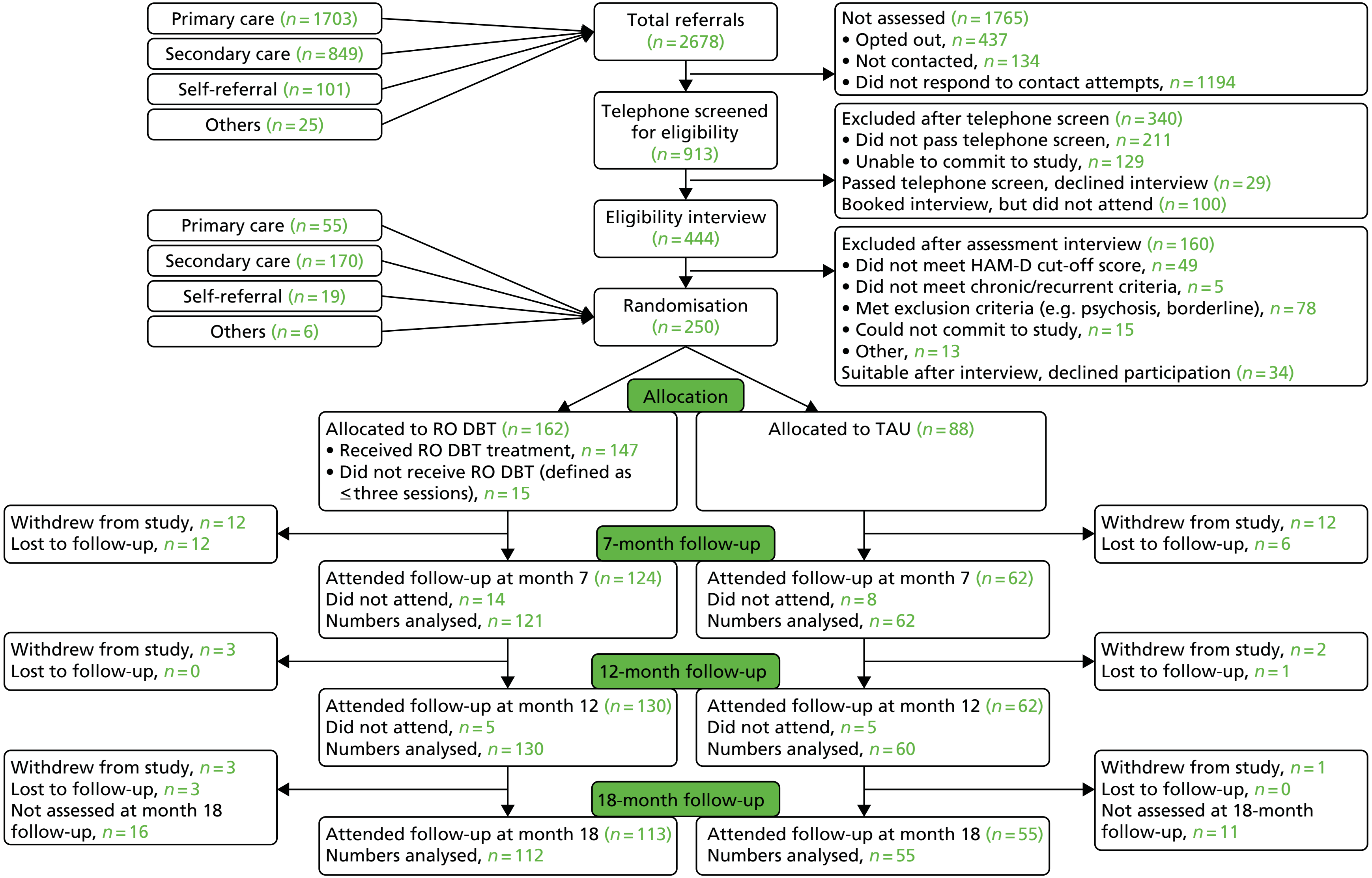
Primary care (n = 1703)
Secondary care (n = 849)
Self-referral (n = 101)
Others (n = 25)
Total referrals (n = 2678)
Not assessed (n = 1765)
• Opted out, n = 437
• Not contacted, n = 134
• Did not respond to contact attempts, n = 1194
Telephone screened for eligibility (n = 913)
Excluded after telephone screen (n = 340)
• Did not pass telephone screen, n = 211
• Unable to commit to study, n = 129
Passed telephone screen, declined interview (n = 29)
Booked interview, but did not attend (n = 100)
Eligibility interview (n = 444)
Excluded after assessment interview (n = 160)
• Did not meet HAM-D cut-off score, n = 49
• Did not meet chronic/recurrent criteria, n = 5
• Met exclusion criteria (e.g. psychosis, borderline), n = 78
• Could not commit to study, n = 15
• Other, n = 13
Suitable after interview, declined participation (n = 34)
Primary care (n = 55)
Secondary care (n = 170)
Self-referral (n = 19)
Others (n = 6)
Randomisation (n = 250)
Allocation
Allocated to RO DBT (n = 162)
• Received RO DBT treatment, n = 147
• Did not receive RO DBT (defined as ≤ three sessions), n = 15
Allocated to TAU (n = 88)
Withdrew from study, n = 12
Lost to follow-up, n = 12
7-month follow-up
Withdrew from study, n = 12
Lost to follow-up, n = 6
Attended follow-up at month 7 (n = 124)
Did not attend, n = 14
Numbers analysed, n = 121
Attended follow-up at month 7 (n = 62)
Did not attend, n = 8
Numbers analysed, n = 62
Withdrew from study, n = 3
Lost to follow-up, n = 0
12-month follow-up
Withdrew from study, n = 2
Lost to follow-up, n = 1
Attended follow-up at month 12 (n = 130)
Did not attend, n = 5
Numbers analysed, n = 130
Attended follow-up at month 12 (n = 62)
Did not attend, n = 5
Numbers analysed, n = 60
Withdrew from study, n = 3
Lost to follow-up, n = 3
Not assessed at month 18 follow-up, n = 16
18-month follow-up
Withdrew from study, n = 1
Lost to follow-up, n = 0
Not assessed at 18-month follow-up, n = 11
Attended follow-up at month 18 (n = 113)
Numbers analysed, n = 112
Attended follow-up at month 18 (n = 55)
Numbers analysed, n = 55

FIGURE 2 The Consolidated Standards of Reporting Trials (CONSORT) flow diagram showing the flow of study participants throughout the study and the numbers analysed per assessment point. Some participants did not attend at month 7 but did attend at month 12, for example. These participants were kept in the study and are not the same as the participants who withdrew.

Inter-rater reliability

Four assessment interviews, conducted by nine different individual assessors, were coded for inter-rater reliability for the HRSD scores. The reliability analyses were conducted at the item level, a more conservative method than using total scores. Across all measurements, Krippendorff's alpha was 0.89 (95% CI 0.86 to 0.92), suggesting that the raters attained 'very good' to 'near perfect' inter-rater reliability.[146] Ratings and any disagreements in coding were subsequently discussed as a group.

In addition, SSQ scores were coded by two independent raters and reliability analyses were performed for each of the three items listing the initials of those the participants identified as being members of their social support network. For each item Krippendorff's alpha was ≥ 0.9.

Demographic and baseline data

In accordance with standard practice, we conducted no statistical tests of demographic variables between groups.

Demographics

Table 5 presents baseline demographic data for the RO DBT and TAU groups. Of the entire sample, 164 participants were female (64%) and 193 participants (77%) were aged between 35 and 65 years. The vast majority of participants ($n = 222$; 93%) were of white British descent and 124 (50%) reported being in a relationship. With regard to highest education level, 113 participants (46%) had higher education experience. As can be seen from *Table 5*, demographics were comparable across groups.

Table 21 in *Appendix 1* shows the mean ages of the participants per group and per site; as can be seen, the age distributions per site per group did not differ greatly. The average age in the TAU group was 48 years (range 20–69 years) and in the RO DBT group was 47 years (range 21–72 years).

Furler *et al.*[147] note that demographic information reported by clinical trials typically fails to include data on participants' socioeconomic status but that this can be important for the interpretation and generalisation of trial results. We used participants' postcodes to look up values from publicly available geographically based indices of deprivation in England and Wales. For England, we used the *English Indices of Deprivation* for 2014[148] and for Wales, the revised *Welsh Index for Multiple Deprivation* for 2014.[149] Data were available for 224 out of the 250 participants. *Figure 8* in *Appendix 4* presents these data split by treatment group; we found no evidence that the treatment groups differed on these indices. Decile scores are not directly comparable between England and Wales.[150] However, overall, the Welsh sample was close to the Welsh national average on the *Welsh Index of Multiple Deprivation* [mean decile 5.5, standard deviation (SD) 2.5]. The two English sites were slightly more deprived relative to other areas [the mean decile for the *English Indices of Deprivation* for Dorset was 5.8 (SD 2.4) and for Hampshire was 7.4 (SD 2.5)].

Clinical backgrounds

Table 6 presents clinical background data for the RO DBT and TAU groups. More than one-third of the participants (38%) reported that they had experienced their first depressive episode before the age of 16 years. Eighty-five per cent of participants were classed as having chronic rather than recurrent depression at the time of interview, and of those that could recall the number of previous episodes that they had experienced, one-third (34%) reported no previous episodes, one-third (35%) reported six or more and the remainder of participants reported between one and five previous episodes. The vast majority of participants (83%) had previously received psychotherapy.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 5 Demographic data per treatment group allocation

Demographic variable	Treatment group TAU *n*	*N*	%	RO DBT *n*	*N*	%
Sex						
Female: true	53	88	60.2	111	162	68.5
Age category (years)						
17–25	2	87	2.3	6	162	3.7
25–35	13	87	14.9	21	162	13.0
35–45	14	87	16.1	33	162	20.4
45–55	38	87	43.7	53	162	32.7
55–65	16	87	18.4	39	162	24.1
65–100	4	87	4.6	10	162	6.2
Income (£)						
0–5000	21	79	26.6	38	142	26.8
5001–10,000	13	79	16.5	23	142	16.2
10,001–20,000	14	79	17.7	32	142	22.5
20,001–32,000	17	79	21.5	26	142	18.3
32,001–50,000	9	79	11.4	14	142	9.9
> 50,000	5	79	6.3	9	142	6.3
Education						
Left school before 16 years	13	84	15.5	24	157	15.3
Finished school at 16 years	26	84	31.0	41	157	26.1
Finished school at 18 years	8	84	9.5	16	157	10.2
Attended university or equivalent	11	84	13.1	20	157	12.7
Completed university or equivalent	14	84	16.7	29	157	18.5
Completed postgraduate qualification	12	84	14.3	27	157	17.2
Ethnicity						
White: British	78	83	94.0	144	155	92.9
White: other	2	83	2.4	8	155	5.2
Other	3	83	3.6	3	155	1.9
Marital status						
Divorced or separated: true	20	88	22.7	42	162	25.9
Married or cohabiting: true	40	88	45.5	66	162	40.7
Single: true	28	88	31.8	54	162	33.3

TABLE 6 Basic clinical data per treatment group allocation

	Treatment group					
	TAU			RO DBT		
Clinical variable	*n*	*N*	%	*n*	*N*	%
Number of past depressive episodes						
0	17	57	29.8	44	111	39.6
1	4	57	7.0	7	111	6.3
2–3	10	57	17.5	14	111	12.6
4–5	3	57	5.3	13	111	11.7
≥ 6	23	57	40.4	33	111	29.7
Chronic vs. recurrent depression						
Chronic	61	70	87.1	118	143	82.5
Recurrent	9	70	12.9	25	143	17.5
Age (years) at onset categorised						
< 16	35	88	39.8	57	162	35.2
16–21	13	88	14.8	36	162	22.2
22–30	19	88	21.6	31	162	19.1
31–50	20	88	22.7	34	162	21.0
≥ 50	1	88	1.1	4	162	2.5
Previous treatment						
Psychotherapy received	72	84	85.7	119	150	79.3
Comorbid DSM-IV disorders						
Axis I						
Generalised anxiety disorder	40	85	47.1	79	154	51.3
Social phobia	42	51	82.4	78	100	78.0
Panic disorder	29	50	58.0	50	95	52.6
Specific phobia	22	34	64.7	42	68	61.8
Agoraphobia without history of panic disorder	10	20	50.0	23	55	41.8
Obsessive–compulsive disorder	15	31	48.4	30	66	45.5
Post-traumatic stress disorder	21	37	56.8	34	76	44.7
Anxiety disorder not otherwise specified	6	16	37.5	5	35	14.3
Anorexia nervosa	2	13	15.4	9	34	26.5
Bulimia nervosa	1	15	6.7	3	26	11.5
Binge-eating disorder	5	16	31.2	14	33	42.4
Alcohol abuse/dependence	1	34	2.9	0	68	0.0
Cocaine abuse/dependence	1	17	5.9	0	48	0.0
Hypochondriasis	2	84	2.4	7	155	4.5
Body dysmorphic disorder	11	84	13.1	15	156	9.6
Pain disorder	4	84	4.8	7	154	4.6
Somatisation disorder	2	84	2.4	6	152	4.0
Somatoform disorder	0	84	0.0	3	154	2.0

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 6 Basic clinical data per treatment group allocation (*continued*)

	Treatment group					
	TAU			RO DBT		
Clinical variable	*n*	*N*	%	*n*	*N*	%
Bipolar I disorder	0	25	0.0	2	64	3.1
Any other current Axis-I disorder	1	7	14.3	0	7	0.0
Axis II						
Schizotypal PD	6	88	6.8	16	161	9.9
Paranoid PD	18	88	20.5	39	161	24.2
Schizoid PD	9	85	10.6	16	159	10.1
Avoidant PD	46	88	52.3	90	162	55.6
Obsessive–compulsive PD	41	87	47.1	75	162	46.3
Dependent PD	6	86	7.0	17	161	10.6
Conduct disorder	6	74	8.1	11	150	7.3
Total[a]						
Axis I						
0	14	88	15.9	19	162	11.7
1	13	88	14.8	34	162	21.0
2	21	88	23.9	39	162	24.1
3	18	88	20.5	29	162	17.9
4	11	88	12.5	17	162	10.5
5	6	88	6.8	11	162	6.8
6	3	88	3.4	6	162	3.7
7	2	88	2.3	7	162	4.3
Axis II						
0	20	88	22.7	33	162	20.4
1	24	88	27.3	46	162	28.4
2	29	88	33.0	48	162	29.6
3	11	88	12.5	23	162	14.2
4	3	88	3.4	8	162	4.9
5	1	88	1.1	3	162	1.9
6	0	88	0.0	1	162	0.6
Axes I and II						
0	4	88	4.5	5	162	3.1
1 or 2	25	88	28.4	44	162	27.2
3 or 4	25	88	28.4	54	162	33.3
5–7	29	88	33.0	41	162	25.3
8–10	5	88	5.7	14	162	8.6
11 or 12	0	88	0.0	4	162	2.5

a Number of patients experiencing several mental health disorders listed separately for Axis I, clinical disorders (ranging from 0 to 7) and Axis II, personality disorders (ranging from 0 to 6), as well as the number of patients experiencing combinations of both Axis I and Axis II disorders (ranging from 0 to 12).

High comorbidity rates were present in the sample: only 3.6% of participants did not report any comorbid disorders; the remaining 96.4% were experiencing an Axis-I, Axis-II or a combination of both types of mental health disorder. Eighty-six per cent reported at least one comorbid Axis-I disorder and 78% reported at least one comorbid Axis-II disorder. The two most common Axis-I disorders were social phobia (80%) and specific phobia (63%), although other anxiety disorders were also very common in this sample. The two most common PDs were avoidant PD (54%) and obsessive–compulsive PD (47%).

As can be seen in *Table 6*, two participants were assessed as having bipolar I disorder. Although established bipolar disorder was an exclusion criterion, these participant's notes indicated that both were experiencing a moderately severe depressive episode at the time of study inclusion (without rapid cycling and without seasonal pattern). They were, thus, deemed eligible, recruited and included in the analyses.

Outcome of adaptive randomisation procedures

The adaptive randomisation procedure appears to have been effective in balancing the clinical backgrounds of the participants across groups (see *Table 6*). Treatment sites were also comparable in the clinical characteristics of participants (see *Table 22* in *Appendix 1*).

Provision of therapy

A total of 22 therapists saw participants for individual treatment, with the number of participants allocated to each therapist ranging from 1 to 19 and the number of participants being seen by each therapist ranging from 1 to 17. The average number of participants seen per therapist was 6.9 and the median was 7.

Patients' treatment attendance

Individual treatment sessions

Table 23 in *Appendix 1* shows treatment attendance for individual sessions at each of the three treatment sites. Between 82% and 91% of planned individual sessions were attended. Participants were 5% more likely to attend (88% vs. 83%) when their regular therapist was available than when the session was offered by a back-up therapist (odds ratio 0.66; $p = 0.04$).

Of the missing sessions that were reported to the study team, the most common reasons for not attending sessions were sickness (29%), holidays or being out of town (28%), patients not wanting to come (17%), patients forgetting (3%) or the therapist not knowing why their client did not attend the session (23%).

As can be seen in *Figure 9* in *Appendix 4*, apart from those participants who prematurely withdrew from treatment, missed sessions did not follow a specific pattern. In other words, there did not seem to be a particular time during the treatment schedule that participants tended to attend less frequently.

Skills groups

Group attendance did not differ between sites. The mean number of sessions attended was 19.7 in Dorset (median 23.0 sessions), 20.0 in Hampshire (median 23.0 sessions) and 20.5 in North Wales (median 21.5 sessions). *Table 24* in *Appendix 1* gives a more detailed overview.

Fidelity to treatment manual: therapists' treatment adherence

Treatment adherence was reassessed throughout the trial by coding a random sample of sessions. A total of 273 (9%) tapes were rated, aiming to rate no more than three tapes from the same therapist–client combination. Eighty-one per cent of all tapes were considered adherent according to the standard DBT adherence rating scale (i.e. had a score of ≥ 4.0 points). The average score was 4.1 ± 0.2 points.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Adverse events

A total of 33 SAEs were reported to the trial manager: four for participants allocated to TAU, 28 for those allocated to RO DBT and one for a participant who had not yet been randomised. Of the 28 events reported within the RO DBT group, seven incidents related to three participants (two participants had two incidents and one participant had three incidents); thus, SAEs were reported for 25 individual participants in the RO DBT group.

Of the total number of reported incidents, it was judged that 17 (TAU, *n* = 4; RO DBT, *n* = 13) were definitely not related to the study intervention (e.g. heart problems and broken leg) and eight were unlikely to be related (e.g. long-standing non-suicidal self-injury, non-fatal overdose without intent to die and suicide attempts). It was judged that, of the remaining eight SAEs, all at which were from the intervention group, five were possibly related to RO DBT and three were probably related to RO DBT (*Table 7*). Nevertheless, none of the resulting eight serious adverse reactions (SARs) was classed as 'suspected unexpected' or required immediate reporting to the Research Ethics Committee.

TABLE 7 Adverse events reported to trial manager

Group	Number of adverse events reported	Source of report	Nature of SAE or SAR
Pre-randomisation (total *n* = 1)			
Definitely not related (SAE)	1	Patient (*n* = 1)	• Hospitalisation because of suicide attempt – pre-randomisation (*n* = 1)
TAU (total *n* = 4)			
Definitely not related (SAE)	4	Patient (*n* = 4)	• Hospitalisation because of physical illness (*n* = 2) • Hospitalisation because of suicide attempt (*n* = 1) • Disclosure of severe suicidal ideation (*n* = 1)
RO DBT (total *n* = 28)			
Definitely not related (SAE)	12	RO DBT clinician (*n* = 9) and patient (*n* = 3)	• Hospitalisation because of physical illness, traffic accident or fall (*n* = 8) • Hospitalisation because of suicide attempt – pre-treatment (*n* = 1) • Disclosure of thoughts about harming others (*n* = 1) • Fear of accidental overdose pain medication – no medical attention required (*n* = 1) • Death because of natural causes (*n* = 1)
Unlikely to be related (SAE)	8	RO DBT clinician (*n* = 6) and patient (*n* = 2)	• Hospitalisation because of suicide attempt – long-standing history (*n* = 4) • Self-harming that required medical attention – long-standing history (*n* = 1) • Disclosure of self-harming that did not require medical attention – long-standing history (*n* = 2) • Disclosure of severe suicidal ideation – long-standing history (*n* = 1)
Possibly related (SAR)	5	RO DBT clinician (*n* = 5)	• Overdose with suicidal intent (*n* = 1) • Overdose without suicidal intent (*n* = 1) • Disclosure of thoughts about harming others (*n* = 3)
Probably related (SAR)	3	RO DBT clinician (*n* = 3)	• Started non-suicidal self-injury (*n* = 2) • Suspected suicide attempt – prevented (*n* = 1)

Thus, all eight SARs occurred in the intervention group (Fisher's exact test; one-tailed; $p < 0.01$). However, TAU group participants were seen by trial assessors only at the three follow-up interviews and SAE reporting relied on the participants volunteering the relevant information. In contrast, RO DBT participants were also seen twice a week by trial therapists. An attempt was made to prevent reporting bias by asking participants' GPs to report any AEs that their patients experienced; however, none of the SAEs was reported by GPs or anyone else outside the RefraMED team. Instead, all four SAEs in the TAU group were reported by the participants themselves either during an assessment or after having been contacted by the trial office. In contrast, of the SAEs reported for participants in the RO DBT group, the vast majority ($n = 23$; 82%) were reported by the treating RO DBT clinicians, whereas only five SAEs were reported by the participants themselves, either during an assessment or after having been contacted by the trial office. It is believed that the imbalance was a result of this gross difference in reporting opportunities, a form of reporting bias. In fact, the DMEC concluded at its final meeting at which SAEs were discussed, in January 2015, that 'there was no reason to suspect that RO DBT had adverse effects on patients'.

Continuous outcomes

Number of participants analysed

All analyses were 'by treatment allocated'. *Figure 9* in *Appendix 4* presents the proportion of data available for analyses at baseline and each of the main follow-up points. For the primary outcome, the HRSD score, data were available for all participants at baseline, for 73% of participants at month 7 and for 76% of patients at month 12. Fewer data points (67%) were available at month 18 by design, because data collection ended when all participants had completed the 12-month follow-up, but not all had completed the 18-month follow-up.

Primary outcome: depression symptoms measured by the Hamilton Rating Scale for Depression

We did not find a significant interaction between treatment allocation and time. Planned between-group contrasts revealed a statistically significant difference between RO DBT and TAU at month 7, but not at months 12 or 18 (*Table 8*). Standardised differences between groups were large at month 7 ($d = 1.03$). At 12 months, the effect size matched that used in the power calculations ($d = 0.41$) but was slightly reduced at month 18 ($d = 0.32$). Model-adjusted predictions for each treatment group are presented in *Figure 3*. These predictions suggest that treatment gains for RO DBT participants were maintained between month 7 and month 18, but that TAU participants also experienced reductions in symptoms between months 7 and 12. The descriptive statistics of the HRSD scores are presented in *Table 9* and the model-adjusted means in *Table 10*.

Variations in therapist performance

The variance partition coefficient was computed for the variance associated with therapists from the random-effects terms in the mixed models; the variance partition coefficient represents the proportion of variance associated with a particular 'level' of the model. Because conventional maximum likelihood fitting can perform poorly when estimating therapist variance components, the variance partition coefficient was calculated using the Bayesian models described in *Chapter 2*, *Clinical predictions for new patients*. Sensitivity analyses showed that these Hamiltonian Monte Carlo-fit models were slightly more conservative than standard maximum-likelihood model fits.

For HRSD score, approximately 12% of the variance in outcome was attributable to differences between therapists (see *Table 25* in *Appendix 1* for further details). Participants allocated to the best-performing therapist benefited by about 2.54 points on the HRSD (95% highest probability density interval –0.75 to 6.68 points) compared with those allocated to the worst performing therapist, and this equates to a standardised difference of 0.48 (95% highest probability density interval –0.14 to 1.28). Relative differences in therapist effectiveness are shown in *Figure 11* in *Appendix 4*.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

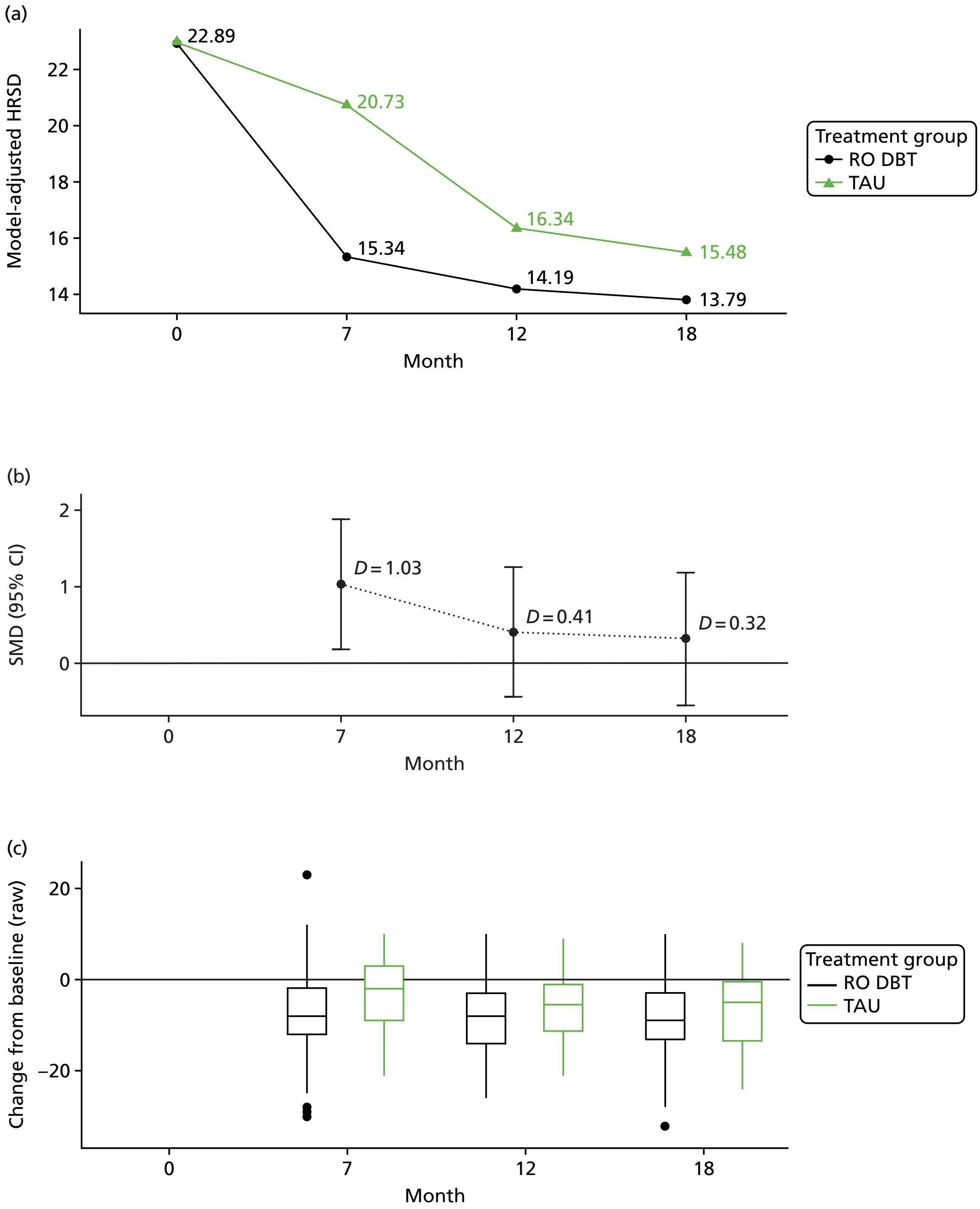

FIGURE 3 Model-adjusted predictions for each treatment group based on HRSD score. (a) Adjusted predictions; (b) standardised mean differences between groups; and (c) change from baseline. SMD, standardised mean difference.

TABLE 8 Planned contrasts between TAU and RO DBT groups at 7, 12 and 18 months for depression scores (as measured using the HRSD), psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ)

Contrast (month)	Delta	Standard error	*t*	95% CI	*p*-value	SMD (Cohen's *d*)	SMD	
							Upper CI	Lower CI
AAQ-II								
7	–3.37	1.30	–2.59	–5.92 to –0.82	*0.010*	–0.49	–0.86	–0.12
12	–4.94	1.31	–3.76	–7.53 to –2.36	*0.000*	–0.72	–1.09	–0.34
18	–5.48	1.66	–3.30	–8.74 to –2.22	*0.001*	–0.79	–1.27	–0.32
EAC								
7	1.50	0.71	2.10	0.09 to 2.90	*0.036*	0.32	0.02	0.62
12	3.55	0.72	4.94	2.14 to 4.95	*0.000*	0.76	0.46	1.06
18	2.98	0.94	3.18	1.14 to 4.82	*0.002*	0.64	0.25	1.04
HRSD								
7	–5.40	1.92	–2.80	–9.85 to –0.94	*0.024*	–1.03	–1.88	–0.18
12	–2.15	1.89	–1.14	–6.59 to 2.28	0.290	–0.41	–1.26	0.44
18	–1.69	2.02	–0.84	–6.22 to 2.84	0.424	–0.32	–1.19	0.54
LIFE-RIFT								
7	–0.72	0.80	–0.90	–2.85 to 1.41	0.412	–0.23	–0.92	0.46
12	0.08	0.77	0.10	–2.10 to 2.25	0.924	0.03	–0.68	0.73
18	0.19	0.88	0.21	–1.92 to 2.30	0.838	0.06	–0.62	0.74
MSSI								
7	–1.84	1.17	–1.57	–4.14 to 0.46	0.117	–0.18	–0.39	0.04
12	–1.44	1.09	–1.32	–3.58 to 0.70	0.187	–0.14	–0.34	0.07
18	0.45	1.39	0.32	–2.29 to 3.18	0.748	0.04	–0.22	0.30
SSQ								
7	1.91	1.62	1.18	–1.93 to 5.75	0.276	0.23	–0.23	0.68
12	2.60	1.63	1.59	–1.25 to 6.44	0.155	0.31	–0.15	0.77
18	3.44	1.87	1.83	–0.63 to 7.51	0.091	0.41	–0.08	0.89

SMD, standardised mean difference.
p-values in italics indicate significant difference between groups at the $\alpha = 0.05$ level.

Other continuous outcomes

We also ran mixed-effect models and computed contrasts at months 7, 12 and 18 for psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ). *Table 8* presents the planned contrasts between TAU and RO DBT at 7, 12 and 18 months for all of the outcomes.

Table 9 presents the descriptive statistics for all variables, whereas *Table 10* presents the adjusted means for these models. We found significant advantages for the RO DBT patients on EAC and AAQ-II scores at months 7, 12 and 18. AAQ-II scores for RO DBT patients again decreased consistently, with medium to large differences between TAU and RO DBT groups (see *Figure 12* in *Appendix 4*). For EAC, the RO DBT group mean score increased consistently and the standardised difference grew from small to moderate/large by 18 months (see *Figure 13* in *Appendix 4*).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 9 Descriptive statistics of depression scores (as measured using the HRSD), psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ) presented per group per month

		Centile						
Outcome, time point	Mean score (SD)	Minimum	10th	25th	Median	75th	90th	Maximum
HRSD								
Month 0								
RO DBT	22.96 (5.27)	15	16.00	19.00	23.0	27.0	30.9	36
TAU	23.48 (5.20)	15	17.00	19.00	23.0	27.0	30.0	37
Month 7								
RO DBT	15.97 (8.80)	1	6.00	9.00	15.0	22.0	28.0	38
TAU	19.74 (7.22)	4	9.10	14.25	21.0	24.8	28.0	35
Month 12								
RO DBT	14.70 (8.86)	0	3.00	8.25	14.0	20.0	26.1	39
TAU	16.05 (6.71)	5	7.90	10.75	15.0	21.2	25.0	29
Month 18								
RO DBT	14.13 (8.34)	0	3.10	7.75	13.0	22.2	25.0	29
TAU	15.67 (7.32)	1	4.00	11.00	17.0	21.5	24.0	30
MSSI								
Month 0								
RO DBT	8.06 (11.00)	0	0.00	0.00	2.0	14.2	26.1	43
TAU	7.09 (9.60)	0	0.00	0.00	2.0	12.5	21.4	42
Month 7								
RO DBT	4.78 (9.12)	0	0.00	0.00	1.0	2.0	18.2	37
TAU	5.25 (9.89)	0	0.00	0.00	1.0	3.0	19.3	47
Month 12								
RO DBT	2.56 (6.26)	0	0.00	0.00	0.0	1.0	5.6	36
TAU	2.30 (6.11)	0	0.00	0.00	1.0	1.0	3.0	32
Month 18								
RO DBT	2.67 (7.31)	0	0.00	0.00	1.0	2.0	3.0	40
TAU	2.45 (8.39)	0	0.00	0.00	0.0	1.0	2.0	47
LIFE-RIFT								
Month 0								
RO DBT	15.82 (3.12)	4	12.00	14.00	16.0	18.0	20.0	22
TAU	16.06 (3.04)	5	13.00	14.00	16.0	18.0	20.0	22
Month 7								
RO DBT	12.82 (4.51)	4	6.00	9.00	13.0	16.0	19.0	20
TAU	13.79 (3.92)	4	8.00	12.00	14.0	17.0	19.0	20
Month 12								
RO DBT	11.92 (4.58)	4	6.00	8.75	12.0	15.0	18.3	22
TAU	12.02 (4.61)	3	6.00	8.00	12.5	16.0	17.0	19

TABLE 9 Descriptive statistics of depression scores (as measured using the HRSD), psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ) presented per group per month *(continued)*

		Centile						
Outcome, time point	Mean score (SD)	Minimum	10th	25th	Median	75th	90th	Maximum
Month 18								
RO DBT	12.16 (4.84)	3	5.00	9.00	12.0	16.0	18.0	20
TAU	11.78 (4.42)	4	6.00	8.00	12.0	15.0	17.0	23
EAC								
Month 0								
RO DBT	15.90 (4.68)	8	10.00	12.00	16.0	19.0	22.0	32
TAU	16.47 (4.60)	8	10.00	13.33	16.5	20.0	23.0	25
Month 7								
RO DBT	18.66 (5.07)	8	12.80	16.00	18.8	23.0	25.0	32
TAU	16.92 (5.41)	8	9.22	11.75	18.0	20.0	23.5	27
Month 12								
RO DBT	18.64 (5.33)	8	12.00	15.00	19.0	22.0	25.0	32
TAU	15.21 (5.78)	3	8.00	10.16	16.0	19.0	23.0	29
Month 18								
RO DBT	18.98 (5.31)	8	11.00	15.50	19.0	23.0	25.6	31
TAU	16.63 (4.74)	8	10.50	13.00	17.0	19.8	22.0	27
AAQ-II								
Month 0								
RO DBT	38.23 (7.18)	17	29.00	33.50	39.0	44.0	47.6	49
TAU	39.10 (6.33)	22	31.00	34.00	39.0	44.0	47.8	49
Month 7								
RO DBT	32.86 (9.75)	7	19.00	27.00	33.0	40.0	46.3	49
TAU	37.05 (7.66)	18	26.00	32.90	39.0	42.2	45.0	49
Month 12								
RO DBT	32.03 (10.27)	9	18.00	24.00	32.0	41.0	45.0	49
TAU	37.11 (8.41)	20	26.00	29.00	38.0	45.0	47.0	49
Month 18								
RO DBT	31.58 (10.61)	9	17.00	23.00	33.0	39.0	46.0	49
TAU	36.11 (8.88)	16	22.50	29.00	38.0	43.0	47.5	49
SSQ								
Month 0								
RO DBT	18.89 (8.79)	0	7.00	13.00	20.0	25.0	30.0	39
TAU	18.88 (7.75)	0	9.00	14.00	19.0	23.0	27.0	37
Month 7								
RO DBT	20.59 (8.90)	3	10.00	14.00	20.0	26.0	32.0	43
TAU	19.67 (8.54)	0	8.40	14.00	21.0	26.0	29.0	42

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 9 Descriptive statistics of depression scores (as measured using the HRSD), psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ) presented per group per month *(continued)*

		Centile						
Outcome, time point	Mean score (SD)	Minimum	10th	25th	Median	75th	90th	Maximum
Month 12								
RO DBT	22.16 (7.90)	3	12.00	16.25	22.0	28.0	32.0	43
TAU	19.64 (9.28)	0	8.00	12.25	20.0	26.0	31.0	40
Month 18								
RO DBT	21.75 (8.64)	2	11.00	16.25	21.0	28.0	32.0	42
TAU	19.33 (9.31)	0	9.20	13.50	19.0	24.0	31.0	44

TABLE 10 Model-adjusted means by group at each time point for depression scores (as measured using the HRSD), psychological inflexibility (as measured using the AAQ-II), emotional approach coping (as measured using the EAC), impaired functioning (as measured using the LIFE-RIFT), suicide ideation (as measured using the MSSI) and social support (as measured using the SSQ)

	Treatment group			
	RO DBT		TAU	
Outcome, time point	Mean score	Standard error	Mean score	Standard error
AAQ-II				
7	33.12	0.742	36.48	1.067
12	32.56	0.762	37.51	1.072
18	30.69	0.786	36.17	1.462
EAC				
7	18.38	0.407	16.89	0.587
12	18.64	0.411	15.10	0.588
18	19.31	0.441	16.33	0.826
HRSD				
7	15.34	0.739	16.89	1.776
12	14.19	0.695	15.10	1.754
18	13.79	0.758	16.33	1.873
LIFE-RIFT				
7	12.73	0.386	13.46	0.701
12	11.96	0.359	11.88	0.681
18	12.38	0.398	12.19	0.788
MSSI				
7	4.66	0.651	6.50	0.975
12	2.46	0.569	3.90	0.931
18	2.04	0.663	1.59	1.223
SSQ				
7	21.50	0.680	19.59	1.465
12	22.33	0.686	19.73	1.482
18	22.49	0.720	19.05	1.731

No significant advantage was found for RO DBT on LIFE-RIFT, MSSI or SSQ scores at any follow-up point. Both treatment groups exhibited substantial reductions in LIFE-RIFT (impaired functioning) scores from baseline (see *Figure 14* in *Appendix 4*). In both treatment groups, the mean MSSI scores at baseline were below eight, indicating low levels of suicidal ideation, and this remained low throughout the trial (see *Figure 15* in *Appendix 4*). Perceived social support increased over time in the RO DBT group but not the TAU group, although this difference did not reach significance (see *Figure 16* in *Appendix 4*).

Other outcomes

Remission rates and reliable change

Remission rates calculated by the primary criteria, and by other criteria reported in the literature, are given in *Table 11* and presented in *Figure 17* in *Appendix 4*. There were substantial differences in absolute rates of remission depending on the criteria adopted, and also whether or not we calculated remission based on model-adjusted predictions or raw scores. This is likely because the mixed models tend to 'shrink' extreme values towards the cluster and grand mean.[151] Depending on the criterion used (from the many described in the literature), the effect of this shrinkage is more or less marked. For the purposes of comparison between groups, it was believed that the model-estimated remission rates are preferable, because they adjust for differences in severity at baseline in the same way as the model-based inferences reported as the primary outcome.

TABLE 11 Remission rates (%) per group and follow-up month using model-adjusted and raw data

	Month (% remission)					
	7		12		18	
Criteria	RO DBT	TAU	RO DBT	TAU	RO DBT	TAU
Adjusted data						
Change						
17.5%	59	27	69	48	59	41
2 SD	14	2	20	14	19	16
50%	7	–	17	2	18	8
RCI	43	16	56	31	49	30
Remission						
Full	1	–	8	–	7	1
Full or partial	23	6	36	22	33	24
Raw data						
Change						
17.5%	49	30	56	41	50	38
2 SD	22	14	30	20	26	19
Change 50%	22	10	31	19	27	19
Change RCI	40	24	48	33	44	26
Remission						
Full	15	3	18	7	17	12
Full or partial	25	12	36	20	31	22

RCI, reliable change index.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

This variation in remission rates as a result of choice of methodology and criterion highlights the substantial disadvantage of dichotomising or trichotomising continuous outcome data; although remission provides a simple and tangible metric of the performance of a treatment, this may come at the cost of oversimplification. Nonetheless, to aid comparison with other trials, remission rates were reported based on a variety of criteria in *Table 11* and *Figure 17* in *Appendix 4*. Using the primary criterion for remission, which was based on both the HRSD and LIFE-RIFT scores, more RO DBT participants achieved full or at least partial remission at months 7, 12 and 18. At the primary follow-up (month 12), and using model-adjusted data, this equated to around eight additional RO DBT participants achieving full remission for every 100 treated, and so had a number needed to treat to achieve an additional fully remitted case of 12.5. For partial remission, an additional 14 participants benefited from RO DBT, with a number needed to treat of 7.1. Using the secondary criterion of 'worthwhile' change from the participant perspective, which equates to a 17.5% reduction in symptoms from baseline,[76] it was found that an additional 21 participants per 100 benefited from RO DBT.

Clinical predictions for new patients

From the perspective of clinicians and patients, it is relevant to consider the likely range of outcomes that an individual patient might experience, given different treatment choices. For example, patients might wish to know what benefit they are likely to obtain in choosing a new treatment, compared with continuing in TAU. By simulation from the primary outcome model, fit by Hamiltonian Monte Carlo, predictions were made for new patients who would be treated with either TAU or RO DBT. These predictions show good convergence, with group means from models fit via maximum likelihood (see *Figure 3*), especially at months 7 and 12. At month 18, the predictions from this Bayesian model are more pessimistic than the equivalent model fit via maximum likelihood, with median predictions at 18 months of 18.12 and 19.17 points for RO DBT and TAU, respectively, which compares with 13.96 (RO DBT) and 16.09 points (TAU) for the equivalent maximum-likelihood model. These discrepancies are believed to be primarily caused by the larger number of missing data points at month 18 (part of the study design) and the different assumptions embodied by the priors in the Bayesian model, which influence how data are pooled across time periods. These predictions are also presented in *Table 26* in *Appendix 1*.

Based on these predictions, we computed the difference between outcomes if patients had been treated with RO DBT rather than TAU and, thus, the *expected additional benefit* of being treated with RO DBT for a new patient. Inspecting the distribution of this predicted additional benefit allows patients to form an expectation for what their outcome might be should they accept RO DBT treatment over TAU. *Figure 4* presents expected changes in symptoms for the 5th through to the 95th centiles of patients graphically, and *Table 27* in *Appendix 1* presents the same figures in a table. The 5th centile represents the additional benefit of RO DBT in the 'best-case scenario' – only 5% of patients would experience a benefit this great.

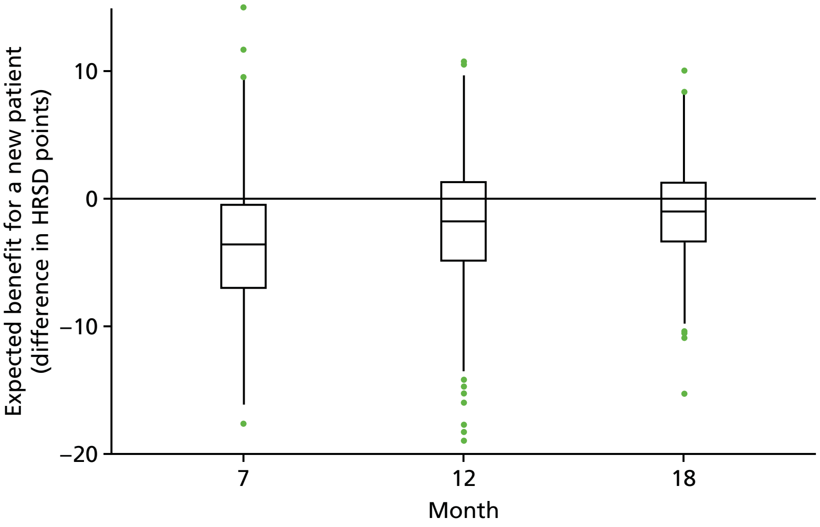

FIGURE 4 The expected benefits for new patients accepting RO DBT rather than TAU at 7, 12 and 18 months.

Conversely, the 95th centile represents the expected change in symptoms for the least-fortunate 5% of patients. The 50th centile represents the expected change for the average patient. Both *Figure 4* and *Table 27* in *Appendix 1* demonstrate how variable prognosis is from the patients' perspective. When choosing RO DBT rather than TAU, the models show that patients should expect to make additional improvements in symptoms, but the intervals on this prediction are wide.

Table 12 and *Figure 18* in *Appendix 4* represent these predicted values as the percentage of participants who achieve 'worthwhile' change, using Button's criteria[76] of a 17.5% change from baseline; in the sample, the mean HRSD value at baseline was 23 points, and so this required, on average, a 4-point change on the HRSD. From this we can say that 50 out of 100 patients would experience a worthwhile benefit from accepting RO DBT, rather than TAU, at the end of treatment. At month 12, despite the improvements in the TAU group over this period, 32 out of 100 patients would still obtain a worthwhile reduction in symptoms from accepting RO DBT. Note that in *Table 12* (and see *Figure 4*) the 'worsening' of symptoms does not refer to an iatrogenic effect of RO DBT but rather reflects the fact that predictions for *new* patients must take into account the natural variation in depressive symptoms; failure to account for this natural variation in symptoms would lead to predictions for prognosis that were unjustifiably confident.[152]

Health economics analyses

At the 12-month follow-up, full service-use data were available for the entire follow-up period for 125 participants in the RO DBT group (77%) and 61 in the TAU group (69%). This is compared with complete data in the RO DBT group for 118 participants (73%) at 7 months and 101 participants (62%) at 18 months and in the TAU group for 61 participants (69%) at 7 months and 51 participants (58%) at 18 months. There were no statistically significant differences between groups in the proportion of missing data ($\chi^2 = 0.25$; $p = 0.61$). Missing resource-use items in completed questionnaires were assumed to indicate no service use and were given a zero value. Multiple imputation methods were used to replace missing EQ-5D-3L, cost and HRSD score values, assuming that data were missing at random.[126]

Resource use

Resource use over the follow-up period is summarised in *Table 13*. Participants who enrolled in the RO DBT group attended an average of 22.8 individual RO DBT sessions (median 26 sessions) and 19.3 group RO DBT sessions (median 22.5 sessions). In addition, the participants attended an average of 3.4 sessions of other types of talking therapy. Participants in the TAU group attended an average of 9.1 sessions of various talking therapies. The use of all other health and social care services, including medications, tended to be lower in the RO DBT group or broadly similar between the two groups. Days reported off work were also lower in the RO DBT group than in the TAU group.

TABLE 12 The prognosis for new patients choosing RO DBT vs. TAU

	Change from baseline, percentage of participants		
Time point (months)	Worthwhile benefit	No change	Worse off
7	49.9	45.8	4.3
12	31.8	57.7	10.5
18	18.8	73.7	7.5

Improvement (worthwhile benefit) or worsening (worse off) of symptoms is judged as a 17.5% change from baseline. Note that 'worsening' of symptoms does not refer to an iatrogenic effect of RO DBT but rather reflects that predictions for *new* patients must take into account the natural variation in depressive symptoms.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 13 Resource use at baseline and at 7, 12 and 18 months' follow-up by group

Time period, resource use	Treatment group			
	RO DBT		TAU	
	Mean number (SD)	% used at least once	Mean number (SD)	% used at least once
Baseline	$n = 162$		$n = 88$	
Talking therapy sessions	9.26 (8.34)	28.40	10.89 (8.32)	43.18
Talking therapy groups	0.14 (0.04)	1.85	0.22 (0.01)	2.27
Hospital services	5.21 (19.01)	72.84	7.69 (21.65)	64.77
Community services	16.80 (18.63)	95.68	21.20 (62.34)	97.73
Antidepressants	–	86.36	–	79.63
Days off work	40.29 (61.52)	37.65	47.82 (64.44)	46.59
Baseline to 7 months' follow-up	$n = 118$		$n = 61$	
RO DBT individual	22.76 (8.05)	96.23	0.00 (0.00)	0.00
RO DBT groups	19.29 (8.25)	96.23	0.00 (0.00)	0.00
Talking therapy sessions	0.39 (0.20)	5.93	9.39 (6.84)	45.90
Talking therapy groups	0.00 (0.00)	0.00	0.00 (0.00)	0.00
Hospital services	1.43 (2.09)	55.08	4.58 (14.42)	63.93
Community services	10.01 (18.86)	69.14	9.51 (17.81)	63.64
Antidepressants	–	77.12	–	80.33
Days off work	27.86 (49.88)	32.20	38.34 (61.27)	39.34
7–12 months' follow-up	$n = 125$		$n = 61$	
Talking therapy sessions	0.72 (2.51)	4.80	8.33 (5.76)	44.3
Talking therapy groups	0.00 (0.00)	0.00	0.26 (0.18)	1.64
Hospital services	2.59 (7.79)	52.8	1.87 (3.14)	57.38
Community services	9.05 (21.80)	72.22	8.65 (20.05)	61.36
Antidepressants	–	81.60	–	95.08
Days off work	38.39 (68.50)	33.60	41.98 (65.60)	45.90
12–18 months' follow-up	$n = 101$		$n = 51$	
Talking therapy sessions	8.64 (10.36)	21.78	9.23 (6.04)	25.49
Talking therapy groups	0.23 (0.06)	4.95	0.31 (0.11)	3.92
Hospital services	2.37 (4.91)	55.45	2.47 (9.10)	41.18
Community services	6.84 (17.07)	58.28	5.45 (10.74)	51.72
Antidepressants	–	79.21	–	88.24
Days off work	35.44 (55.65)	39.60	53.37 (66.65)	49.02

Costs

Table 14 summarises health and social care costs from the NHS and personal social services perspective over the 6 months prior to trial entry and over the 12- and 18-month follow-up periods. Disaggregated costs are based on complete-case data, as data imputation was conducted at the aggregate level. The costs of all health and social care services used over the 6 months prior to trial entry were lower in the RO DBT group than in the TAU, group, but the difference was not significant (mean cost difference –£1029, bootstrapped 95% CI –£2465 to £407; $p = 0.160$). Excluding the cost of RO DBT, the cost of all health and social care services used was significantly lower over the 12-month follow-up period (adjusted mean cost difference –£909, bootstrapped 95% CI –£1799 to –£19; $p = 0.045$) and was also lower at the 18-month follow-up period, although not significantly so (adjusted mean cost difference of –£901, bootstrapped 95% CI –£2755 to £952; $p = 0.340$). The RO DBT intervention cost per person was approximately £5000, including the cost of both individual and group sessions, resulting in total costs that were significantly higher in the RO DBT group than TAU over both the 12-month follow-up period (adjusted mean cost difference £4566, bootstrapped 95% CI £3691 to £5440; $p < 0.001$) and the 18-month follow-up period (adjusted mean cost difference £4463, bootstrapped 95% CI £2915 to £6011; $p < 0.001$).

TABLE 14 Disaggregated mean costs by group at baseline and each follow-up

Time point, health and social care costs	Treatment group, mean cost (SD)		Mean cost difference[a]	95% CI[a]	*p*-value[a]
	RO DBT	TAU			
Baseline	$n = 162$	$n = 88$			
Talking therapy	£364 (£795)	£692 (£1112)			
Hospital services	£1305 (£4280)	£1968 (£5140)			
Community services	£713 (£733)	£755 (£852)			
Medications	£17 (£12)	£19 (£11)			
Total NHS/PSS costs	£2397 (£4418)	£3426 (£5812)	–£1029	–£2463 to £407	0.160
Absenteeism	£3821 (£5563)	£4648 (£6277)			
Presenteeism	£3382 (£1749)	£2676 (£2771)			
Total societal costs	£9600 (£6469)	£10,750 (£8354)	–£1830	–£3763 to £104	0.064
Baseline to month 7	$n = 118$	$n = 61$			
RO DBT individual	£3095 (£1095)	£0 (£0)			
RO DBT groups	£1910 (£817)	£0 (£0)			
Total RO DBT	£5005 (£1809)	£0 (£0)			
Talking therapy	£217 (£182)	£762 (£1427)			
Hospital services	£379 (£510)	£1019 (£3120)			
Community services	£554 (£848)	£641 (£1087)			
Medications	£15 (£11)	£17 (£11)			
Total NHS/PSS costs	£6170 (£1090)	£2439 (£4935)	£2187	£1420 to £2955	< 0.001
Absenteeism	£1653 (£3067)	£2484 (£5452)			
Presenteeism	£2572 (£2275)	£2498 (£2080)			
Total societal costs	£10,377 (£3754)	£7421 (£7639)	£2255	£1191 to £3320	< 0.001

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 14 Disaggregated mean costs by group at baseline and each follow-up *(continued)*

Time point, health and social care costs	Treatment group, mean cost (SD)		Mean cost difference[a]	95% CI[a]	*p*-value[a]
	RO DBT	TAU			
Baseline to month 12	*n* = 125	*n* = 61			
Talking therapy	£256 (£137)	£1317 (£2388)			
Hospital services	£934 (£1803)	£1216 (£2400)			
Community services	£986 (£1527)	£966 (£1411)			
Medications	£29 (£18)	£35 (£20)			
Total NHS/PSS costs	£7210 (£3343)	£3534 (£4240)	£4566	£3691 to £5440	< 0.001
Absenteeism	£2415 (£5248)	£4063 (£9145)			
Presenteeism	£5641 (£2662)	£4827 (£2419)			
Total societal costs	£15,266 (£5072)	£12,424 (£6764)	£2657	£1217 to £4098	< 0.001
Baseline to month 18	*n* = 101	*n* = 51			
Talking therapy	£501 (£253)	£1633 (£1561)			
Hospital services	£1419 (£1824)	£2004 (£1793)			
Community services	£1419 (£1824)	£1407 (£2399)			
Medications	£45 (£46)	£55 (£48)			
Total NHS/PSS costs	£8389 (£4357)	£5099 (£7677)	£4463	£2915 to £6011	< 0.001
Absenteeism	£4050 (£9616)	£5833 (£13,224)			
Presenteeism	£7634 (£6925)	£7718 (£7818)			
Total societal costs	£20,073 (£6967)	£18,650 (£12,262)	£1062	–£1762 to £3886	0.461

PSS, personal social services.

[a] All analyses adjusted and bootstrapped.

Outcomes

The mean EQ-5D-3L scores and QALYs are presented in *Table 15*, together with unadjusted differences and differences after imputation for missing data. The EQ-5D-3L scores improved in both groups over the 18 months' follow-up. Over the 7- and 12-month follow-up periods, the EQ-5D-3L scores and QALYs were slightly higher in the RO DBT group than in the TAU group, but at 18 months' follow-up the scores were slightly higher in the TAU group. Differences were small and non-significant at all follow-up points.

Cost-effectiveness

The ICERs, based on the bootstrapped adjusted differences in imputed costs and QALYs with missing data imputed, are as follows:

- at 7 months, the ICER was £707,600 (difference in costs, £7076; difference in QALYs, 0.010)
- at 12 months, the ICER was £220,250 (difference in costs, £7048; difference in QALYs, 0.032)
- at 18 months, the ICER was £320,000 (difference in costs, £7360; difference in QALYs, 0.023).

The base-case 12-month additional cost of £7048 was associated with a difference of 0.032 QALYs, yielding an ICER of £220,250 per QALY; this ICER is well above the NICE upper willingness-to-pay threshold of £30,000 per QALY. ICERs were well above the maximum willingness-to-pay threshold at all time points.

TABLE 15 The EQ-5D-3L scores and QALYs over the 18-month follow-up period

	Treatment group		Difference	
Outcome, time point	RO DBT	TAU	Unadjusted	Adjusted and imputed
EQ-5D-3L score				
Baseline	0.422 (0.291)	0.395 (0.329)	0.027	–
7 months	0.549 (0.318)	0.455 (0.337)	0.094	0.058
12 months	0.552 (0.339)	0.547 (0.307)	0.005	0.008
18 months	0.564 (0.311)	0.596 (0.309)	–0.032	0.005
QALYs				
7 months	0.339 (0.187)	0.334 (0.241)	0.005	0.010
12 months	0.534 (0.315)	0.496 (0.349)	0.038	0.032
18 months	0.702 (0.067)	0.763 (0.453)	–0.061	0.023

The cost-effectiveness plane for RO DBT versus TAU from the perspective of NHS and personal social services at the primary 12-month end point is shown in *Figure 5*. This illustrates the scatterplot of 30,000 bootstrapped cost and effectiveness pairs for RO DBT versus TAU. All scatter points lie above the *x*-axis, illustrating that total health and social care costs are higher for the RO DBT group than the TAU group in all cases. The majority of scatter points fall in the north-eastern quadrant in which RO DBT is more effective than TAU but also more costly. Uncertainty around the ICER was explored in a cost-effectiveness acceptability curve, which is shown in *Figure 6*. This graph illustrates that RO DBT in its current format has zero probability of being cost-effective compared with TAU at the NICE willingness-to-pay threshold of £30,000.

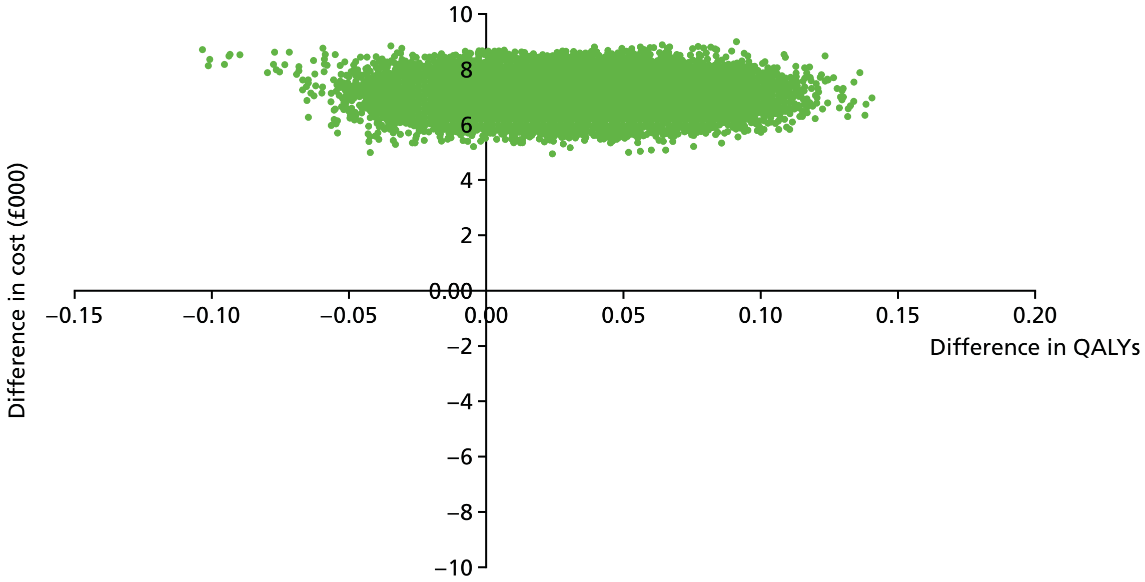

FIGURE 5 A scatterplot of differences in costs vs. differences in QALYs for RO DBT vs. TAU at 12 months' follow-up, taking the NHS/personal social services perspective.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

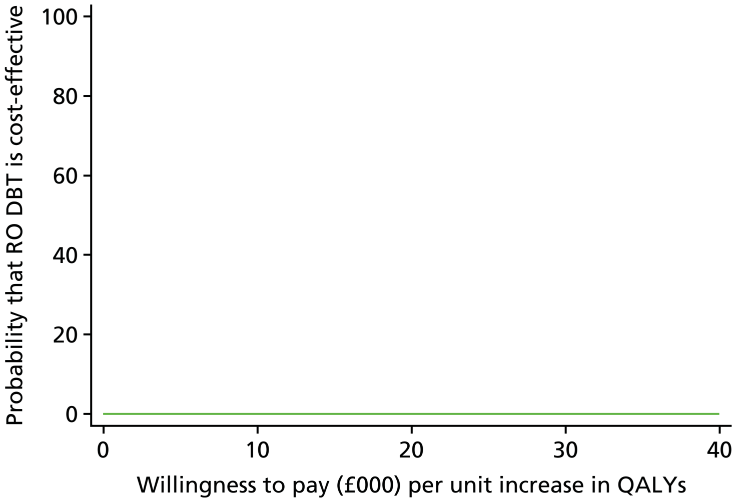

FIGURE 6 The cost-effectiveness acceptability curve for QALYs showing the probability that RO DBT is cost-effective compared with TAU at 12 months' follow-up, taking the NHS/personal social services perspective.

Sensitivity analyses

The primary economic analysis (also known as the 'base case') was based on a cost–utility analysis conducted at the 12-month follow-up point, from the perspective of NHS and personal social services with imputation of missing data. *Table 16* summarises how variation in methods and assumptions affected the ICERs. In sensitivity analyses 1–4 (complete-case analysis, inclusion of productivity losses, reduction in the cost of RO DBT group sessions and analysis at the 18-month follow-up point), with all other factors being equal to the base case, RO DBT remained cost-ineffective compared with TAU, with costs per QALY all well above the NICE willingness-to-pay threshold of £30,000 per QALY. At 18 months' follow-up (scenario 4), the TAU group achieved more QALYs at a lower cost, thereby dominating RO DBT.

In the final scenario, a cost-effectiveness analysis was undertaken using the primary clinical outcome as the measure of effect (the HRSD score), with all other factors being equal to the base-case analysis (the 12-month follow-up point), from the perspective of the NHS and personal social services, with imputation of missing data. This analysis yielded an ICER of £7048 per unit improvement on the HRSD. *Figure 19* in *Appendix 4* illustrates the scatterplot of 30,000 bootstrapped cost and effectiveness pairs for RO DBT versus TAU based on differences in HRSD score using the £1000 per HRSD unit threshold proposed by Romeo *et al.*[153] The results were very similar to the cost–utility analysis, with all replications falling above the threshold of willingness to pay and the cost-effectiveness acceptability curve indicating a zero probability that RO DBT was cost-effective compared with TAU (*Figure 20*, *Appendix 4*).

TABLE 16 Summary of sensitivity analyses

	Difference in		
Scenario	**Costs**	**Outcomes (QALYs)**	**ICER**
Base case	£7048	0.032	£220,250 per QALY
Complete-case analysis	£4566	0.013	£351,231 per QALY
Including productivity losses	£6805	0.030	£224,526 per QALY
RO DBT group costs reduced	£5381	0.026	£206,961 per QALY
Analysis at 18 months	£7360	0.023	£320,000 per QALY
HRSD as a measure of effect	£7048	1.71 points	£4122 per point improvement in HRSD

Mechanisms analyses using instrumental variables

Single-mediator model

The results obtained using the single-mediator model for each mediator variable in turn are shown in *Table 17*. Results were unadjusted for baseline predictors and adjusted for the same baseline predictors (as used elsewhere): the HRSD score prior to randomisation, site, age at depression onset and PD status.

Focusing on those results obtained without using the IVs (i.e. assuming there is no unobserved confounding), it can be seen that the indirect effects mediated by skills (DBT-WCCL) and alliance (California Psychotherapy Alliance Scales) are significant in the sense of having 95% CIs that exclude zero. However, the width of these CIs is wide, and it is clear that the model struggled to accurately decompose the total effect using the information available from this small data set.

Moving to the results obtained using the IVs (i.e. assuming that there is unobserved confounding), it can be seen that this situation deteriorates further. The CIs of the various effects have increased in size compared with those obtained without using the IVs. Although all the decomposed effects, bar those for exposure, are significantly different from zero, the estimates are too large to be plausible and the precision is poor. This could be because of collinearity between the predicted value of exposure and the direct effect.

The reason for this is not because the average direct and indirect effects are non-identified, but because these effects are poorly estimated from the sample data. Running the weak-IV diagnostics using the *ivregress* post-estimation command, the three IVs for the second part of the mediator in *Equation 2* give Cragg–Donald statistics no larger than one. The critical values provided by the command tell us that we can expect the bias of the estimator in both cases to be considerably greater than 30% (to expect a bias of < 10%, it would have needed a Cragg–Donald statistic of at least 9.08), and the actual size of the significance test (nominally at the 5% level) is actually much greater than 25% (the critical value for a size < 10% is 22.30), which indicates a high chance of obtaining a false positive.

TABLE 17 Estimates of the total, direct and indirect effects using single-mediator models

	Effects (95% CI)		
Mediator variable	**Total**	**Direct**	**Indirect**
Exposure (*X*)	*n* = 183		
No unobserved confounding	–3.83 (–6.2 to –1.5)	–0.69 (–6.9 to 7.1)	–3.14 (–12 to 2.8)
Using IVs	–3.83 (–5.5 to –1.7)	14.54 (–26 to 194)	–18.37 (–199 to 23)
Skills (*S*)	*n* = 158		
No unobserved confounding	–3.58 (–6.3 to –1.1)	11.27 (–7.4 to 15)	–14.85 (–19 to –12)
Using IVs	–3.59 (–6.5 to –0.6)	25.43 (15 to 81)	–29.03 (–84 to –16)
Alliance (*A*)	*n* = 176		
No unobserved confounding	–3.83 (–6.1 to –2.2)	15.59 (3.5 to 28)	–19.42 (–30 to –7.7)
Using IVs	–3.77 (–6.0 to –1.7)	–15.85 (–26 to 57)	–19.61 (–52 to 26)

Note

CIs were obtained using bias-corrected accelerated non-parametric bootstrap to allow for interval asymmetry as a result of residual non-normality. All results are adjusted for the following baseline predictors: site, PD status, age at depression onset and HRSD score. The IV estimates of the indirect and direct effects are subject to severe weak-instrument bias and so have no meaningful interpretation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

All-mediator model

The results (unadjusted and then adjusted) for the joint model are shown in *Table 18*. Note that the total mediated effect through exposure does not equal the indirect effect through exposure in *Table 17*. This is because the sample size for the all-mediator model is smaller than that for the single-mediator model for exposure.

If the IVs are not used and the assumption of no unobserved confounding is reasonable, these results suggest that the pathway from RO DBT through exposure to skills has a protective effect on patients' mental health. However, the results obtained using IVs are again implausible and inaccurate. This demonstrates, again, that the IVs are too weak for us to be able to decompose the pathway effects as had been hoped.

Participant perspectives

Eight months after randomisation participants were asked to complete an anonymous survey asking about their experiences of the study and the treatment if they received RO DBT and for suggestions for improvements if they had any. We received 126 (50%) completed forms, 81 (64%) of which were from participants who had attended RO DBT. No formal analysis was conducted, but several themes emerged from the responses and the most common ones are reported in the following sections. For each theme, we present up to three comments from patients who best represent the issue.

Radically open dialectical behaviour therapy was considered helpful

The vast majority of patients who attended RO DBT found it helpful, with only 5 out of 81 RO DBT participants stating that they did not think it was helpful or that it made their depression worse:

> *I find it hard to imagine what life would have been like without this intervention. The group sessions and the skills file have been so useful. I refer to the work every day. I have enjoyed embracing the skills/learning. The CD [compact disc]/loving kindness and other structures have helped shape my problem solving. Through RO DBT I have learnt to cope with myself without harsh judgement. I think I now like myself more and more.*
>
> Patient 64

> *My experience of it has exceeded my expectations. It has allowed me to eradicate a way of thinking and behaving that was detrimental to my mental health. Challenging at times but as it progressed it all made sense and became intensely rewarding.*
>
> Patient 81

> *I don't feel it has done much; it has just made me more aware of what makes my life stressful. A lot of time and effort went into this for no reward or change.*
>
> Patient 120

TABLE 18 Estimates of the direct and indirect effects mediated through exposure using the all-mediators *Equations 4–7*

		Effect		
			Indirect	
Approach	Total mediated ($n = 150$)	Direct	Alliance	Skills
No unobserved confounding	–5.59 (–21 to 5.9)	2.10 (–10.2 to 12.6)	–1.24 (–7.1 to 1.1)	–6.46 (–12.8 to –2.04)
Using IVs	–36.8 (–4143 to 49)	–62.8 (–118,730 to 20)	41.6 (–55 to 7829)	–15.5 (–1089 to 1017)

Note

CIs were obtained using bias-corrected accelerated non-parametric bootstrap to allow for interval asymmetry as a result of residual non-normality. All results are adjusted for the following baseline predictors: site, PD status, age at depression onset and HRSD score. The IV estimates of the indirect and direct effects are subject to severe weak-instrument bias and so have no meaningful interpretation.

Suggestion 1: increase treatment duration or provide follow-up sessions

The most common suggestions for improvement were to increase the number of treatment sessions, allow participants to go through the skills training twice or offer several follow-up sessions at the end of treatment:

> *I received therapy for 6 months and felt that was the ideal length for me. For some people this may be too short and would benefit from a longer time. It would be ideal to have gone back after 4–6 weeks to see a therapist to see how you are getting on. Be given something (support material) at the end when you left.*
>
> Patient 7

> *Very positive and helpful but would have been better if it was longer as many things came to the surface near the end of the therapy which I didn't have a chance to deal with.*
>
> Patient 122

> *Therapy could be improved by giving patients who have completed a check-up session every 3 to 4 months with a therapist for a period of up to 1 year. When the final session took place it felt quite final. A follow-up would certainly be of some reassurance.*
>
> Patient 117

Suggestion 2: increase skills group duration or change format slightly

Although two people found that the duration of the skills groups was too long, at least nine people felt that the sessions were too short and another five felt that less time should be spent on going through individual homework and instead focus on the new skills to be learned:

> *I felt as though some of the group sessions were rushed. There is a lot to take in and it sometimes feels that some parts of the course needed a little more time to study and digest.*
>
> Patient 26

> *Less time could be spent going over the homework each week in the group therapy so there is more time to spend on learning and understanding the skills. Our skills were often taught in a rushed way because we had run out of time in the group session.*
>
> Patient 27

Suggestion 3: simplify language and homework

Although the vast majority reported that the homework was very beneficial to them, 7 out of 81 RO DBT participants said that they found the homework difficult or that the terms used were hard to understand:

> *I think that the jargon and mental health vocabulary could be improved. The pressure to name things in a certain way I feel is not necessary to get the best out of the therapy.*
>
> Patient 99

> *I found the homework worrying i.e. had to make sure I did it. I did not always understand what was being put across which worried me. The 'terms' were very off putting. I felt I had to keep asking for explanations. I probably inhibited some helpful features of the experience.*
>
> Patient 2

Suggestion 4: develop application or other supporting materials

At least seven people suggested that additional supporting materials, such as mobile phone applications (apps), a forum or CDs, should be offered to assist with learning skills:

> *The creation of a phone app so I have to hand, on the go, a reminder of the skills.*
>
> Patient 16

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

> *An app or apps so that skills could be used as reminders/aid on mobiles. Or PDFs [portable document format files] for some purpose.*
>
> Patient 86

> *Simplified version. CD to listen to as sometimes sitting down to read something is hard.*
>
> Patient 111

Open group structure

Finally, four people commented on the open group structure, saying that they would have preferred it to be a closed group with participants all starting at the same time:

> *Joining the group at different times was difficult. I think all should attend + start at the same time.*
>
> Patient 14

> *The only other thing I found hard was people starting therapy at different times. It's nerve racking when everyone seems to know about skills not covered already and at the other end it was really difficult to watch friends leave.*
>
> Patient 40

Chapter 4 Discussion

Principal findings

The primary aim of this study was to examine the efficacy and mechanisms of RO DBT for refractory depression. The primary outcome was HRSD scores at 12 months. Although RO DBT showed an advantage over TAU at 12 months, with a 'medium' effect size with a *d* of 0.41, the difference between groups was not statistically significant. However, the trial did find a statistically significant advantage for RO DBT in reducing depressive symptoms compared with TAU at 7 months after randomisation (immediately after treatment), equivalent to a 'very large' effect size (i.e. standardised mean difference), with a *d* of 1.03. Symptom severity continued to decrease in both groups, yielding a 'small' effect size, with a *d* of 0.32 for the comparison between groups at 18 months. Depressive symptoms in RO DBT participants decreased faster than in control participants, as in a trial of RO DBT-E.[48] More RO DBT participants experienced full or partial remission and change that was deemed 'worthwhile' than TAU participants, from the patient perspective.

The expectation that many participants would have comorbid PDs was justified: 86% of participants reported at least one comorbid Axis-I disorder (anxiety problems in particular) and 78% reported at least one comorbid overcontrolled PD. Secondary analyses indicated that those participants who received RO DBT achieved significant improvements in psychological flexibility, emotional approach coping and expressiveness than those who received TAU, and these differences between groups increased over time. However, economic analyses revealed that RO DBT was not cost-effective on the basis of NICE criteria, despite costing no more than standard dialectical behaviour therapy.[72]

Depressive symptoms

The finding that depressive symptoms at the end of treatment were significantly fewer after RO DBT than after TAU reinforces findings from another trial of RO DBT-E, which reported effect sizes with a *d* of 0.49 and 0.34 after the end of group sessions and individual treatments, respectively.[49] That a very large effect was found at the end of treatment in this trial is especially important, as the TAU participants were free to access any available treatment, including psychotherapy. Furthermore, the effect size is much larger than those reported by previous trials of psychotherapy for refractory depression, which were either small (CBASP)[9] or medium sized (CBT, the Re-ChORD programme).[26,28]

More than half of the TAU participants received some form of psychotherapy and were not solely reliant on ADM. *Table 13* shows that between 7 and 12 months, 44% of TAU participants accessed 'talking therapy' at least once, in comparison with < 5% of RO DBT participants. In contrast, control groups received only ADM in most other studies, including both RO DBT-E pilot studies[48,49] and several CBASP studies.[21–23] A recent meta-analysis suggested that TAU in the UK 'may be considered to be of a high standard encompassing pharmacological and psychological treatment' (p. 158).[154] The subgroup analyses of Watts *et al.*[154] comparing the effects of CBT with different types of TAU (collaboration between primary and secondary care, multiple service providers, psychotherapy alone, minimal contact and others that were not well described) revealed that the effect size of CBT relative to TAU depended on the nature of that TAU. The smallest differences were for comparisons between CBT and primary–secondary care collaboration and the largest were for comparison with 'minimal contact'. Surprisingly, there was a larger difference between CBT and TAU including psychotherapy than between CBT and TAU managed in primary care. One possible explanation is that primary care-managed TAU often includes combined pharmacological and psychological therapy, which is close to current best practice in the UK.[154]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Personality matters

Radically open dialectical behaviour therapy is the first treatment specifically to target maladaptive personality traits in the treatment of refractory depression related to overcontrol. This study suggests that, when psychopathology is long-standing and does not respond to efficacious first-line treatments, overlearned patterns of coping – either overcontrolled or undercontrolled – are interfering with attempts to change. Importantly, RO DBT does not seek to rid clients of their overcontrolled personality style, instead it focuses on changing maladaptive coping linked to overcontrol, such as rigid rule-governed behaviour, inhibited emotional expression and aloof or distant relationships, which are hypothesised to exacerbate emotional loneliness and depression.

Surprisingly, this study does not suggest that the effect of RO DBT depends on the presence of comorbid PD. As 78% of the sample showed at least one overcontrolled PD, it is possible that this reflects the homogeneity of the sample. A meta-analysis of 34 studies concluded that the occurrence of a PD in people with depression is about twice as likely to be associated with a poor response than no PD.[155] This review also found that this result did not differ across treatment types such as electroconvulsive therapy, drugs, psychotherapy or drugs and psychotherapy. A later study by Fournier *et al.*[156] compared the response rates of depressed patients with and without PD who were randomised to either ADM or cognitive therapy. People with PD responded better to ADM (44%) than to cognitive therapy (66%), whereas those without PD responded better to cognitive therapy (70%) than ADM (49%). Finally, a recent large study with 1249 participants within the UK IAPT initiative showed that higher scores on a screening measure for PD were associated with poorer outcomes on measures of depression, anxiety and social functioning and reduced recovery rates at the end of treatment.[157]

Remission rates

As a secondary outcome, we compared remission rates between groups at each time point. Using the primary criteria for full or partial remission, which combined scores on both the HRSD and the LIFE-RIFT, low rates of full remission were found in both groups: at 12 months, 8% of RO DBT participants and no TAU participants achieved full remission. The rates of partial remission were higher, that is, 36% for RO DBT participants and 22% for TAU participants at month 12. For both these criteria and several others, there were higher rates of remission for participants treated by RO DBT than by TAU.

Compared with the two earlier trials of RO DBT-E for refractory depression, full remission rates are much lower in both RO DBT and TAU groups. In Lynch *et al.*'s first study,[49] 71% of RO DBT-E recipients met criteria for remission compared with 47% of controls, and after 6 months the corresponding percentages were 75% and 31%. However, this earlier study may have included potentially less severe patients; for example, the mean age at first depression in the RefraMED trial was much lower, at 23 years versus 36 years. Lynch *et al.*'s second trial[48] recruited patients with PD and comorbid depression and reported remission rates of 71% for RO DBT-E and 50% for control participants. However, the authors defined remission as a HRSD score of ≤ 10 points, whereas the present study used a more conservative method and defined full remission as a HRSD score of ≤ 7 points plus a LIFE-RIFT score of ≤ 12 points. Lynch, the treatment developer, conducted both pilot trials at Duke University, a prestigious academic research centre, led treatment teams and personally provided treatment for study patients. In the present study, in contrast, treatment occurred in non-academic NHS clinics and Lynch did not participate in the delivery of treatment or directly lead clinical teams.

Other psychotherapy trials for refractory depression report a wide range of remission rates, varying between 9.4% and 57%. However, remission rates are not helpful in comparing studies because there is wide variation in the definitions of full and partial remission, the methods used to assess that participants remitted, the duration of follow-up and the study samples. For example, the rates of comorbid Axis-I and Axis-II disorders and the age at depression onset, both of which are factors that can influence treatment response, vary by study. Furthermore, because remission rates dichotomise continuous scores, these factors may also interact in an unexpected fashion. Therefore, to aid interpretation of the findings, we simulated the likely outcomes for new patients choosing between TAU and RO DBT: this suggests that 32 in every 100 patients would make 'worthwhile' gains after 12 months from choosing RO DBT rather than TAU.

Suicide ideation

Despite the moderate to severe depression scores and high comorbidity in the sample, participants reported medium to low suicidal ideation throughout the trial. The odds ratio of a suicide attempt is particularly high among individuals with a history of three or more mental health disorders (odds ratio 19.7)[158] and refractory depression is one of the strongest predictors of suicidal acts.[159] As 44% of the participants had three or more comorbid Axis-I disorders, the number of suicide attempts (7) in the sample is remarkably low.

Mechanisms of change

Instrumental variables analysis

The aims of this analysis were to divide the effect of RO DBT between the pathways through which it is hypothesised to work and to estimate the contribution of each pathway by adjusting for confounding variables and IVs.

We found that, if there was no unobserved confounding, single-mediator models would suggest that the pathway through skills or through alliance had a positive effect on participants' HRSD scores at 7 months. The all-mediators model suggests that skills played the main role, although further exploration of measures of alliance taken early in treatment may aid understanding of the inter-relationships between these variables.

If it is assumed that there is unobserved confounding then the IVs must be incorporated into the analysis. However, no conclusions can be drawn from the IV analysis because the results are severely affected by 'weak' IV bias and so are highly misleading. Although effects could have been identified in a larger study or if the IVs had been far more strongly correlated with the mediator variable, the results highlight the challenge of applying IV methods in this context rather than any specific issue with the design. In particular, we identify a tension between the desire to introduce manipulations that are effective in altering levels of mediators of interest and the desire of clinicians and treatment developers to provide the best possible care and to test new interventions at full dose. Future studies may need to be bolder and, rather than randomising to additional interventions, which have only weak effects on mediators, subdivide the intervention itself. This would randomise participants between variants of the treatment with more or less exposure to core elements and provide greater certainty that these IVs would be strong enough to mitigate bias from unobserved confounding.

In summary, using the IVs incorporated into the design of the study resulted in statistically significant but implausible and imprecise estimates. Standard diagnostics for IVs indicated that these IVs were weak and that the resulting estimates are likely to be biased and unreliable. Thus, there is, as yet, no proof of principle for IV mediation analysis.

Psychological flexibility

One of the proposed mechanisms of RO DBT is that a stable sense of emotional well-being engenders openness, flexibility and social connectedness. Thus, RO DBT aims to help rigid and rule-governed patients learn psychological flexibility. Indeed, the RO DBT skill most frequently emphasised in training is 'flexible mind'. The results show that those participants who received RO DBT had significantly better psychological flexibility (AAQ-II) scores than those in TAU and this effect was present at all three follow-up time points, indicating that those who received RO DBT continued to experience the benefits of treatment in this area.

Emotional approach coping and expressiveness

Radically open dialectical behaviour therapy targets masked and constrained social signalling as the primary source of emotional loneliness, isolation and misery. Therefore, a core aim of treatment is to reinforce context-appropriate disclosure and uninhibited expression of emotion. The results indicate that those participants who received RO DBT scored higher on emotional approach coping and expressiveness

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

(as measured via the EAC scale) than those who received TAU. The difference between groups was nearly significant at month 7 and increased at months 12 and 18.

Health economics

Radically open dialectical behaviour therapy was not cost-effective compared with TAU alone in the treatment of patients with refractory depression, either from the perspective of the NHS and personal social services or when productivity losses were included. RO DBT is a resource-intensive intervention that does not achieve sufficient gains in outcomes or savings in the use of other health and social services to justify its additional cost. Cost-ineffectiveness was driven by the intensive contact in RO DBT, which may, in part, have been because of the comprehensive care available to patients enrolled in a clinical trial. Whether or not such a costly service could be provided in routine mental health services in the UK NHS is debatable. Therefore, it may be worth exploring the effect on participant outcomes of a reduction in the number of sessions provided, an increase in the number of participants per group or a focus on the group element of the intervention. This requires further research into a briefer version of RO DBT or a stepped-care approach offering group sessions initially and individual sessions only to those who fail to respond. Small studies of RO DBT skills training alone have reported promising results.[160,161]

Nevertheless, RO DBT achieved cost-savings throughout the 18 months' follow-up. The use of other health and social care services tended to be higher in the TAU group and the cost of these other services were significantly higher than in the RO DBT group over the 12 months' follow-up. This leads to the hypothesis that, if these cost-savings continue in the longer term, the cost-savings may eventually outweigh the cost of the intervention and RO DBT may become more cost-effective than TAU. This is a question for future research to consider.

This is the first economic evaluation of RO DBT for depression and one of very few studies to explore the economic implications of a DBT-informed approach for an Axis-I disorder. Previous economic evaluations have focused on standard DBT for PD.[72,162] One economic study, alongside a RCT, compared standard DBT with TAU for self-harming patients with PD.[72] Although the economic component focused on the cost of the intervention and other health and social care services and did not include a formal cost-effectiveness analysis, it is a useful comparator. Excluding the cost of standard DBT, the mean cost of all health and social care services used over 12 months by the PD population (DBT group, about £2500 per patient; control group, about £3400) was very similar to that in the depressed population (RO DBT group about £2200; control group about £3500). The cost of standard DBT was lower, at £3159 per patient, than for RO DBT, at £5005 per patient, in the study, but this is still substantially higher than the cost of therapies such as CBT, which is commonly provided by the UK NHS. The authors concluded that standard DBT can be effective in reducing self-harm in patients with PD, but acknowledged that this would incur higher total treatment costs.[72]

The remaining economic evaluations were part of a systematic review and preliminary economic evaluation of evidence for psychological therapies for BPD.[162] Brazier *et al.*[162] reported four individual economic evaluations combining data from four RCTs[163–166] of standard DBT for BPD, with cost estimations using data from the trials and other published sources. The results were mixed. Two evaluations suggested that standard DBT dominated the comparison arm [TAU in one study (based on Linehan *et al.*[164]), client-centred therapy in the other (based on Turner *et al.*[163])], achieving better outcomes at a lower cost per patient. A third evaluation (based on Van Den Bosch *et al.*[165]) found slightly higher costs but better outcomes (parasuicide events avoided) for standard DBT compared with TAU. The final evaluation (based on Koons *et al.*[166]) found that standard DBT was more costly than TAU but delivered only moderate gains in outcomes (parasuicide events avoided and QALYs), with an ICER considerably above the NICE willingness-to-pay threshold. All of these analyses exhibited considerable uncertainty and limited generalisability to the UK context.[167]

The authors used data from trials with small sample sizes (range 20–44 patients analysed) and relied on assumptions and data from sources outside the original trials, as none provided economic data. The authors themselves concluded that the findings do not support the cost-effectiveness of standard DBT, although they suggest that it has the potential to be cost-effective for BPD.

Strengths and limitations

Generalisability of findings

Although the RefraMED trial has much in common with clinical effectiveness studies (e.g. few exclusion criteria), it is best viewed as a hybrid efficacy trial of a promising treatment. For example, centralised control of assessments and use of IVs characterise efficacy trials, whereas cost-effectiveness analyses and exclusion of the treatment developer from providing or supervising therapy characterise effectiveness trials.

A real strength of the study was the inclusion of patients with depression that is difficult to treat, including chronic and recurrent presentations, PDs and high levels of suicidal ideation. This inclusion strategy considerably increases the generalisability of the findings. As mental health practitioners treat numerous patients with multiple mental health needs and PDs, transdiagnostic treatments, such as RO DBT, are needed to meet the demands of clinical practice.

Therapist variation

The random allocation of therapists to participants is both a scientific strength and a practical limitation. Allocating therapists to participants at random (rather than allocating the most difficult participants to the most skilled therapists, as in clinical practice) enhanced the observed performance of the best therapists while reducing that of the least skilled. The study found that 12% of the variability in the primary outcome was attributable to variation between therapists, with a difference of 2.6 points on the HRSD scale (equivalent to an effect size with a *d* of 0.43) between the best and weakest therapists. This substantial variation between therapists has decreased the statistical power of the study. Our power analysis assumed that for less variability in the primary outcome would be attributable to the variation between therapists; combined with more loss to follow-up than expected, this reduced the power of the study for the primary outcome at month 12.

Sample size and design considerations

As most studies in this field use the end of treatment as their primary end point, rather than a predefined time as used in study, it is difficult to compare the primary outcome with other studies that have investigated treatments for refractory depression. For example, the Re-ChORD trial did not collect sufficient follow-up data to investigate post-treatment effects[26] and of the six studies[19,21–25] included in the CBASP meta-analysis,[9] only one study reported follow-up data 12 months after therapy.[22] Similar to the findings 12 months after the end of therapy (i.e. at month 18), the Research Evaluating the Value of Augmenting Medication With Psychotherapy (REVAMP) study[22] reported no significant differences between treatment groups, although both groups showed significant changes over time for depressive symptom scores. The Tavistock Adult Depression Study[27] assessed patients at 6-monthly intervals during 18 months of treatment and 2 years' follow-up. The study reported that the intervention group started to show significantly larger decreases in HRSD scores than those shown by the TAU group at 6 months after therapy, with medium effect sizes. The CoBAlT study[28] showed that BDI scores after 6 and 12 months were significantly lower in the intervention group than in the control group.

Despite the best efforts to achieve a sample size of 276 participants, only 250 were randomised and only 190 were analysable at the primary end point of 12 months. Although the desired effect size was achieved at month 12 (with a *d* of 0.41) and a spectacular effect size at month 7 (with a *d* of 1.03), the shortage of power was disappointing.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Weak instrumental variables

The choice of IVs for the RefraMED trial was necessarily speculative. To our knowledge, the RefraMED trial is the first mental health trial to examine the potential advantage of adding IVs. The results of the mechanisms analysis using IVs are consistent with those obtained using 'weak' IVs, which lead to biased and misleading conclusions. Nevertheless, the work has highlighted an important consideration for future researchers in this area: stronger IVs are needed if studies wish to subdivide the direct and indirect effects of treatment.

Patient and public involvement

Recruitment and membership

The Mental Health Research Network and INVOLVE helped us recruit service users: two for the TSC and two for the TMG. Of these four service users, one attended all meetings and two attended one or two meetings, but we lost contact with one, who did not attend any meetings. These service users actively engaged with the development of patient information leaflets by providing detailed feedback on the readability and clarity, the management of the study by attending trial management meetings and the dissemination of the study by providing text for and feedback on the report, in particular the plain English summary and *Chapter 2, Patient and public involvement*.

Effect of patient and public involvement

First, the Mental Health Research Network provided anonymous feedback on the wording of the patient information leaflets, thus improving the intelligibility and accessibility of the materials. Second, input from the service users who were members of the TSC and, in particular, the TMG, informed the recruitment and dissemination strategies; this improved how the study communicated about the treatment and the findings. Third, during the trial participants were asked for feedback about the treatment; this will inform the further development of RO DBT training and delivery.

Chapter 5 Conclusions

Implications for health care

This study has demonstrated that depressive symptoms of participants receiving RO DBT decreased significantly over the 7 months of treatment and these reductions in symptoms were maintained at 12 and 18 months. That differences between RO DBT and TAU groups after 12 and 18 months were not statistically significant was, in part, a result of a smaller analysable sample than had been sought and the resulting lack of power to detect a real difference between these groups and, in part, as a result of improvement in depressive symptoms after 7 months in the TAU group, who received more treatment between months 7 and 12 than the RO DBT group. Despite the higher cost of RO DBT than treatments typically offered within the NHS, it can improve depressive symptoms in a highly symptomatic population who suffer from many mental health problems and PDs and may require a specialised treatment for their problems.

Radically open dialectical behaviour therapy is the first treatment specifically to target maladaptive personality traits in the treatment of refractory depression related to overcontrol. We suggest that when psychopathology is long-standing and does not respond to efficacious first-line treatments, this may signal that overlearned patterns of coping are interfering with attempts at change.

The patient group that was addressed by this study has a high level of distress and functional impairment that affects themselves, their employers, their carers and those they care for, especially their children. An effective intervention will reduce suffering in patients and those around them with potential effects in preventing loss of productivity in the economy and the development of mental health problems in new generations. The less stringent inclusion criteria adopted by the RefraMED trial reflects clinical practice in the real world and maximises the representativeness of the data set.

Radically open dialectical behaviour therapy is the first transdiagnostic treatment designed specifically for overcontrolled disorders, including refractory depression, anorexia nervosa and obsessive–compulsive PD, which are much more common than undercontrolled disorders such as BPD and bipolar disorder. Despite the higher prevalence of overcontrolled disorders, to date most treatments have focused on undercontrolled disorders, aiming to improve emotion regulation skills, reduce impulsivity or achieve a state of equanimity. In contrast, RO DBT introduces the notion that what matters more for long-term psychological well-being is being socially connected rather than feeling internally regulated or achieving a state of equanimity.

Radically open dialectical behaviour therapy offers a new way of thinking about psychological well-being and targets a group of problem behaviours that have, thus far, been under-recognised and undertreated. Novel approaches designed to enhance social connectedness include skills designed to activate a neurologically based social safety engagement system, signal co-operation, encourage genuine self-disclosure, break down overlearned expressive inhibitory barriers, enhance forgiveness and change unhelpful envy or bitterness. The emphasis on social signalling and activation of neural substrates differentiates RO DBT from other treatments that target thoughts and behaviours, such as CBT. RO DBT suggests a unique mechanism of therapeutic change by linking neuroregulatory theory and the communicative functions of emotional expression to the formation of close social bonds that translate into novel skills targeting social signalling and changing psychophysiological arousal – a key component differentiating RO DBT from other treatments. Although more work is needed to further investigate its clinical effectiveness for refractory depression and other disorders, the findings demonstrate that RO DBT has great potential to contribute to improved well-being for those suffering from overcontrolled disorders.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Recommendations for future research

Improving cost-effectiveness

This study concluded that RO DBT in its current form is not cost-effective on the basis of NICE criteria. Hence, research should address how RO DBT can be refined to achieve the same gains but at a lower cost and for a longer period of time. There are many ways in which the delivery of RO DBT could be varied to reduce costs and it is important that future work addresses the clinical effectiveness and cost-effectiveness of these amended treatment protocols. For example, treatment might be tapered from weekly to fortnightly or monthly attendance after an initial intensive period. Another important avenue of future research will be the development and evaluation of a mobile- or web-based RO DBT support programme for patients. Several participants in this study indicated that this would support them both during the active treatment phase and afterwards.

Furthermore, RO DBT skills groups alone – without individual treatment sessions – may achieve important clinical outcomes. A recent non-randomised trial of overcontrolled personality dysfunction in treatment-resistant adults[161] found that those who had learned RO DBT skills in 18 3-hour skill classes over 9 weeks showed significantly greater improvements in severity of psychological symptoms (large effect size), social safeness (medium effect size) and effective use of coping skills (large effect size) than those receiving TAU while on the waiting list for RO DBT. More thorough investigation of the most effective RO DBT skills, including exploratory analyses of the RefraMED data set, will be critical in developing a shorter version of the skills programme for use in NHS settings with limited funding.

Skill consolidation after end of treatment

It is also important to ensure that patients who have completed RO DBT continue to benefit from the skills that they have learned. One possible way of achieving this would be to offer RO DBT maintenance treatment. Feedback from the participants shows that they would welcome a maintenance programme (see *Chapter 3*, *Patient perspectives*). These 'graduate groups' could be led by clinicians or peers, as a recent meta-analytic review reported that peer-led skills groups show promise in reducing depressive symptoms.[168] The Dorset research site is currently developing a peer-led approach. Graduate programmes designed to consolidate skills learned in therapist-led group sessions have been trialled: a recent pilot demonstrated that patients with predominantly depressive symptoms who attended a weekly skills group for 9 months experienced significantly improved mood and a greater number of patients were able to return to work.[169] The authors reported a gradual reduction in attendance with patients, initially with weekly participation but reducing their attendance thereafter.

Longer follow-up periods

This was the first multisite trial of RO DBT for refractory depression and the largest to date. It demonstrated that RO DBT is much better than TAU at improving depressive symptoms over the 7 months of active treatment and that the benefits of RO DBT are maintained over the 12 months following the end of treatment, possibly owing to maintained skills in patients. However, we need a better understanding of whether or not and how these gains are maintained over a longer period of time, especially because of the recurrent nature of refractory depression and the associated risk of suicide: the recurrence of major depression increases the risk of suicide attempts by seven times in the 2 years after discharge from hospital.[159] It is not yet known whether or not RO DBT and control patients differ in relapse risk long after treatment. Furthermore, although RO DBT is expensive in its current form, the cost of other health and social care services were significantly higher in the control group than in the RO DBT arm. It is possible that if these cost-savings continue over a long time, RO DBT may become more cost-effective than TAU. Future studies using a longer follow-up period would shed light on these issues.

Transdiagnostic studies

Finally, although RO DBT has predominantly been investigated in adults with depression or eating disorders, it is clinically being used much more widely, including for adolescents and populations with anxiety or autism spectrum disorders. Future studies should aim to test the clinical effectiveness of RO DBT in these other overcontrolled populations, when possible using a transdiagnostic approach.

Acknowledgements

For generous and assiduous support of the trial we thank the trial participants, the trial therapists (Michelle Batcock, Laura Brummer, Ellie Campbell, Sara Clarke, Susi Collins, Laura Dannahy, Sarah Farman, Alex Fowke, Michelle Gormley, Stephanie Hastings, Veronica Hicks, Carolyn Hinds, James Lea, Maria Ozanne, Donna Rutherford, Pip Thomas, Frankie Vanderbeek, Sandy Waite and Hannah Wilson), the trial research assistants (Alison Firth, Anna Knight, Helena Maher, Ray Manning, Zoe McAndrews, Meghan Owens, Nadia Peppa, Rebecca Phillips, Katie Rose and Davina Wong), the trial administrative staff (Carol Hill, Sue McDonald, Karlien Paas and Gill Schofield), the clinical studies officers (Eimear Corrigan, Emma Dobson, Julia Marandu, Caroline Mulvaney-Jones, Julia Roberts and Claire Wellsted), several voluntary research assistants, the independent members of the TSC [Mark Williams (chairperson), Francesco Palma, Seren Roberts and Rebecca Walwyn] and the DMEC [Jonathan Evans (chairperson), Gary Latchford and John Norrie (editor for the NIHR Journals Library)], the adherence raters (Christine Dunkley, Amy Gaglia and Stine Laberg), and administrative and research and development staff at Dorset HealthCare University NHS Foundation Trust, Southern Health NHS Foundation Trust and Betsi Cadwaladr University Health Board.

Contributions of authors

Thomas R Lynch (Professor of Clinical Psychology; Chief Investigator) developed RO DBT and contributed to all aspects of the design and management of the study and the writing of this report.

Roelie J Hempel (Senior Research Fellow) contributed to the design of the study, acted as trial manager throughout the study and co-ordinated and contributed to the writing of this report.

Ben Whalley (Lecturer in Psychology) contributed to the design of the study, undertook data analyses and contributed to the writing of the grant proposal and this report.

Sarah Byford (Professor of Health Economics) contributed to the design of the study, was responsible for the design and management of the economic evaluation and contributed to drafting the economic results.

Rampaul Chamba (Patient and Public Representative) contributed to managing the study and writing this report, especially with regard to patient and public engagement and the plain English summary.

Paul Clarke (Professor of Social Statistics) was responsible for the design, analysis and writing of the IV analysis.

Susan Clarke (Visiting Professor, Consultant Clinical Psychologist, Principal Investigator and Clinical Lead for the Dorset site until 2014) contributed to the design of the study and clinical methods and supervision of research assistants and therapists and delivered the therapy.

David Kingdon (Professor of Mental Health Care Delivery, Site Principal Investigator for the Hampshire site) contributed to the design of the study and the supervision of research assistants and to writing this report.

Heather O'Mahen (Senior Lecturer in Clinical Psychology; Assessment Lead) trained assessors in clinical assessment, provided clinical supervision to the research assistants, carried out reliability analyses and contributed to writing this report.

Bob Remington (Emeritus Professor in Psychology) contributed to the writing of the grant proposal and this report and to managing the study.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Sophie C Rushbrook (Consultant Clinical Psychologist, Clinical Lead for the Dorset site from 2014) was responsible for selecting and supervising therapists providing treatment, delivered therapy and contributed to the preparation of results for publication.

James Shearer (Lecturer in Health Economics) undertook economic analyses, prepared the results for publication and contributed to writing this report.

Maggie Stanton (Consultant Clinical Psychologist, Clinical Lead for the Hampshire site) conducted the selection and supervision of the therapists providing treatment, delivered therapy, contributed clinical advice and contributed to the writing of this report.

Michaela Swales (Consultant Clinical Psychologist, Reader in Clinical Psychology, Site Principal Investigator for the North Wales site) was responsible for selecting and supervising therapists providing treatment, delivered therapy and contributed to the supervision of research assistants and the preparation of results for publication.

Alan Watkins (Associate Professor of e-Trials Research) designed and managed the randomisation process, drafted the statistical analysis plan and reviewed its implementation and contributed to validation of study data.

Ian T Russell (Professor of Clinical Trials) was the trial methodologist and contributed to the design and management of the study and to the writing of the grant proposal and this report.

Publications

Lynch TR, Whalley B, Hempel RJ, Byford S, Clarke P, Clarke S, *et al.* Refractory depression – Mechanisms and Evaluation of radically open Dialectical behaviour therapy (RO-DBT) [REFRAMED]: protocol for randomised trial. *BMJ Open* 2015;**5**:e00885.

Hughes-Morley A, Young B, Hempel RJ, Russell IT, Waheed W, Bower P. What can we learn from trial decliners about improving recruitment? Qualitative study. *Trials* 2016;**17**:494.

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to available anonymised data may be granted following review.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation.

References

1. Greer TL, Kurian BT, Trivedi MH. Defining and measuring functional recovery from depression. *CNS Drugs* 2010;**24**:267–84. https://doi.org/10.2165/11530230-000000000-00000
2. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLOS Med* 2006;**3**:e442. https://doi.org/10.1371/journal.pmed.0030442
3. Berlim MT, Turecki G. What is the meaning of treatment resistant/refractory major depression (TRD)? A systematic review of current randomized trials. *Eur Neuropsychopharmacol* 2007;**17**:696–707. https://doi.org/10.1016/j.euroneuro.2007.03.009
4. Souery D, Amsterdam J, de Montigny C, Lecrubier Y, Montgomery S, Lipp O, *et al.* Treatment resistant depression: methodological overview and operational criteria. *Eur Neuropsychopharmacol* 1999;**9**:83–91. https://doi.org/10.1016/S0924-977X(98)00004-2
5. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, *et al.* Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry* 2006;**163**:1905–17. https://doi.org/10.1176/ajp.2006.163.11.1905
6. Thase ME, Friedman ES, Biggs MM, Wisniewski SR, Trivedi MH, Luther JF, *et al.* Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: a STAR*D report. *Am J Psychiatry* 2007;**164**:739–52. https://doi.org/10.1176/ajp.2007.164.5.739
7. Crown WH, Finkelstein S, Berndt ER, Ling D, Poret AW, Rush AJ, Russell JM. The impact of treatment-resistant depression on health care utilization and costs. *J Clin Psychiatry* 2002;**63**:963–71. https://doi.org/10.4088/JCP.v63n1102
8. Fournier JC, DeRubeis RJ, Shelton RC, Hollon SD, Amsterdam JD, Gallop R. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *J Consult Clin Psychol* 2009;**77**:775–87. https://doi.org/10.1037/a0015401
9. Negt P, Brakemeier EL, Michalak J, Winter L, Bleich S, Kahl KG. The treatment of chronic depression with cognitive behavioral analysis system of psychotherapy: a systematic review and meta-analysis of randomized-controlled clinical trials. *Brain Behav* 2016;**6**:e00486. https://doi.org/10.1002/brb3.486
10. Fava M. Diagnosis and definition of treatment-resistant depression. *Biol Psychiatry* 2003;**53**:649–59. https://doi.org/10.1016/S0006-3223(03)00231-2
11. Thomas L, Kessler D, Campbell J, Morrison J, Peters TJ, Williams C, *et al.* Prevalence of treatment-resistant depression in primary care: cross-sectional data. *Br J Gen Pract* 2013;**63**:e852–8. https://doi.org/10.3399/bjgp13X675430
12. Burrows GD, Norman TR. Treatment-Resistant Unipolar Depression. In Lader M, Naber D, editors. *Difficult Clinical Problems in Psychiatry*. London: Martin Dunitz; 1999. pp. 57–73.
13. McCrone PR, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. *Paying the Price: The Cost of Mental Health Care in England to 2026*. London: The King's Fund; 2008. URL: www.kingsfund.org.uk (accessed 2 August 2016).
14. Ivanova JI, Birnbaum HG, Kidolezi Y, Subramanian G, Khan SA, Stensland MD. Direct and indirect costs of employees with treatment-resistant and non-treatment-resistant major depressive disorder. *Curr Med Res Opin* 2010;**26**:2475–84. https://doi.org/10.1185/03007995.2010.517716
15. Cuijpers P, van Straten A, Schuurmans J, van Oppen P, Hollon SD, Andersson G. Psychotherapy for chronic major depression and dysthymia: a meta-analysis. *Clin Psychol Rev* 2010;**30**:51–62. https://doi.org/10.1016/j.cpr.2009.09.003

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

16. Shinohara K, Honyashiki M, Imai H, Hunot V, Caldwell DM, Davies P, *et al.* Behavioural therapies versus other psychological therapies for depression. *Cochrane Database Syst Rev* 2013;**10**:CD008696. https://doi.org/10.1002/14651858.CD008696.pub2
17. McPherson S, Cairns P, Carlyle J, Shapiro DA, Richardson P, Taylor D. The effectiveness of psychological treatments for treatment-resistant depression: a systematic review. *Acta Psychiatr Scand* 2005;**111**:331–40. https://doi.org/10.1111/j.1600-0447.2004.00498.x
18. McCullough Jr JP. *Treatment for Chronic Depression: Cognitive Behavioral Analysis System of Psychotherapy (CBASP)*. New York, NY: The Guildford Press; 2003.
19. Michalak J, Schultze M, Heidenreich T, Schramm E. A randomized controlled trial on the efficacy of mindfulness-based cognitive therapy and a group version of cognitive behavioral analysis system of psychotherapy for chronically depressed patients. *J Consult Clin Psychol* 2015;**83**:951–63. https://doi.org/10.1037/ccp0000042
20. Brakemeier EL, Engel V, Schramm E, Zobel I, Schmidt T, Hautzinger M, *et al.* Feasibility and outcome of cognitive behavioral analysis system of psychotherapy (CBASP) for chronically depressed inpatients: a pilot study. *Psychother Psychosom* 2011;**80**:191–4. https://doi.org/10.1159/000320779
21. Keller MB, McCullough JP, Klein DN, Arnow B, Dunner DL, Gelenberg AJ, *et al.* A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *N Engl J Med* 2000;**342**:1462–70. https://doi.org/10.1056/NEJM200005183422001
22. Kocsis JH, Gelenberg AJ, Rothbaum BO, Klein DN, Trivedi MH, Manber R, *et al.* Cognitive behavioral analysis system of psychotherapy and brief supportive psychotherapy for augmentation of antidepressant nonresponse in chronic depression: the REVAMP Trial. *Arch Gen Psychiatry* 2009;**66**:1178–88. https://doi.org/10.1001/archgenpsychiatry.2009.144
23. Schramm E, Zobel I, Dykierek P, Kech S, Brakemeier EL, Kulz A, *et al.* Cognitive behavioral analysis system of psychotherapy versus interpersonal psychotherapy for early-onset chronic depression: a randomized pilot study. *J Affect Disord* 2011;**129**:109–16. https://doi.org/10.1016/j.jad.2010.08.003
24. Wiersma JE, Van Schaik DJ, Hoogendorn AW, Dekker JJ, Van HL, Schoevers RA, *et al.* The effectiveness of the cognitive behavioral analysis system of psychotherapy for chronic depression: a randomized controlled trial. *Psychother Psychosom* 2014;**83**:263–9. https://doi.org/10.1159/000360795
25. Schramm E, Zobel I, Schoepf D, Fangmeler T, Schnell K, Walter H, *et al.* Cognitive behavioral analysis system of psychotherapy versus escitalopram in chronic major depression. *Psychother Psychosom* 2015;**84**:227–40. https://doi.org/10.1159/000381957
26. Murray G, Michalak EE, Axler A, Yaxley D, Hayashi B, Westrin A, *et al.* Relief of chronic or resistant depression (Re-ChORD): a pragmatic, randomized, open-treatment trial of an integrative program intervention for chronic depression. *J Affect Disord* 2010;**123**:243–8. https://doi.org/10.1016/j.jad.2009.10.015
27. Fonagy P, Rost F, Carlyle JA, McPherson S, Thomas R, Pasco Fearon RM, *et al.* Pragmatic randomized controlled trial of long-term psychoanalytic psychotherapy for treatment-resistant depression: the Tavistock Adult Depression Study (TADS). *World Psychiatry* 2015;**14**:312–21. https://doi.org/10.1002/wps.20267
28. Wiles N, Thomas L, Abel A, Ridgway N, Turner N, Campbell J, *et al.* Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care based patients with treatment resistant depression: results of the CoBalT randomised controlled trial. *Lancet* 2013;**381**:375–84. https://doi.org/10.1016/S0140-6736(12)61552-9

29. Wiles NJ, Thomas L, Turner N, Garfield K, Kounali D, Campbell J, *et al.* Long-term effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: follow-up of the CoBalT randomised controlled trial. *Lancet Psychiatry* 2016;**3**:137–44. https://doi.org/10.1016/S2215-0366(15)00495-2

30. Hollon SD, Ponniah K. A review of empirically supported psychological therapies for mood disorders in adults. *Depress Anxiety* 2010;**27**:891–932. https://doi.org/10.1002/da.20741

31. Crijnen AA, Achenbach TM, Verhulst FC. Comparisons of problems reported by parents of children in 12 cultures: total problems, externalizing, and internalizing. *J Am Acad Child Adolesc Psychiatry* 1997;**36**:1269–77. https://doi.org/10.1097/00004583-199709000-00020

32. Fava M, Farabaugh AH, Sickinger AH, Wright E, Alpert JE, Sonawalla S, *et al.* Personality disorders and depression. *Psychol Med* 2002;**32**:1049–57. https://doi.org/10.1017/S0033291702005780

33. Klein DN, Riso LP, Donaldson SK, Schwartz JE, Anderson RL, Ouimette PC, *et al.* Family study of early-onset dysthymia. Mood and personality disorders in relatives of outpatients with dysthymia and episodic major depression and normal controls. *Arch Gen Psychiatry* 1995;**52**:487–96. https://doi.org/10.1001/archpsyc.1995.03950180073010

34. Riso LP, du Toit PL, Blandino JA, Penna S, Dacey S, Duin JS, *et al.* Cognitive aspects of chronic depression. *J Abnorm Psychol* 2003;**112**:72–80. https://doi.org/10.1037/0021-843X.112.1.72

35. Candrian M, Schwartz F, Farabaugh A, Perlis RH, Ehlert U, Fava M. Personality disorders and perceived stress in major depressive disorder. *Psychiatry Res* 2008;**160**:184–91. https://doi.org/10.1016/j.psychres.2007.06.014

36. Asendorpf JB, Denissen JJ, van Aken MA. Inhibited and aggressive preschool children at 23 years of age: personality and social transitions into adulthood. *Dev Psychol* 2008;**44**:997–1011. https://doi.org/10.1037/0012-1649.44.4.997

37. Chapman BP, Goldberg LR. Replicability and 40-year predictive power of childhood ARC types. *J Pers Soc Psychol* 2011;**101**:593–606. https://doi.org/10.1037/a0024289

38. Eisenberg N, Fabes RA, Guthrie IK, Reiser M. Dispositional emotionality and regulation: their role in predicting quality of social functioning. *J Pers Soc Psychol* 2000;**78**:136–57. https://doi.org/10.1037/0022-3514.78.1.136

39. Anderluh M, Tchanturia K, Rabe-Hesketh S, Collier D, Treasure J. Lifetime course of eating disorders: design and validity testing of a new strategy to define the eating disorders phenotype. *Psychol Med* 2009;**39**:105–14. https://doi.org/10.1017/S0033291708003292

40. Chapman AL, Lynch TR, Rosenthal MZ, Cheavens JS, Smoski MJ, Krishnan KRR. Risk aversion among depressed older adults with obsessive compulsive personality disorder. *Cognit Ther Res* 2007;**31**:161–74. https://doi.org/10.1007/s10608-006-9114-x

41. Lynch TR. *Radically Open Dialectical Behavior Therapy: Theory and Practice for Treating Disorders of Overcontrol*. Reno, NV: Context Press; 2018.

42. Lynch TR, Hempel RJ, Dunkley C. Radically open-dialectical behavior therapy for disorders of over-control: signaling matters. *Am J Psychother* 2015;**69**:141–62. https://doi.org/10.1176/appi.psychotherapy.2015.69.2.141

43. Porges SW. The polyvagal theory: phylogenetic substrates of a social nervous system. *Int J Psychophysiol* 2001;**42**:123–46. https://doi.org/10.1016/S0167-8760(01)00162-3

44. Porges SW. Social Engagement and Attachment: A Phylogenetic Perspective. In King JA, Ferris CF, Lederhendler II, editors. *Roots of Mental Illness in Children*. New York, NY: New York Academy of Sciences; 2003. pp. 31–47. https://doi.org/10.1196/annals.1301.004

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

45. Butler EA, Egloff B, Wilhelm FH, Smith NC, Erickson EA, Gross JJ. The social consequences of expressive suppression. *Emotion* 2003;**3**:48–67. https://doi.org/10.1037/1528-3542.3.1.48

46. Mauss IB, Shallcross AJ, Troy AS, John OP, Ferrer E, Wilhelm FH, *et al.* Don't hide your happiness! Positive emotion dissociation, social connectedness, and psychological functioning. *J Pers Soc Psychol* 2011;**100**:738–48. https://doi.org/10.1037/a0022410

47. Schneider KG, Hempel RJ, Lynch TR. That 'poker face' just might lose you the game! The impact of expressive suppression and mimicry on sensitivity to facial expressions of emotion. *Emotion* 2013;**13**:852–66. https://doi.org/10.1037/a0032847

48. Lynch TR, Cheavens JS, Cukrowicz KC, Thorp SR, Bronner L, Beyer J. Treatment of older adults with co-morbid personality disorder and depression: a dialectical behavior therapy approach. *Int J Geriatr Psychiatry* 2007;**22**:131–43. https://doi.org/10.1002/gps.1703

49. Lynch TR, Morse JQ, Mendelson T, Robins CJ. Dialectical behavior therapy for depressed older adults: a randomized pilot study. *Am J Geriatr Psychiatry* 2003;**11**:33–45. https://doi.org/10.1097/00019442-200301000-00006

50. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;**37**:53–72. https://doi.org/10.1016/0168-8510(96)00822-6

51. National Institute for Health and Clinical Excellence (NICE). *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.

52. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**:1173–82. https://doi.org/10.1037/0022-3514.51.6.1173

53. Smoski MJ, Felder J, Bizzell J, Green SR, Ernst M, Lynch TR, Dichter GS. fMRI of alterations in reward selection, anticipation, and feedback in major depressive disorder. *J Affect Disord* 2009;**118**:69–78. https://doi.org/10.1016/j.jad.2009.01.034

54. Smoski MJ, Lynch TR, Rosenthal MZ, Cheavens JS, Chapman AL, Krishnan RR. Decision-making and risk aversion among depressive adults. *J Behav Ther Exp Psychiatry* 2008;**39**:567–76. https://doi.org/10.1016/j.jbtep.2008.01.004

55. Kasch KL, Rottenberg J, Arnow BA, Gotlib IH. Behavioral activation and inhibition systems and the severity and course of depression. *J Abnorm Psychol* 2002;**111**:589–97. https://doi.org/10.1037/0021-843X.111.4.589

56. Riso LP, Miyatake RK, Thase ME. The search for determinants of chronic depression: a review of six factors. *J Affect Disord* 2002;**70**:103–15. https://doi.org/10.1016/S0165-0327(01)00376-7

57. Folkman S, Moskowitz JT. Positive affect and the other side of coping. *Am Psychol* 2000;**55**:647–54. https://doi.org/10.1037/0003-066X.55.6.647

58. Fredrickson BL. The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. *Am Psychol* 2001;**56**:218–26. https://doi.org/10.1037/0003-066X.56.3.218

59. Lynch TR, Whalley B, Hempel RJ, Byford S, Clarke P, Clarke S, *et al.* Refractory depression: mechanisms and evaluation of radically open dialectical behaviour therapy (RO-DBT) [REFRAMED]: protocol for randomised trial. *BMJ Open* 2015;**5**:e008857. https://doi.org/10.1136/bmjopen-2015-008857

60. Scott NW, McPherson GC, Ramsay CR, Campbell MK. The method of minimization for allocation to clinical trials. A review. *Control Clin Trials* 2002;**23**:662–74. https://doi.org/10.1016/S0197-2456(02)00242-8

61. Miller SD, Duncan BL, Brown J, Sorrell R, Chalk MB. Using formal client feedback to improve retention and outcome: making ongoing, real-time assessment feasible. *J Brief Therapy* 2006;**5**:5–22.

62. Russell D, Hoare ZS, Whitaker R, Whitaker CJ, Russell IT. Generalized method for adaptive randomization in clinical trials. *Stat Med* 2011;**30**:922–34. https://doi.org/10.1002/sim.4175

63. Williams JB, Kobak KA, Bech P, Engelhardt N, Evans K, Lipsitz J, *et al.* The GRID-HAMD: standardization of the Hamilton Depression Rating Scale. *Int Clin Psychopharmacol* 2008;**23**:120–9. https://doi.org/10.1097/YIC.0b013e3282f948f5

64. First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. (SCID-I/P)*. New York, NY: Biometrics Research, New York State Psychiatric Institute; 2002.

65. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. *Commun Methods Meas* 2007;**1**:77–89. https://doi.org/10.1080/19312450709336664

66. Whalley B. *Running Online Studies with SignalBox*. 2014. URL: www.thesignalbox.net (accessed 10 August 2017).

67. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013;**310**:2191–4. https://doi.org/10.1001/jama.2013.281053

68. National Institute for Health Research Clinical Research Network. *A Pocket Guide to Good Clinical Practice, including the Declaration of Helsinki.* Version 2.1. Leeds: National Institute for Health Research Clinical Research Network; 2011.

69. Lynch TR. *The Skills Training Manual for Radically Open Dialectical Behavior Therapy: A Clinicians Guide for Treating Disorders of Overcontrol*. Reno, NV: Context Press; 2018.

70. Linehan MM. *Skills Training Manual for Treating Borderline Personality Disorder*. New York, NY: Guilford Press; 1993.

71. Safer DL, Robinson AH, Jo B. Outcome from a randomized controlled trial of group therapy for binge eating disorder: comparing dialectical behavior therapy adapted for binge eating to an active comparison group therapy. *Behav Ther* 2010;**41**:106–20. https://doi.org/10.1016/j.beth.2009.01.006

72. Priebe S, Bhatti N, Barnicot K, Bremner S, Gaglia A, Katsakou C, *et al.* Effectiveness and cost-effectiveness of dialectical behaviour therapy for self-harming patients with personality disorder: a pragmatic randomised controlled trial. *Psychother Psychosom* 2012;**81**:356–65. https://doi.org/10.1159/000338897

73. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991;**59**:12–19. https://doi.org/10.1037/0022-006X.59.1.12

74. Leon AC, Solomon DA, Mueller TI, Turvey CL, Endicott J, Keller MB. The Range of Impaired Functioning Tool (LIFE-RIFT): a brief measure of functional impairment. *Psychol Med* 1999;**29**:869–78. https://doi.org/10.1017/S0033291799008570

75. Kriston L, von Wolff A. Not as golden as standards should be: interpretation of the Hamilton Rating Scale for Depression. *J Affect Disord* 2011;**128**:175–7. https://doi.org/10.1016/j.jad.2010.07.011

76. Button KS, Kounali D, Thomas L, Wiles NJ, Peters TJ, Welton NJ, *et al.* Minimal clinically important difference on the Beck Depression Inventory – II according to the patient's perspective. *Psychol Med* 2015;**45**:3269–79. https://doi.org/10.1017/S0033291715001270

77. Beck AT, Steer RA, Brown GK. *Beck Depression Inventory: Second Edition Manual*. San Antonio, TX: Psychological Corporation; 1996.

78. Miller IW, Norman WH, Bishop SB, Dow MG. The Modified Scale for Suicidal Ideation: reliability and validity. *J Consult Clin Psychol* 1986;**54**:724–5. https://doi.org/10.1037/0022-006X.54.5.724

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

79. Linehan MM. *Suicidal Behaviors Questionnaire*. Seattle, WA: University of Washington; 1981.
80. Kuyken W, Byford S, Taylor RS, Watkins E, Holden E, White K, *et al.* Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. *J Consult Clin Psychol* 2008;**76**:966–78. https://doi.org/10.1037/a0013786
81. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D, *et al.* The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med* 2003;**45**:156–74. https://doi.org/10.1097/01.jom.0000052967.43131.51
82. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. https://doi.org/10.1097/00005650-199711000-00002
83. Richardson G, Manca A. Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. *Health Econ* 2004;**13**:1203–10. https://doi.org/10.1002/hec.901
84. Mulhern B, Mukuria C, Barkham M, Knapp M, Byford S, Soeteman D, *et al.* Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. *Br J Psychiatry* 2014;**205**:236–43. https://doi.org/10.1192/bjp.bp.112.122283
85. Hart CM, Ritchie TD, Hepper EG, Gebauer JE. The Balanced Inventory of Desirable Responding Short Form (BIDR-16). *SAGE Open* 2015;**5**:1–9. https://doi.org/10.1177/2158244015621113
86. Letzring TD, Block J, Funder DC. Ego-control and ego-resiliency: generalization of self-report scales based on personality descriptions from acquaintances, clinicians, and the self. *J Res Pers* 2005;**39**:395–422. https://doi.org/10.1016/j.jrp.2004.06.003
87. Mountford V, Corstorphine E, Tomlinson S, Waller G. Development of a measure to assess invalidating childhood environments in the eating disorders. *Eat Behav* 2007;**8**:48–58. https://doi.org/10.1016/j.eatbeh.2006.01.003
88. Neuberg SL, Newsom JT. Personal need for structure: individual differences in the desire for simpler structure. *J Pers Soc Psychol* 1993;**65**:113–31. https://doi.org/10.1037/0022-3514.65.1.113
89. Whiteside SP, Lynam DR, Miller JD, Reynolds SK. Validation of the UPPS impulsive behaviour scale: a four-factor model of impulsivity. *Eur J Pers* 2005;**19**:559–74. https://doi.org/10.1002/per.556
90. First MB, Gibbon M, Spitzer RL, Williams JBW, Benjamin LS. *Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II)*. Washington, DC: American Psychiatric Press, Inc.; 1997.
91. Wilson BA, Emslie H, Foley J, Shiel A, Watson P, Hawkins K, *et al. The Cambridge Prospective Memory Test*. London: Harcourt-Assessment; 2005.
92. McCrae RR, Costa PT. A contemplated revision of the NEO Five-Factor Inventory. *Pers Individ Dif* 2004;**36**:587–96. https://doi.org/10.1016/S0191-8869(03)00118-1
93. Frost RO, Marten P, Lahart C, Rosenblate R. The dimensions of perfectionism. *Cognit Ther Res* 1990;**14**:449–68. https://doi.org/10.1007/BF01172967
94. DiBartolo PM, Frost RO, Chang P, LaSota M, Grills AE. Shedding light on the relationship between personal standards and psychopathology: the case for contingent self-worth. *J Ration Emot Cogn Behav Ther* 2004;**22**:237–50. https://doi.org/10.1023/B:JORE.0000047310.94044.ac
95. DiBartolo PM, Li CY, Frost RO. How do the dimensions of perfectionism relate to mental health? *Cognit Ther Res* 2007;**32**:401–17. https://doi.org/10.1007/s10608-007-9157-7
96. Levinson CA, Rodebaugh TL, Shumaker EA, Menatti AR, Weeks JW, White EK, *et al.* Perception matters for clinical perfectionism and social anxiety. *J Anxiety Disord* 2015;**29**:61–71. https://doi.org/10.1016/j.janxdis.2014.11.002

97. Schwartz SH. Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. In Zanna MP, editor. *Advances in Experimental Social Psychology, Vol. 25*. San Diego, CA: Academic Press; 1992. pp. 1–65. https://doi.org/10.1016/S0065-2601(08)60281-6
98. Parker G, Roussos J, Hadzi-Pavlovic D, Mitchell P, Wilhelm K, Austin MP. The development of a refined measure of dysfunctional parenting and assessment of its relevance in patients with affective disorders. *Psychol Med* 1997;**27**:1193–203. https://doi.org/10.1017/S003329179700545X
99. Bond FW, Hayes SC, Baer RA, Carpenter KM, Guenole N, Orcutt HK, *et al.* Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: a revised measure of psychological inflexibility and experiential avoidance. *Behav Ther* 2011;**42**:676–88. https://doi.org/10.1016/j.beth.2011.03.007
100. King LA, Emmons RA. Conflict over emotional expression: psychological and physical correlates. *J Pers Soc Psychol* 1990;**58**:864–77. https://doi.org/10.1037/0022-3514.58.5.864
101. Neacsiu AD, Rizvi SL, Vitaliano PP, Lynch TR, Linehan MM. The dialectical behavior therapy ways of coping checklist: development and psychometric properties. *J Clin Psychol* 2010;**66**:563–82. https://doi.org/10.1002/jclp.20685
102. Kim Y, Pilkonis PA. Selecting the most informative items in the IIP scales for personality disorders: an application of item response theory. *J Pers Disord* 1999;**13**:157–74. https://doi.org/10.1521/pedi.1999.13.2.157
103. Sarason IG, Sarason BR, Shearin EN, Pierce GR. A brief measure of social support: practical and theoretical implications. *J Soc Pers Relatsh* 1987;**4**:497–510. https://doi.org/10.1177/0265407587044007
104. Wegner DM, Zanakos S. Chronic thought suppression. *J Pers* 1994;**62**:616–40. https://doi.org/10.1111/j.1467-6494.1994.tb00311.x
105. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;**16**:606–13. https://doi.org/10.1046/j.1525-1497.2001.016009606.x
106. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 1988;**54**:1063–70. https://doi.org/10.1037/0022-3514.54.6.1063
107. Stanton AL, Kirk SB, Cameron CL, Danoff-Burg S. Coping through emotional approach: scale construction and validation. *J Pers Soc Psychol* 2000;**78**:1150–69. https://doi.org/10.1037/0022-3514.78.6.1150
108. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry* 2000;**31**:73–86. https://doi.org/10.1016/S0005-7916(00)00012-4
109. Thompson ER. Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *J Cross Cult Psychol* 2007;**38**:227–42. https://doi.org/10.1177/0022022106297301
110. Marmar CR, Weiss DS, Gaston L. Toward the validation of the California Therapeutic Alliance Rating System. *Psycholl Assess* 1989;**1**:46–52. https://doi.org/10.1037/1040-3590.1.1.46
111. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. *Eval Program Plann* 1982;**5**:233–7. https://doi.org/10.1016/0149-7189(82)90074-X
112. Wampold BE, Minami T, Baskin TW, Callen Tierney S. A meta-(re)analysis of the effects of cognitive therapy versus ‘other therapies’ for depression. *J Affect Disord* 2002;**68**:159–65. https://doi.org/10.1016/S0165-0327(00)00287-1

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

113. Bates D, Maechler M, Bolker B, Walker S. *lme4: Linear Mixed-Effects Models Using 'Eigen' and S4*. URL: https://cran.r-project.org/web/packages/lme4/index.html (accessed 1 August 2016).

114. Luke SG. Evaluating significance in linear mixed-effects models in R. *Behav Res Methods* 2017;**49**:1494–502.

115. McElreath R. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Boca Raton, FL: CRC Press; 2016.

116. Curtis L, Burns A. *Unit Costs of Health and Social Care 2015*. Canterbury: PSSRU, University of Kent; 2015.

117. NHS Business Services Authority. *NHS Prescription Services*. NHS; 2016. URL: www.nhsbsa.nhs.uk/PrescriptionServices/ (accessed 2 August 2016).

118. Department of Health and Social Care (DHSC). *NHS Reference Costs 2014 to 2015*. London: DHSC; 2015. URL: www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015 (accessed 2 August 2016).

119. Curtis L, Burns A. *Unit Costs of Health and Social Care 2014*. Canterbury: PSSRU, University of Kent; 2014.

120. Third Sector. *Analysis: Counting the Cost of Reform at Samaritans*. 2013. URL: www.thirdsector.co.uk/analysis-counting-cost-reform-samaritans/management/article/1175711 (accessed 2 August 2016).

121. NHS Business Services Authority. *Drug Tariff Part VIII*. URL: www.nhsbsa.nhs.uk/PrescriptionServices/1821.aspx (accessed 2 August 2016).

122. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 2005.

123. Woo JM, Kim W, Hwang TY, Frick KD, Choi BH, Seo YJ, *et al.* Impact of depression on work productivity and its improvement after outpatient treatment with antidepressants. *Value Health* 2011;**14**:475–82. https://doi.org/10.1016/j.jval.2010.11.006

124. Netten A, Knight J, Dennett J, Cooley R, Slight A. *A Ready Reckoner for Staff Costs in the NHS*. Canterbury: University of Kent; 1998.

125. Barrett B, Byford S. The Challenges of Estimating the Unit Cost of Group Based Therapies. In Curtis L, Burns A, editors. *Unit Costs of Health and Social Care 2008*. Canterbury: PSSRU, University of Kent; 2008.

126. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;**30**:377–99. https://doi.org/10.1002/sim.4067

127. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ* 1994;**3**:309–19. https://doi.org/10.1002/hec.4730030505

128. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. *Br J Psychiatry* 2005;**187**:106–8. https://doi.org/10.1192/bjp.187.2.106

129. Kessler RC, Barber C, Birnbaum HG, Frank RG, Greenberg PE, Rose RM, *et al.* Depression in the workplace: effects on short-term disability. *Health Aff* 1999;**18**:163–71. https://doi.org/10.1377/hlthaff.18.5.163

130. Emsley R, Dunn G, White IR. Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. *Stat Methods Med Res* 2010;**19**:237–70. https://doi.org/10.1177/0962280209105014

131. Robins J. Correcting for non-compliance in randomized trials using structural nested mean models. *Commun Stat Theory Methods* 1994;**23**:2379–412. https://doi.org/10.1080/03610929408831393

132. Have TR, Joffe MM, Lynch KG, Brown GK, Maisto SA, Beck AT. Causal mediation analyses with rank preserving models. *Biometrics* 2007;**63**:926–34. https://doi.org/10.1111/j.1541-0420.2007.00766.x

133. Bloom HS. *Learning More from Social Experiments: Evolving Analytic Approaches*. New York, NY: Russell Sage Foundation; 2005.

134. MacKinnon DP. *Introduction to Statistical Mediation Analysis*. New York, NY: Lawrence Erlbaum Associates; 2008.

135. Small D. Mediation analysis without sequential ignorability: using baseline covariates interacted with random assignment as instrumental variables. *J Stat Res* 2012;**46**:91–103.

136. Hansen LP. Large sample properties of generalized method of moments estimators. *Econometrica* 1982;**50**:1029–54. https://doi.org/10.2307/1912775

137. Anker MG, Duncan BL, Sparks JA. Using client feedback to improve couple therapy outcomes: a randomized clinical trial in a naturalistic setting. *J Consult Clin Psychol* 2009;**77**:693–704. https://doi.org/10.1037/a0016062

138. Wampold BE. *The Great Psychotherapy Debate: Models, Methods, and Findings*. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.

139. Hersoug AG, Høglend P, Monsen JT, Havik OE. Quality of working alliance in psychotherapy: therapist variables and patient/therapist similarity as predictors. *J Psychother Pract Res* 2001;**10**:205–16.

140. Herman SM. Therapist–client similarity on the multimodal structural profile inventory as a predictor of early session impact. *J Psychother Pract Res* 1997;**6**:139–44.

141. Hunt DD, Carr JE, Dagadakis CS, Walker EA. Cognitive match as a predictor of psychotherapy outcome. *Psychother Theor Res Pract Train* 1985;**22**:718–21. https://doi.org/10.1037/h0085558

142. Stock J, Yogo M. Testing for Weak Instruments in Linear IV Regression. In Donald WKA, editor. *Identification and Inference for Econometric Models*. New York, NY: Cambridge University Press; 2005. pp. 80–108. https://doi.org/10.1017/CBO9780511614491.006

143. McCambridge J, Kypri K, Elbourne D. In randomization we trust? There are overlooked problems in experimenting with people in behavioral intervention trials. *J Clin Epidemiol* 2014;**67**:247–53. https://doi.org/10.1016/j.jclinepi.2013.09.004

144. Preference Collaborative Review Group. Patients' preferences within randomised trials: systematic review and patient level meta-analysis. *BMJ* 2008;**337**:a1864. https://doi.org/10.1136/bmj.a1864

145. Cook TD, Campbell DT. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago, IL: Houghton Mifflin; 1979.

146. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74. https://doi.org/10.2307/2529310

147. Furler J, Magin P, Pirotta M, van Driel M. Participant demographics reported in 'Table 1' of randomised controlled trials: a case of 'inverse evidence'? *Int J Equity Health* 2012;**11**:14. https://doi.org/10.1186/1475-9276-11-14

148. Ministry of Housing, Communities and Local Government. *English Indices of Deprivation*. 2015. URL: http://opendatacommunities.org/data/societal-wellbeing/imd/indices (accessed 1 July 2017).

149. Welsh Government. *Welsh Index for Multiple Deprivation*. 2014. URL: http://gov.wales/statistics-and-research/welsh-index-multiple-deprivation/?lang=en (accessed 1 July 2017).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

150. Payne R, Abel G. UK indices of multiple deprivation – a way to make comparisons across constituent countries easier. *Health Stat Q* 2012;**53**:22–37.

151. Gelman A. Analysis of variance – why it is more important than ever. *Ann Stat* 2005;**33**:1–53. https://doi.org/10.1214/009053604000001048

152. Soyer E, Hogarth RM. The illusion of predictability: how regression statistics mislead experts. *Int J Forecast* 2012;**28**:695–711. https://doi.org/10.1016/j.ijforecast.2012.02.002

153. Romeo R, Patel A, Knapp M, Thomas C. The cost-effectiveness of mirtazapine versus paroxetine in treating people with depression in primary care. *Int Clin Psychopharmacol* 2004;**19**:125–34. https://doi.org/10.1097/00004850-200405000-00002

154. Watts SE, Turnell A, Kladnitski N, Newby JM, Andrews G. Treatment-as-usual (TAU) is anything but usual: a meta-analysis of CBT versus TAU for anxiety and depression. *J Affect Disord* 2015;**175**:152–67. https://doi.org/10.1016/j.jad.2014.12.025

155. Newton-Howes G, Tyrer P, Johnson T. Personality disorder and the outcome of depression: meta-analysis of published studies. *Br J Psychiatry* 2006;**188**:13–20. https://doi.org/10.1192/bjp.188.1.13

156. Fournier JC, DeRubeis RJ, Shelton RC, Gallop R, Amsterdam JD, Hollon SD. Antidepressant medications v. cognitive therapy in people with depression with or without personality disorder. *Br J Psychiatry* 2008;**192**:124–9. https://doi.org/10.1192/bjp.bp.107.037234

157. Goddard E, Wingrove J, Moran P. The impact of comorbid personality difficulties on response to IAPT treatment for depression and anxiety. *Behav Res Ther* 2015;**73**:1–7. https://doi.org/10.1016/j.brat.2015.07.006

158. Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. *Arch Gen Psychiatry* 1999;**56**:617–26. https://doi.org/10.1001/archpsyc.56.7.617

159. Oquendo MA, Currier D, Mann JJ. Prospective studies of suicidal behavior in major depressive and bipolar disorders: what is the evidence for predictive risk factors? *Acta Psychiatr Scand* 2006;**114**:151–8. https://doi.org/10.1111/j.1600-0447.2006.00829.x

160. Chen EY, Segal K, Weissman J, Zeffiro TA, Gallop R, Linehan MM, *et al.* Adapting dialectical behavior therapy for outpatient adult anorexia nervosa – a pilot study. *Int J Eat Disord* 2015;**48**:123–32. https://doi.org/10.1002/eat.22360

161. Keogh K, Booth R, Baird K, Gibson J, Davenport J. The Radical Openness Group: a controlled trial with 3-month follow-up. *Pract Innov* 2016;**1**:129. https://doi.org/10.1037/pri0000023

162. Brazier J, Tumur I, Holmes M, Ferriter M, Parry G, Dent-Brown K, *et al.* Psychological therapies including dialectical behaviour therapy for borderline personality disorder: a systematic review and preliminary economic evaluation. *Health Technol Assess* 2006;**10**(35). https://doi.org/10.3310/hta10350

163. Turner RM. Naturalistic evaluation of dialectical behavior therapy-oriented treatment for borderline personality disorder. *Cogn Behav Pract* 2000;**7**:413–19. https://doi.org/10.1016/S1077-7229(00)80052-8

164. Linehan MM, Armstrong HE, Suarez A, Allmon D, Heard HL. Cognitive-behavioral treatment of References 58 chronically parasuicidal borderline patients. *Arch Gen Psychiatry* 1991;**48**:1060–4. https://doi.org/10.1001/archpsyc.1991.01810360024003

165. Van den Bosch LM, Verheul R, Schippers GM, Van den Brink W. Dialectical behavior therapy of borderline patients with and without substance use problems. Implementation and long-term effects. *Addict Behav* 2002;**27**:911–23. https://doi.org/10.1016/S0306-4603(02)00293-9

166. Koons CR, Robins CJ, Tweed JL, Lynch TR, Gonzalez AM, Morse JQ, *et al.* Efficacy of dialectical behavior therapy in women veterans with borderline personality disorder. *Behav Ther* 2001;**32**:371–90. https://doi.org/10.1016/S0005-7894(01)80009-5

167. National Institute for Health and Clinical Excellence (NICE). *Borderline Personality Disorder: Treatment and Management*. London: NICE; 2009.

168. Bryan AE, Arkowitz H. Meta-analysis of the effects of peer-administered psychosocial interventions on symptoms of depression. *Am J Community Psychol* 2015;**55**:455–71. https://doi.org/10.1007/s10464-015-9718-y

169. Lopez A, Chessick CA. DBT graduate group pilot study: a model to generalize skills to create a 'life worth living'. *Social Work in Mental Health* 2013;**11**:141–53. https://doi.org/10.1080/15332985.2012.755145

170. Harley R, Sprich S, Safren S, Jacobo M, Fava M. Adaptation of dialectical behavior therapy skills training group for treatment-resistant depression. *J Nerv Ment Dis* 2008;**196**:136–43. https://doi.org/10.1097/NMD.0b013e318162aa3f

171. Shedler J, Mayman M, Manis M. The illusion of mental health. *Am Psychol* 1993;**48**:1117–31. https://doi.org/10.1037/0003-066X.48.11.1117

172. Myers LB. The importance of the repressive coping style: findings from 30 years of research. *Anxiety Stress Coping* 2010;**23**:3–17. https://doi.org/10.1080/10615800903366945

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 1 Additional tables

TABLE 19 The effect sizes for RO DBT for TRD from pilot studies (for group differences in change on the HRSD score from baseline to the end of treatment)

Study	Treatment	Number of participants	Group difference	Cohen's *d*
Lynch *et al.*[49]	RO DBT-E	31	2.75	0.71
Lynch *et al.*[48]	RO DBT-E	31	5.47	0.85
Harley *et al.*[170]	DBT for TRD	19	3.32	1.45

TABLE 20 A breakdown of study and treatment withdrawals

Reasons for withdrawal/loss to follow-up	Time point (*n*)							
	Immediately after randomisation		Month					
			7		12		18	
	TAU	RO DBT	TAU	RO DBT	TAU	RO DBT	TAU	RO DBT
Did not withdraw from study, only from RO DBT treatment		1 (withdrawn)						
Unable to commit because of life/work/family circumstances	2	5 (withdrawn)	2	1	1	1		3
Unable to commit because of physical illness: not related to depression				1 (withdrawn)				
Unhappy with RO DBT treatment				3; 1 (withdrawn)		1		
Unhappy with TAU allocation	7		1		1			
Did not want to give reason				1 (withdrawn)		1	1	
Lost to follow-up: did not respond to contact attempts	6	5 (withdrawn)		5; 1 (withdrawn)	1			3
Lost to follow-up: death				1				
Total	15	11	3	14	3	3	1	6

Withdrawn: if patients attended ≤ 3 sessions, they were classified as withdrawn from RO DBT treatment.

TABLE 21 Age (years) by site and group

Site	Mean age (years) (SD)	Number of participants
TAU		
1	42 (11.4)	43
2	53 (8.1)	33
3	51 (9.9)	12
RO DBT		
1	49 (12.0)	74
2	46 (10.1)	62
3	46 (14.2)	26

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 22 Basic clinical data per site

Clinical variable	Site *n* 1	 2	 3	 *N* 1	 2	 3	 % 1	 2	 3
Number of past depressive episodes									
0	26	28	7	61	81	26	42.6	34.6	26.9
1	5	3	3	61	81	26	8.2	3.7	11.5
2–3	9	9	6	61	81	26	14.8	11.1	23.1
4–5	9	4	3	61	81	26	14.8	4.9	11.5
≥ 6	12	37	7	61	81	26	19.7	45.7	26.9
Chronic vs. recurrent depression									
Chronic	77	74	28	91	86	36	84.6	86.1	77.8
Recurrent	14	12	8	91	86	36	15.4	14.0	22.2
Age (years) at onset categorised									
< 16	44	39	9	117	95	38	37.6	41.1	23.7
16–21	27	16	6	117	95	38	23.1	16.8	15.8
22–30	20	20	10	117	95	38	17.1	21.1	26.3
31–50	22	19	13	117	95	38	18.8	20.0	34.2
≥ 50	4	1	0	117	95	38	3.4	1.1	0.0
Previous psychotherapy received	94	69	28	109	88	37	86.2	78.4	75.7
Comorbid DSM-IV Axis disorder									
Axis I									
Generalised anxiety disorder	55	46	18	110	93	36	50.0	49.5	50.0
Social phobia	54	52	14	63	68	20	85.7	76.5	70.0
Panic disorder	35	31	13	54	69	22	64.8	44.9	59.1
Specific phobia	21	32	11	30	58	14	70.0	55.2	78.6
Agoraphobia without history of panic disorder	11	17	5	20	47	8	55.0	36.2	62.5
Obsessive–compulsive disorder	24	14	7	36	49	12	66.7	28.6	58.3
Post-traumatic stress disorder	23	25	7	38	60	15	60.5	41.7	46.7
Anxiety disorder not otherwise specified	8	2	1	19	28	4	42.1	7.1	25.0
Anorexia nervosa	3	8	0	14	28	5	21.4	28.6	0.0
Bulimia nervosa	3	1	0	13	23	5	23.1	4.4	0.0
Binge-eating disorder	9	7	3	15	27	7	60.0	25.9	42.9
Alcohol abuse/dependence	0	1	0	30	60	12	0.0	1.7	0.0
Cocaine abuse/dependence	0	0	1	14	47	4	0.0	0.0	25.0
Hypochondriasis	2	6	1	111	92	36	1.8	6.5	2.8
Body dysmorphic disorder	14	10	2	111	93	36	12.6	10.8	5.6
Pain disorder	1	8	2	110	92	36	0.9	8.7	5.6
Somatisation disorder	0	6	2	109	91	36	0.0	6.6	5.6
Somatoform disorder	0	3	0	110	92	36	0.0	3.3	0.0
Bipolar I disorder	2	0	0	32	52	5	6.3	0.0	0.0
Any other current Axis-I disorder	1	0	0	4	8	2	25.0	0.0	0.0

TABLE 22 Basic clinical data per site (*continued*)

	Site								
	n			*N*			%		
Clinical variable	1	2	3	1	2	3	1	2	3
Axis II									
Schizotypal PD	5	16	1	116	95	38	4.3	16.8	2.6
Paranoid PD	18	30	9	116	95	38	15.5	31.6	23.7
Schizoid PD	8	15	2	112	94	38	7.1	16.0	5.3
Avoidant PD	58	62	16	117	95	38	49.6	65.3	42.1
Obsessive–compulsive PD	51	50	15	116	95	38	44.0	52.6	39.5
Dependent PD	10	12	1	114	95	38	8.8	12.6	2.6
Conduct disorder	6	7	4	97	92	35	6.2	7.6	11.4
Total[a]									
Axis I									
0	17	11	5	117	95	38	14.5	11.6	13.2
1	21	18	8	117	95	38	18.0	19.0	21.1
2	34	18	8	117	95	38	29.1	19.0	21.1
3	21	17	9	117	95	38	18.0	17.9	23.7
4	10	12	6	117	95	38	8.6	12.6	15.8
5	11	6	0	117	95	38	9.4	6.3	0.0
6	2	5	2	117	95	38	1.7	5.3	5.3
7	1	8	0	117	95	38	0.9	8.4	0.0
Axis II									
0	32	10	11	117	95	38	27.4	10.5	29.0
1	35	23	12	117	95	38	29.9	24.2	31.6
2	37	30	10	117	95	38	31.6	31.6	26.3
3	7	23	4	117	95	38	6.0	24.2	10.5
4	4	6	1	117	95	38	3.4	6.3	2.6
5	2	2	0	117	95	38	1.7	2.1	0.0
6	0	1	0	117	95	38	0.0	1.1	0.0
Axes I and II									
0	5	2	2	117	95	38	4.3	2.1	5.3
1–2	39	19	11	117	95	38	33.3	20.0	29.0
3–4	38	27	14	117	95	38	32.5	28.4	36.8
5–7	29	31	10	117	95	38	24.8	32.6	26.3
8–10	5	13	1	117	95	38	4.3	13.7	2.6
11–12	1	3	0	117	95	38	0.8	3.2	0.0

a Number of patients experiencing several mental health disorders listed separately for Axis I, clinical disorders (ranging from 0 to 7) and Axis II, personality disorders (ranging from 0 to 6), as well as the number of patients experiencing combinations of both Axis I and Axis II disorders (ranging from 0 to 12).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 23 The number of treatment attendances by site

Site	Number of treatment attendances				
	Mean		SD		
	Attended	Missed	Attended	Missed	Attended (%)
Dorset	22.27	4.96	8.58	5.54	81.78
Hampshire	21.98	2.28	9.04	2.43	90.61
North Wales	25.33	4.96	2.91	2.33	83.63

TABLE 24 The RO DBT group session attendance across treatment sites

Site	Number of participants who attended at least one skills group session	Mean number of sessions		
		Attended	SD	Median
Dorset	71	19.67	8.03	23.00
Hampshire	56	20.03	8.05	23.00
North Wales	24	20.50	5.30	21.50

TABLE 25 The variance explained at each level of the primary outcome model with 95% highest probability density interval

Level	% of variance	95% CI
Therapist	12	0 to 23
Patient	61	50 to 73
Residual	27	21 to 33

TABLE 26 The range of expected prognoses for patients receiving TAU or RO DBT based on final HRSD score

Group	Centile				
	5th	25th	50th	75th	95th
Month 7					
RO DBT	8.17	12.52	15.57	18.67	22.64
TAU	13.03	16.67	19.66	22.28	26.25
Month 12					
RO DBT	6.66	10.99	14.12	17.15	21.70
TAU	9.17	13.19	15.92	18.74	23.15
Month 18					
RO DBT	11.46	15.40	18.12	20.71	24.53
TAU	13.19	16.75	19.17	21.62	25.05

Columns refer to the percentile of expected outcomes. The 5th centile would be the final HRSD score achieved by the most fortunate 5% of patients. The 50th centile refers to the median outcome.

TABLE 27 The expected additional change in symptoms for the 5th to 95th centiles of patients when treated with RO DBT rather than TAU, based on final HRSD score

	Centile				
Month	5th	25th	50th	75th	95th
7	–11.71	–7.05	–4.03	–0.83	3.64
12	–9.73	–4.80	–1.89	1.19	5.82
18	–6.73	–3.38	–1.13	1.07	4.89

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 2 Trial procedure for reporting serious adverse events

The Refractory depression: Mechanisms and Efficacy of RO DBT trial procedure for evaluating and reporting adverse events and serious adverse events

Governance identification numbers.

Funder: Efficacy and Mechanism Evaluation report number 09/150/12.

National Institute for Health Research UK Clinical Research Network: 10234.

Research Ethics Committee: reference number 11/SC/0146.

International Standard Randomised Controlled Trial Number: ISRCTN85784627.

Sponsor: University of Southampton 8014.

Background on suicide risk in patients with treatment-resistant depression

To date, there are no discernible adverse reactions to the psychosocial treatment offered in the study during the pilot stage, nor has this been a problem in the two prior RCTs[48,49] examining DBT for refractory depression. Suicidal behaviour, however, is a potential risk when treating severely depressed patients. In addition, the RefraMED trial does not exclude patients with a history of suicide because (1) the experimental condition (DBT) is well-established as a treatment for suicidal behaviour (e.g. see NICE's guidelines for BPD[162]) and expert DBT therapists with considerable experience in treating suicidal behaviour were recruited as study therapists and (2) the RefraMED trial protocol is purposely designed to include the most difficult-to-treat patients. Although the majority of prior RCTs for TRD or chronic depression have either excluded or not measured comorbid PD, suicidal behaviour, prior psychotherapy treatment and patients with high rates of depression relapse, these are all being addressed in this study. This maximises ecological validity and allows for the determination of potential moderators of treatment response.

Given that the odds ratio of a suicide attempt is particularly high among individuals with a history of three or more disorders (odds ratio 19.7) (see Kessler *et al.*[158]) we can expect that the RefraMED trial will have a higher rate of reported suicides than other studies examining TRD or chronic depression who purposely exclude PDs, comorbid disorders, or patients with suicidal behaviour. Interestingly, an estimated 40–60% of unipolar depressed patients meet the criteria for comorbid PD (e.g. Fava *et al.*,[32] Klein *et al.*[33] and Riso *et al.*[34]) with overcontrolled PDs being the most frequent, that is, paranoid, avoidant and obsessive–compulsive PD (e.g. Candrian *et al.*,[35] Fava *et al.*[32] and Fournier *et al.*[34]). Overcontrolled PDs are expected to be highly represented in the RefraMED trial, and these PDs are associated with emotional avoidance, emotional constriction, stoicism and high social desirability. That is, even when obviously distressed, individuals may avoid revealing emotional or physical distress. They share features with a subgroup of individuals who exhibit what has been referred to as an *illusion of mental health*.[171] These 'defensive-deniers' report low levels of distress on standardised mental health measures (possibly up to ≈30% of community samples), but are rated as highly distressed by clinicians while exhibiting high verbal defensiveness and heightened physiological reactivity.[171] This illusory style shares many of the characteristics associated with a *repressive-defensive* style of coping that has been linked with numerous psychological and physical problems and is characterised by low scores on trait anxiety scales, high

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

scores on defensiveness scales and/or social desirability scales and high psychophysiological arousal (see Myers *et al.*[172] for a review).

The expectation that the RefraMED trial will include a high number of defensive-deniers among the recruited sample presents methodological challenges in the assessment of risk in the two conditions. First, the research design of the RefraMED trial is such that the subjects in the experimental condition (DBT) will be likely to receive considerably more therapist/care-giver contact relative to TAU. In addition, the DBT treatment condition is designed to encourage honest patient reporting regarding suicidal behaviour, as well as a range of other emotionally relevant information. In fact, a willingness to discuss these oftentimes previously avoided topics is considered a sign of therapeutic progress. However, this could prove potentially problematic when it comes to reporting AEs and SAEs as the treatment and attention provided in the experimental condition may arbitrarily inflate the reporting of AEs and SAEs relative to TAU, simply as a function of therapeutic progress. Essentially, if, for some subjects, self-report measures of mental health assess not health but defensiveness, then an effect of successful psychotherapy (i.e. DBT) might be to lower scores on these measures (or conversely, to raise scores on measures of distress). For less defended subjects, successful therapy should raise scores on measures of mental health. Failure to distinguish between these groups could, therefore, cloud the results of the RefraMED trial, obscuring treatment effects that may actually be present and/or inadvertently heightening concern regarding the safety of the DBT experimental condition. To help manage this issue, the RefraMED trial is the first clinical trial, to our knowledge, to take into account this potentially unique moderator/mediator related to under-reporting psychological distress by including a measure of social desirability. However, it is also important that the protocols for reporting AEs and SAEs account for the abovementioned potential '*reporting bias*' among the experimental condition.

Definitions of adverse event, adverse reaction, serious adverse event, serious adverse reaction and suspected unexpected serious adverse reaction

All AEs will be assessed for causality, seriousness and expectedness that is, whether or not the AE is related to the behavioural intervention, whether or not the AE is serious and whether or not the AE was unexpected. *Table 28* shows how these three features apply to the main types of AEs.

Seriousness

Any untoward medical occurrence will be deemed serious if it:

- results in death
- is life-threatening
- requires hospitalisation, accident and emergency (A&E) department visit or prolongation of existing hospitalisation
- results in persistent or significant disability or incapacity
- results in a congenital anomaly or birth defect.

TABLE 28 Types of AEs

	Types of AE				
Relatedness, seriousness and expectancy	AEs	Adverse reactions	SAEs	SARs	SUSARs
Is the clinical occurrence considered to be related to trial intervention?	No	Yes	No	Yes	Yes
Is the clinical occurrence serious?	No	No	Yes	Yes	Yes
Is the clinical occurrence unexpected?	No	No	No	No	Yes

Attempted suicide and acts of self-harm that are life-threatening or that require hospital inpatient care will also be considered serious in this study.

All serious events not considered to have a causal relationship with the treatment will be classified and reported as SAEs. All serious events that are considered to have a causal relationship with the treatment will be classified and reported as SARs.

Causality

Causality is the degree to which an untoward medical occurrence can be attributed to the intervention and can be classed as unknown, definitely not related, remotely related, possibly related, probably related or definitely related (see below for definitions). Only untoward medical occurrences that are considered to be possibly, probably or definitely related to the intervention will be reported as having a causal relationship.

If the untoward medical occurrence *is not* considered to have a causal relationship with the treatment at the time of the event (i.e. it is not believed to be a consequence of DBT) this will be classified as an AE. However, if it is considered to have a causal relationship with DBT at the time of the event it will be classified as an adverse reaction.

The investigator is responsible for defining, in his/her best judgement, the relationship of the AE/SAE to the study behavioural interventions. The degree of certainty for which the AE/SAE is attributed to the study interventions or alternative causes (e.g. natural history of the underlying disease, concomitant therapies) should be determined by how well the experience can be understood in terms of one or more of the following:

- Exposure – is there evidence that the subject was actually exposed to the behavioural interventions?
- Timing of the study behavioural interventions – did the AE/SAE follow in a reasonable temporal sequence from administration of the behavioural intervention?
- Alternative explanations for the AE such as concomitant medications, concurrent illness, non-medicinal therapies, diagnostic tests, procedures or other confounding findings.
- Response to discontinuation of the study behavioural interventions.

Terms and definitions to be used in assessing the behavioural intervention relationship to the AE/SAE are:

- unknown – use this category only if the cause of the AE/SAE is not possible to determine
- definitely not related – the subject did not receive the behavioural intervention, the temporal sequence of the AE/SAE onset relative to administration of the behavioural intervention is not reasonable, or there is another obvious cause of the AE/SAE
- remotely related – there is evidence of exposure to the behavioural intervention or there is another more likely cause of the AE/SAE
- possibly related – there is evidence of exposure to the behavioural interventions, the temporal sequence of the AE/SAE onset relative to administration of the behavioural interventions is reasonable, but the AE/SAE could have been because of another equally likely cause
- probably related – there is evidence of exposure to the behavioural interventions, the temporal sequence of the AE/SAE onset relative to administration of the behavioural interventions is reasonable, and the AE/SAE is more likely explained by the behavioural interventions than by any other cause
- definitely related – there is evidence of exposure to the behavioural interventions, the temporal sequence of the AE/SAE onset relative to administration of the behavioural interventions is reasonable, the AE/SAE is more likely explained by the behavioural interventions than by any other cause, and the AE/SAE shows a pattern consistent with previous knowledge of the behavioural interventions or behavioural interventions class.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Expectedness

An untoward medical occurrence will be considered to be ‘unexpected’ if its nature and severity are not consistent with the information in the summary of product characteristics for that treatment. If an AE is (1) considered to be related to the behavioural intervention, (2) is serious and (3) is unexpected, then it will be classed as a SUSAR.

Collection and reporting of adverse events and serious adverse events

All AEs will be initially categorised by severity and source of report (e.g. assessor, therapist, subject). Any AE that cannot clearly be categorised as both non-serious and non-study or non-intervention related is subsequently reviewed and categorised by the principal investigator and the chief investigator. Serious and study-related AEs are then entered in a database. All investigator correspondence is also retained.

1. Record AEs as soon known.
2. Report the severity of the event following the guidance below.
3. Report the relatedness of the event to the study agent administration.

Adverse events that will be reported to the DMEC in the annual report:

- formal complaints
- non-suicidal self-injury (non-life-threatening)
- self-injury with suicidal intention (non-life-threatening).

Serious adverse events that will be reported to the DMEC immediately:

- all deaths
- actual suicide
- hospitalisation or A&E visit
- serious events leading to long-term disability or incapacity.

Reporting bias

In order to obtain potentially useful data to more fully understand the ‘reporting bias’, we will also report the source of the AE and SAE data, specifically whether the information was obtained from the patient, therapist, assessor, non-study health-care provider or patient record. In addition, assessors will be trained to assess factors related to self-disclosure, in particular (1) has this event happened before and, if so, (2) is this the first time the patient has disclosed this to someone?

Adverse events that we expect may differ between conditions because of the ‘reporting bias’:

- self-reported suicidal behaviours
- self-reported non-suicidal self-injury.

Serious adverse events that may differ between conditions because of the ‘reporting bias’:

- self-reported hospitalisations, A&E visits.

Responsibility for reporting

The reporting requirements differ depending on the causality, seriousness and expectedness of the medical occurrence, as summarised in the flow diagram of safety reporting (*Figure 7*).

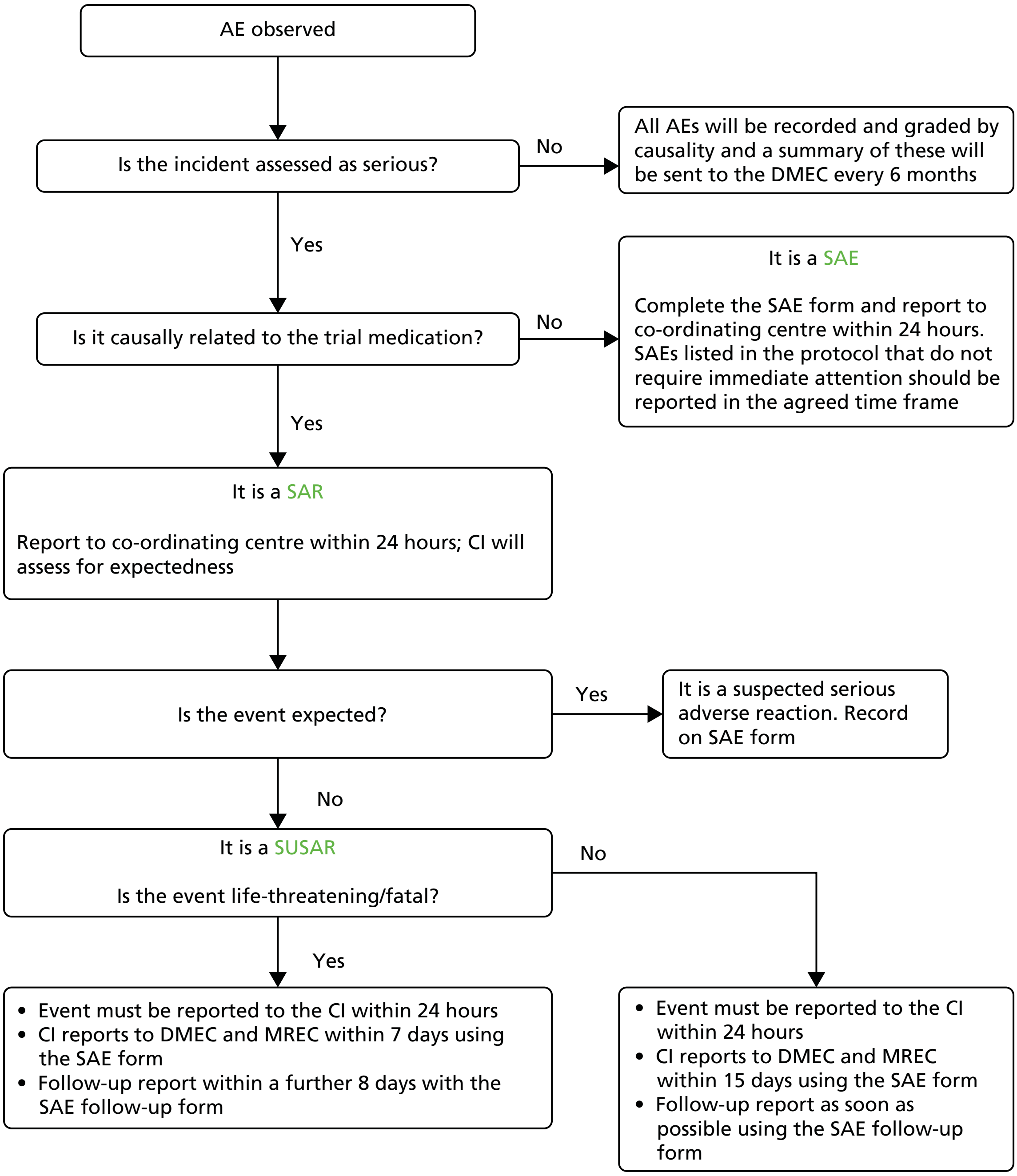
AE observed
Is the incident assessed as serious?
No
All AEs will be recorded and graded by causality and a summary of these will be sent to the DMEC every 6 months
Yes
Is it causally related to the trial medication?
No
It is a SAE
Complete the SAE form and report to co-ordinating centre within 24 hours. SAEs listed in the protocol that do not require immediate attention should be reported in the agreed time frame
Yes
It is a SAR
Report to co-ordinating centre within 24 hours; CI will assess for expectedness
Is the event expected?
Yes
It is a suspected serious adverse reaction. Record on SAE form
No
It is a SUSAR
Is the event life-threatening/fatal?
No
Yes
• Event must be reported to the CI within 24 hours
• CI reports to DMEC and MREC within 7 days using the SAE form
• Follow-up report within a further 8 days with the SAE follow-up form
• Event must be reported to the CI within 24 hours
• CI reports to DMEC and MREC within 15 days using the SAE form
• Follow-up report as soon as possible using the SAE follow-up form

FIGURE 7 Flow chart of safety reporting. CI, chief investigator.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

RefraMED SAE–SUSAR Report Form

REFRAMED SAE-SUSAR report form

SAE ID	PATIENT ID	DATE OF ONSET	ADVERSE EVENT / ADVERSE REACTION	RANDOMISATION DATE	STUDY STATUS	TIME AFTER RANDIMISATION	EVENT/ REACTION	CLINICAL ACTION

REFRAMED Trial office phone: 023 8059 5077

Participant study ID: SAE/SUSAR ID:

Once a SUSAR or SAE has been identified, please **complete sections 1-5** on this form with as much information as possible before emailing it (encrypted) to the Trial Team: reframed@soton.ac.uk within 24 hours of knowledge of the event.

Please update all sections as information becomes available and complete **sections 6-8** before resending the forms at the follow-up reporting period **within 5 days of sending the initial or previous report** until the SUSAR has been resolved or a decision for no further follow up has been taken (if further follow up is required please copy and use the last page if this form.

1. Details of the study					
Full title:	REFRActory depression - Mechanisms and Efficacy of Dialectical Behaviour Therapy				
Study Centre (Trust):					
R&D No:					
Ethics No:	11/SC/0146				
UKCRN No:	10234				
ISCRTN No:	ISRCTN85784627				
RGO ID Southampton	8014				

2. Details of participant affected by SAE/SUSAR

Study ID:	Initials:	DOB (dd/mm/yyyy)	Gender:

Disease history: prior diseases suffered by the participant not being treated by the current study

Disease name	Start Date	End Date	Continuing (yes/no/unknown)

Drug history: ALL medication taken within last 6 months

Drug name	Start Date	End Date

3. Reaction Details

Country of origin:	

Reaction/SAE	Outcome (recovered, recovering, not recovered, fata, unknown)	Start Date	End date

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Narrative: detailed description of SAE/SUSAR			
Setting (e.g. hospital, out-patient clinic, home):			
Body sites(s):			
Diagnosis (if available)			
Other information (including severity):			
Seriousness			
☐ Death ☐ Life threatening ☐ Hospitalisation ☐ Disabling ☐ Congenital anomaly ☐ Other			
Details of medical tests undertaken relevant to SAE/SUSAR			
Test	Result	Unit	Test Date

4. Action Taken at 24 h (initial reporting interval)

5. Details of Principal Investigator at this site writing the initial report details (at 24 h)		
Name:		
Job title/Role in Study:		
Contact Address:		
Email Address		
Telephone No:		
Signature		Date:
Additional Information		
Section Number	Further Information	

6. Update on outcome assessed at follow-up reporting interval (within 5 days of the initial or previous report)*
☐ Resolved ☐ Ongoing ☐ Died
*Give details of outcome indicated (including post-mortem details in case of death)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

7. Details of any further action taken since first report

8. Details of Principal Investigator or delegated physician at this site writing the follow-up report (if different from Section 5)		
Name:		
Job title/Role in Study:		
Contact Address:		
Email Address		
Telephone No:		
Signature		Date:

Serious adverse event and serious adverse reaction

If an AE occurs that is anticipated or no more severe than expected, it will be summarised in the progress report reported to the DMEC, TSC, sponsor and the Efficacy and Mechanism Evaluation programme at 6 monthly intervals.

Suspected unexpected serious adverse reaction

If the event is considered to be a SUSAR, the DMEC, sponsor and Research Ethics Committee will be notified immediately (*Table 29*).

All AEs will be recorded on the trial database, evaluated by the principal investigator or other designated person responsible for the clinical aspect of the trial at each site and included in the annual safety report to the Medical Research and Ethics Committee (MREC) and the Medicines and Healthcare products Regulatory Agency.

However, for all events considered to be serious additional reporting is required. The trial co-ordinating centre must be informed within 24 hours of the site principal investigator's knowledge of the event. A 'Serious Adverse Event Form' must be completed and sent to the trial coordinating centre as soon as possible. The trial manager will immediately contact the clinical and methodological chief investigators who have been delegated the responsibility of reporting any SAEs on behalf of the sponsor (University of Southampton). A decision must then be made by the investigator with the clinical responsibility for aspects of the patient's care that are relevant to the trial whether or not an AE is related to DBT. The investigator should decide the degree to which the event is caused by the intervention. If a decision cannot be made, the investigator must contact the co-ordinating centre. Advice will be sought from the clinical chief investigator and other clinicians may be asked if a decision cannot be reached.

TABLE 29 Event types, actions and deadlines

Event type	Reporting actions	Timeline
All AEs	Record on trial database	Annual safety report to be sent to the DMEC
SAEs	Record on trial database	
	Complete SAE form and PI reports to TM/CI	Within 24 hours
	CI to decide causality and expectedness	
Serious but expected reactions or non-SARs	Record on trial database	
	Complete SAE form and PI reports to TM/CI	Within 24 hours
Life-threatening SUSAR	Record on trial database	
	Complete SAE form and PI reports to TM/CI	Within 24 hours
	CI to decide causality and expectedness	Within 7 calendar days
	SAE form to the DMEC and MREC	Additional 8 calendar days
	Follow-up and final SAE form to the DMEC and MREC	
Non-fatal/non-life-threatening SUSAR	Record on trial database	
	Complete SAE form and PI reports to TM/CI	Within 24 hours
	CI to decide causality and expectedness	
	SAE form to the DMEC and MREC	Within 15 calendar days
	Follow-up and final SAE form to the DMEC and MREC	As soon as possible

CI, chief investigator; MREC, Medical Research Ethics Committee; PI, principal investigator; TM, trial manager.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

If it is decided that the AE is a SAR, then it must be determined whether or not the reaction is expected. If it is unexpected and thus considered a SUSAR, then the SAE form will be completed and reported to the DMEC and MREC. Serious but expected reactions or non-SARs will be recorded in the database and reported in the usual way.

For fatal or life-threatening SUSARs the DMEC and MREC will be notified as soon as possible but no later than 7 calendar days after the sponsor has first knowledge of the event. In each case, relevant follow-up information will be sought and a report completed as soon as possible. The follow-up information will be sent to the DMEC and the MREC within an additional 8 calendar days.

Non-fatal and non-life-threatening SUSARs will be reported to the DMEC and the MREC as soon as possible but no later than 15 calendar days after the sponsor has first knowledge of the event. Further relevant follow-up information will be given as soon as possible.

Appendix 3 Radically open dialectical behaviour therapy skills group lesson plan

TABLE 30 Radically open dialectical behaviour therapy skills group lesson plan

Week	Lesson contents	Source
Mindfulness (2 weeks)		
1	Mindfulness handouts 1A, 1B, 1C, 1D and 1E	RO DBT skills (Lynch[69])
2	Mindfulness handouts 2 and 3	Standard DBT skills (Linehan[70])
Emotion regulation (6 weeks)		
3	Emotion regulation handout 1, 2, 3 and homework sheet 1	Standard DBT skills (Linehan[70])
4	Emotion regulation handout 4, 5 and homework sheet 2	
5	Emotion regulation handout 6, 7, 8 and homework sheet 3	
6	Emotion regulation handout 9 and homework sheet 3	
7	Emotion regulation handout 10 and homework sheet 3	
8	Envy, resentment, revenge, bitterness worksheets 1A and 1B	RO DBT skills (Lynch[69])
Mindfulness (2 weeks)		
9	Mindfulness handouts 1A, 1B, 1C, 1D and 1E	RO DBT skills (Lynch[69])
10	Mindfulness handouts 2 and 3	Standard DBT skills (Linehan[70])
Radical openness (8 weeks)		
11	Radical openness handouts 1A, 1B, 1C, 1D and 1E	RO DBT skills (Lynch[69])
12	Radical openness handouts 2A, 2B and 2C	
13	Radical openness worksheet 3A, 3B and 3C	
14	Radical openness handout 4A, 4B and 4C	
15	Radical openness handout 5A and 5B	
16	Radical openness handout 6A, 6B, 6C, 6D and 6E	
17	Radical openness handout 7A and 7B	
18	Radical openness handout 8A and 8B	
Distress tolerance (1 week)		
19	Distress tolerance handout 1 (only page 167 'Self-Soothe') and 2, 3, 4 and 5 (only p. 177)	Standard DBT skills (Linehan[70])
Mindfulness (2 weeks)		
20	Mindfulness handouts 1A, 1B, 1C, 1D and 1E	RO DBT skills (Lynch[69])
21	Mindfulness handouts 2 and 3	Standard DBT skills (Linehan[70])

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 30 Radically open dialectical behaviour therapy skills group lesson plan (*continued*)

Interpersonal effectiveness (6 weeks)		
22	Interpersonal effectiveness handout 1, 2 and homework sheet 1	Standard skills manual (Linehan[70])
23	Interpersonal effectiveness handout 3, 4 and homework sheet 2	
24	Interpersonal effectiveness handout 5 and 6	
25	Interpersonal effectiveness handout 7, 8 and homework sheet 3	
26	Interpersonal effectiveness handout 9	
27	Interpersonal effectiveness handout 10	

Appendix 4 Additional figures

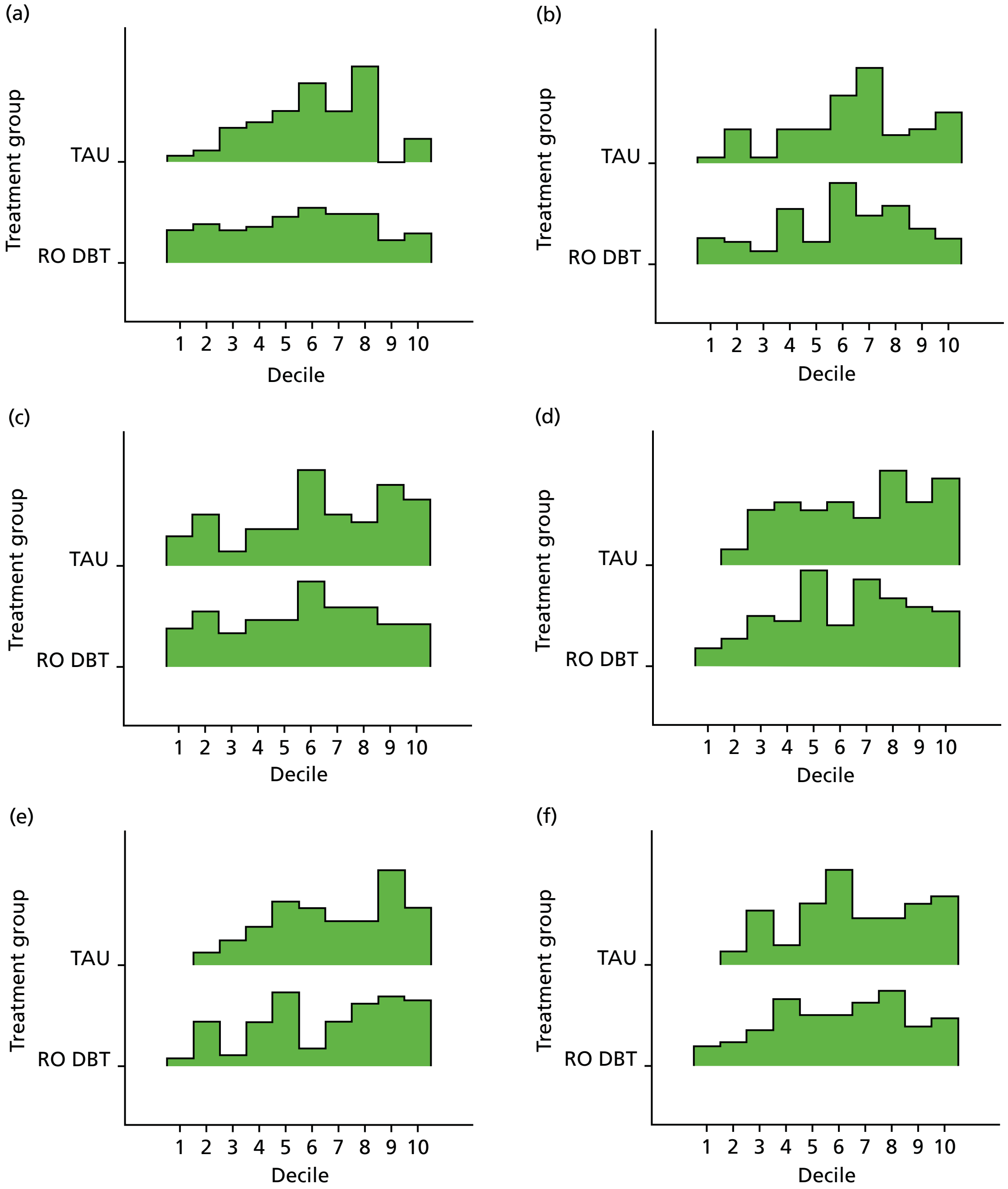

FIGURE 8 Histograms showing the distribution of RefraMED patients across deciles for geographical indices of deprivation, by group. (a) Barriers to housing and services; (b) crime; (c) education and skills; (d) employment; (e) health and disability; (f) income; (g) Index of Multiple Deprivation; (h) living environment; and (i) WIMD 2014 overall. WIMD, Welsh Index of Multiple Deprivation.[149] (*continued*)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

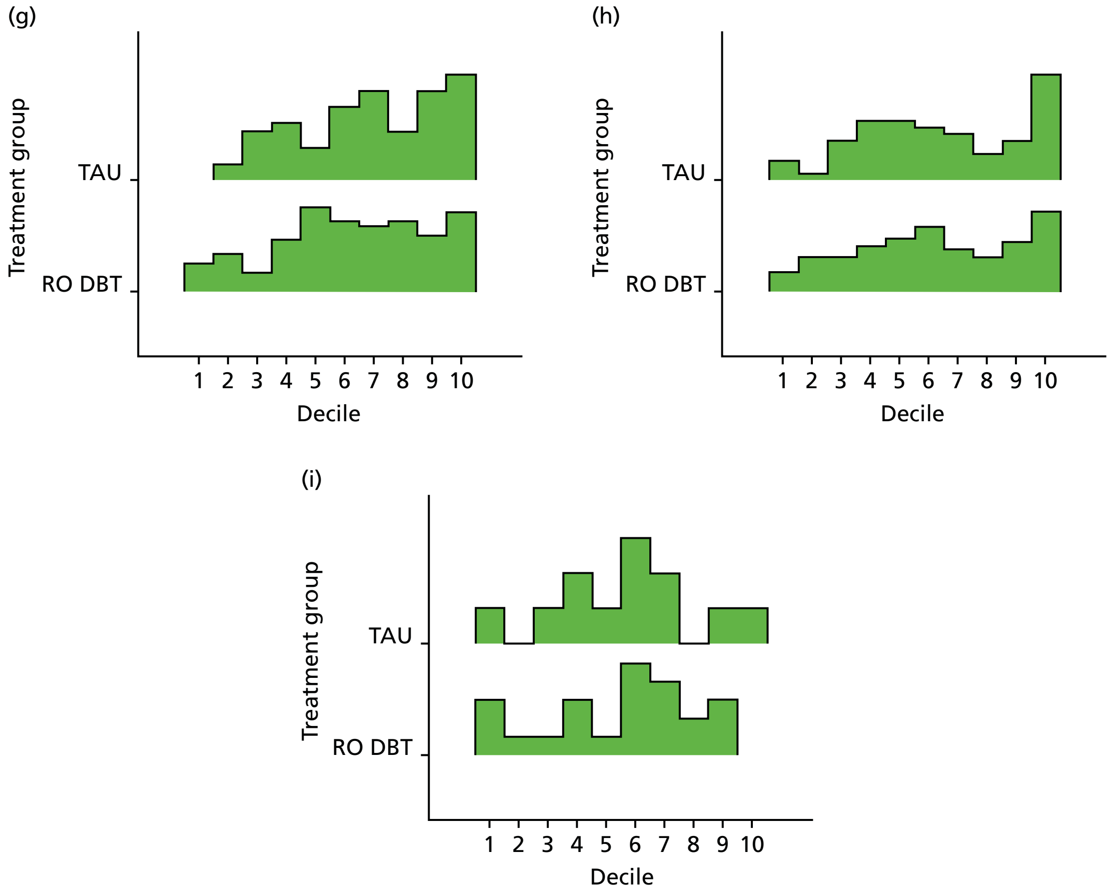

FIGURE 8 Histograms showing the distribution of RefraMED patients across deciles for geographical indices of deprivation, by group. (a) Barriers to housing and services; (b) crime; (c) education and skills; (d) employment; (e) health and disability; (f) income; (g) Index of Multiple Deprivation; (h) living environment; and (i) WIMD 2014 overall. WIMD, Welsh Index of Multiple Deprivation.[149]

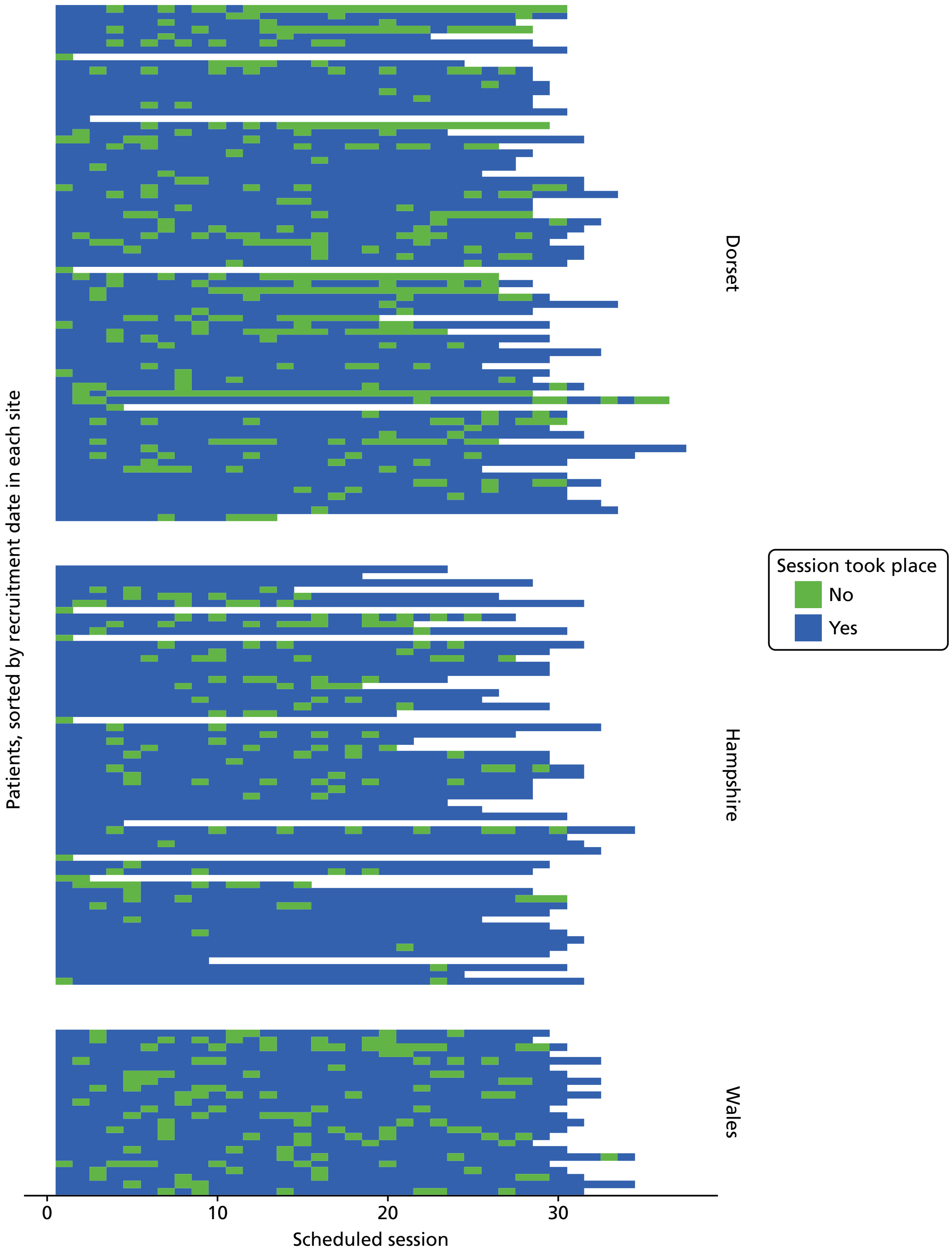

FIGURE 9 The non-monotonic pattern of non-attendance of individual therapy sessions by site.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

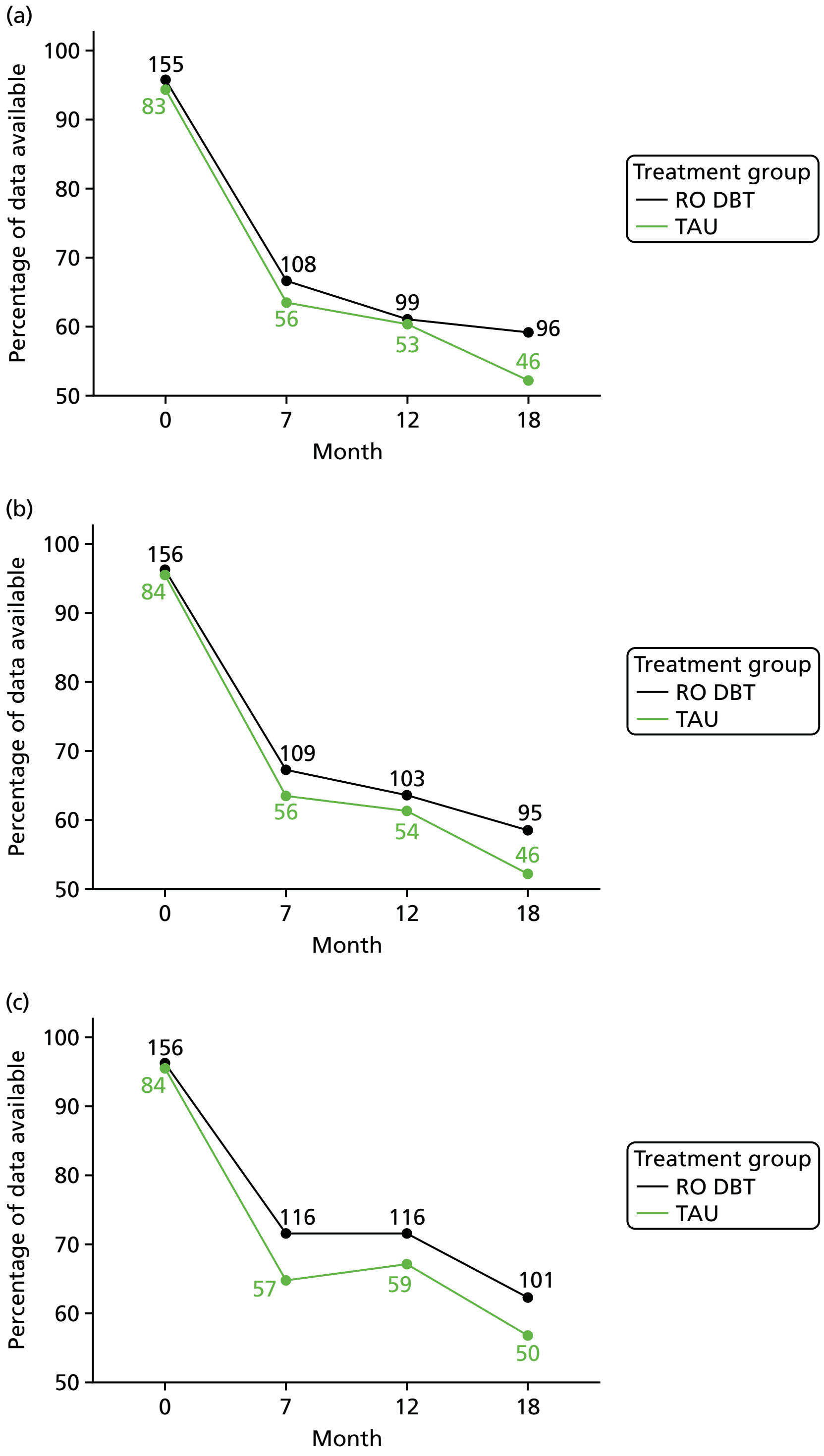

FIGURE 10 The percentage of available continuous data (text labels indicate the number of data points available for analysis). (a) AAQ-II; (b) EAC; (c) EQ-5D-3L; (d) HRSD; (e) LIFE-RIFT; and (f) MSSI. Percentages are calculated based on the full sample of 250. (*continued*)

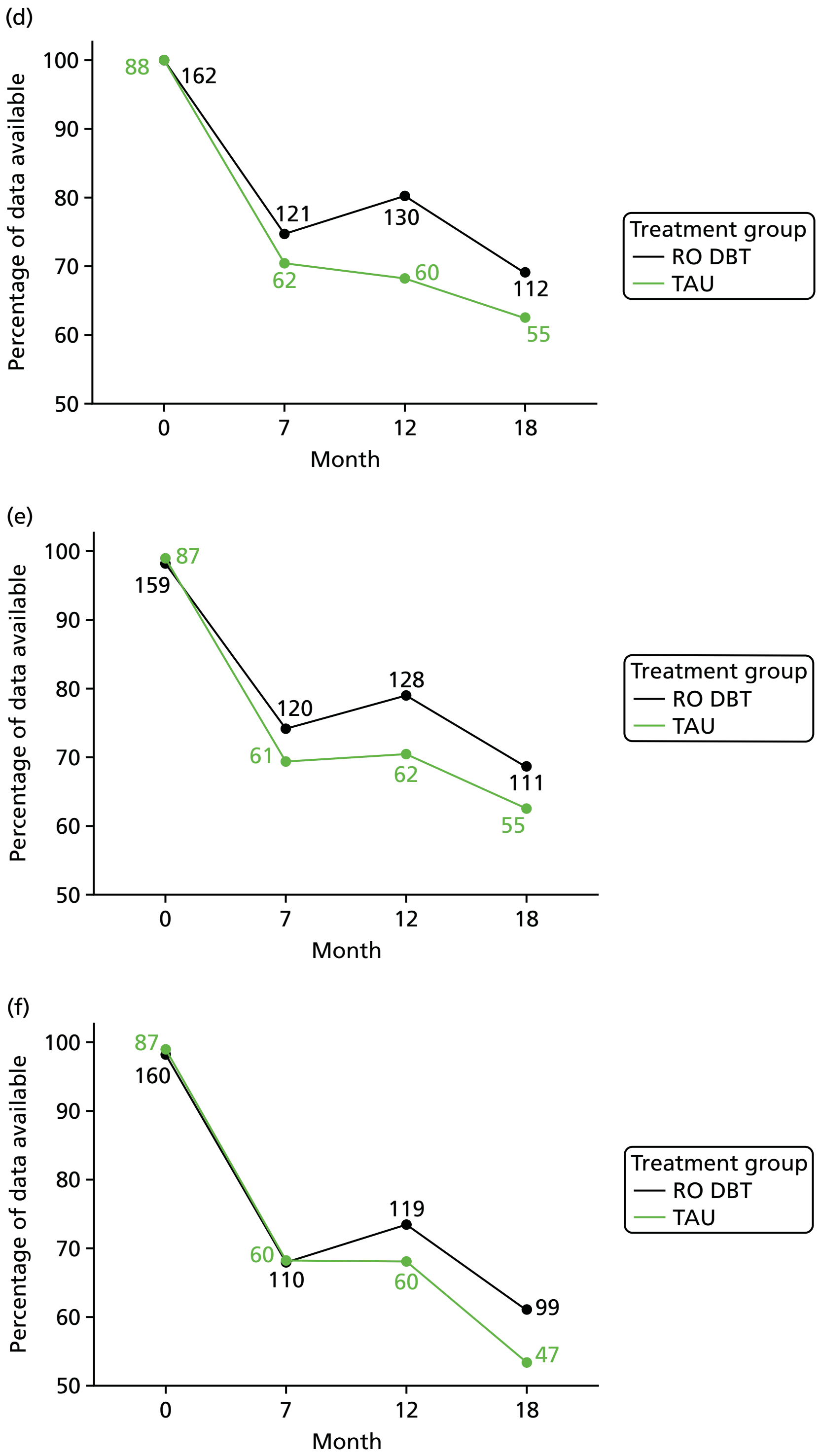

FIGURE 10 The percentage of available continuous data (text labels indicate the number of data points available for analysis). (a) AAQ-II; (b) EAC; (c) EQ-5D-3L; (d) HRSD; (e) LIFE-RIFT; and (f) MSSI. Percentages are calculated based on the full sample of 250.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

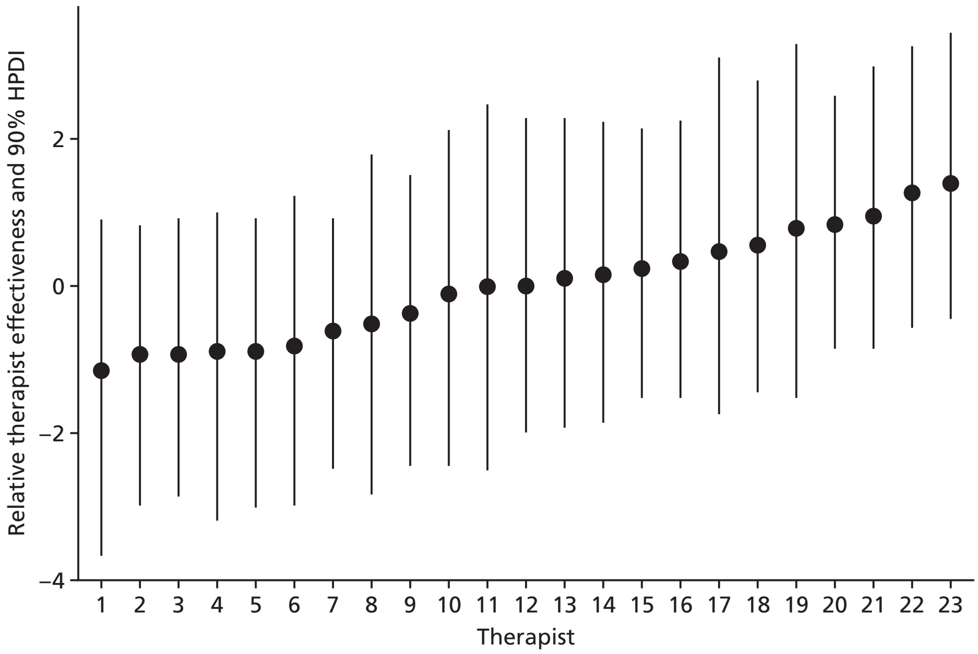

FIGURE 11 The relative differences in therapist effectiveness. HPDI, Highest Probability Density Interval.

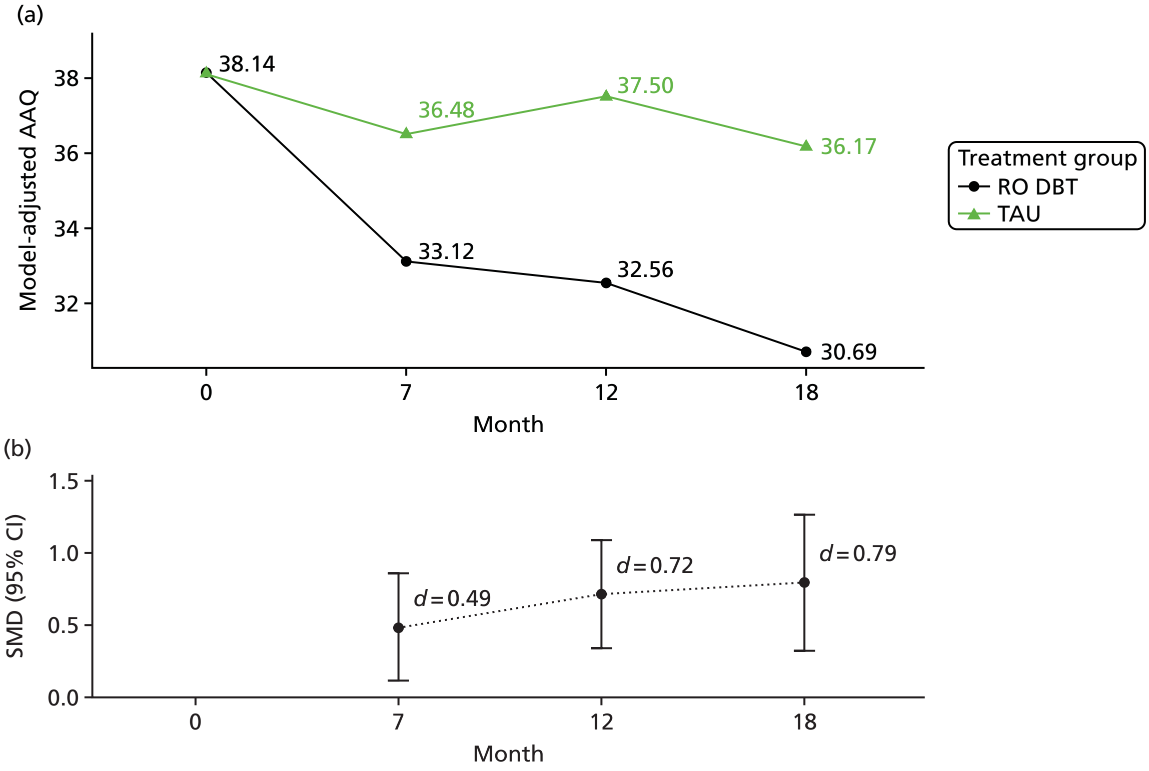

FIGURE 12 The AAQ-II. (a) Adjusted predictions; (b) standardised mean differences between groups; and (c) change from baseline. SMD, standardised mean difference. (*continued*)

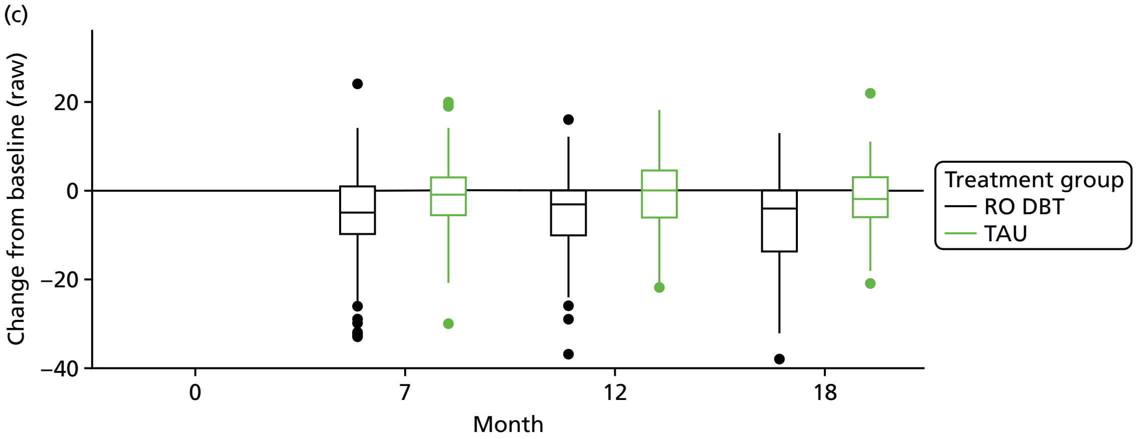

FIGURE 12 The AAQ-II. (a) Adjusted predictions; (b) standardised mean differences between groups; and (c) change from baseline. SMD, standardised mean difference.

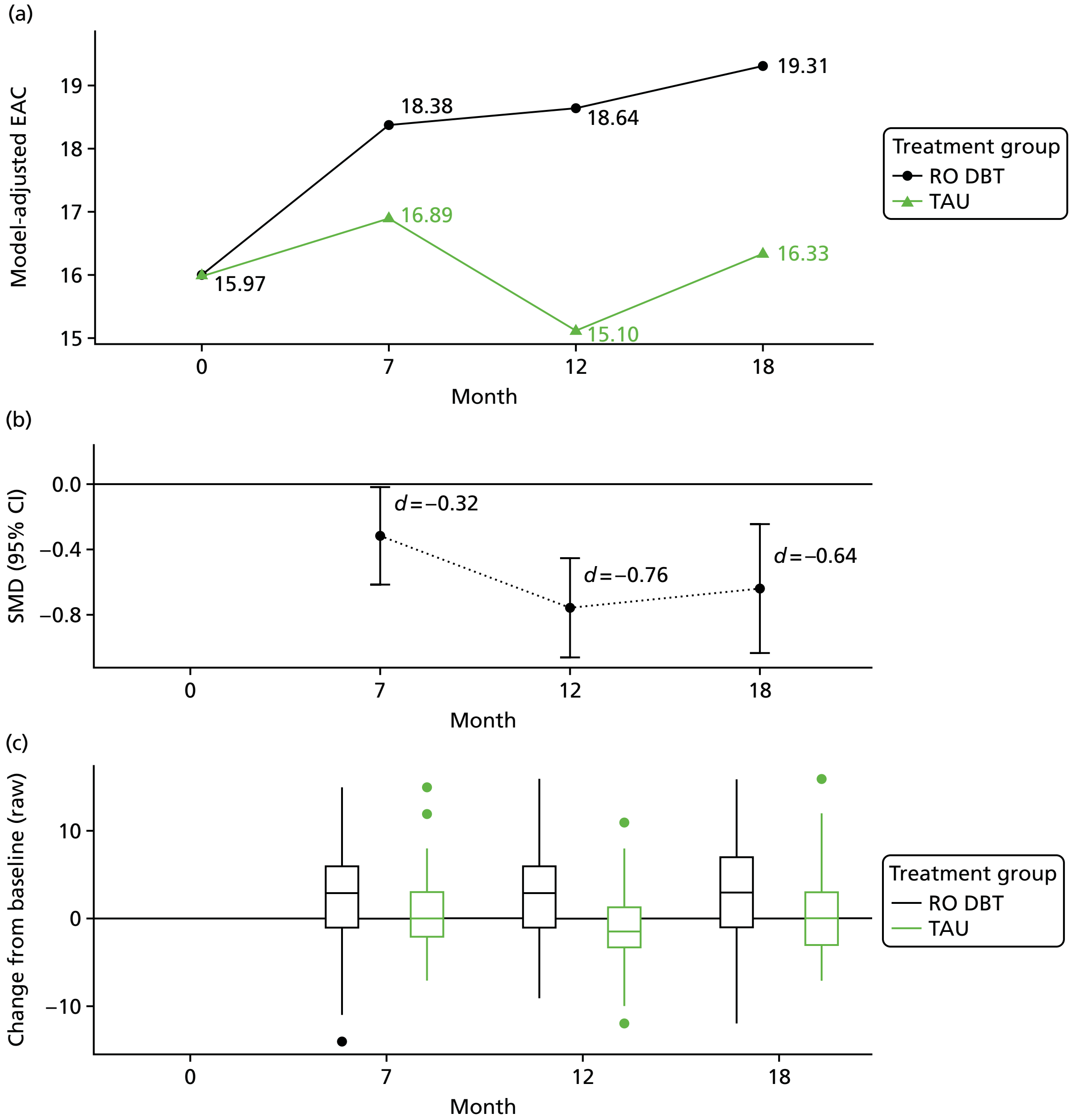

FIGURE 13 The EAC. (a) Adjusted predictions; (b) standardised mean differences between groups; and (c) change from baseline. SMD, standardised mean difference.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

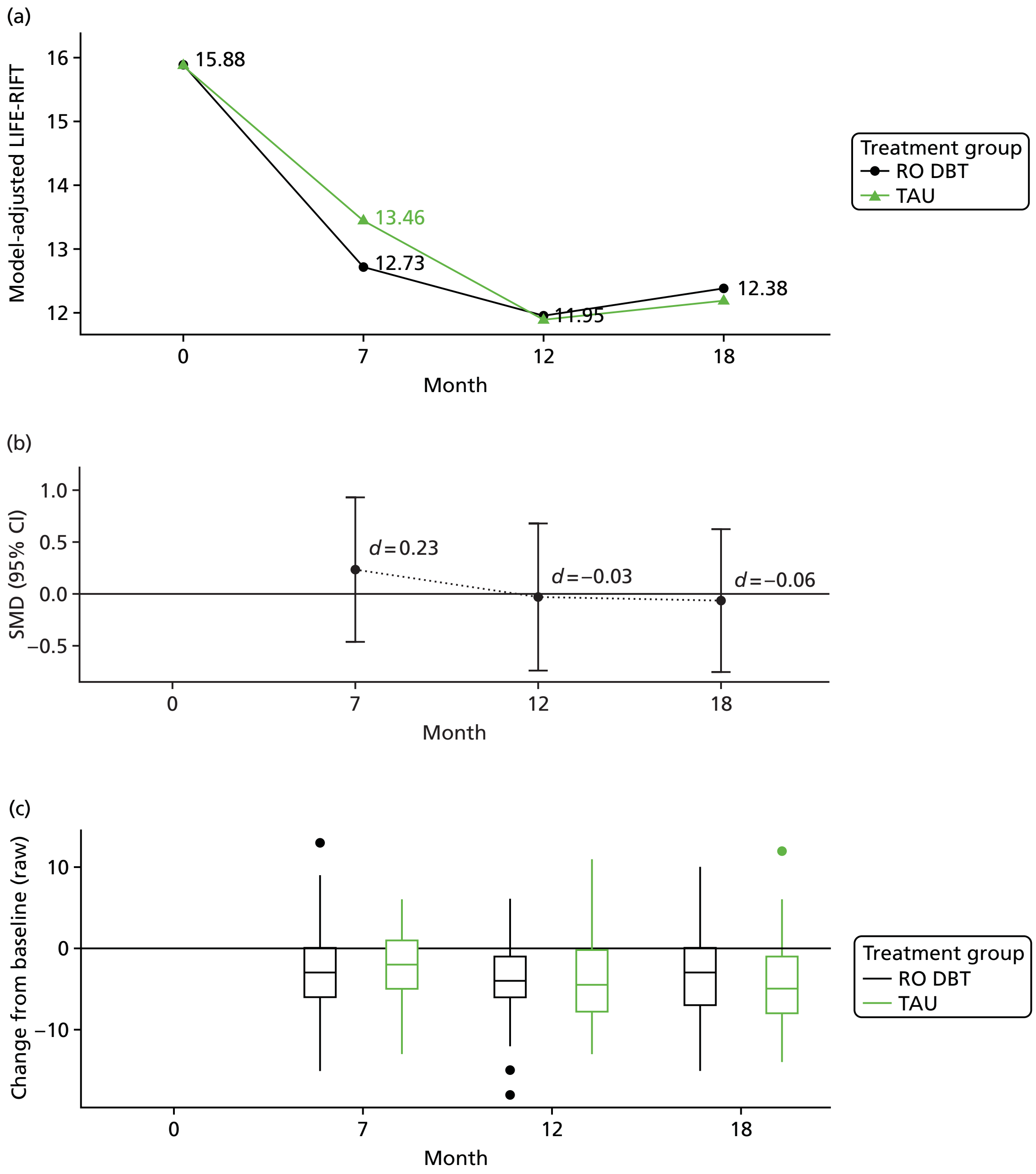

FIGURE 14 The LIFE-RIFT. (a) Adjusted predictions; (b) standardised mean differences between groups; and (c) change from baseline. SMD, standardised mean difference.

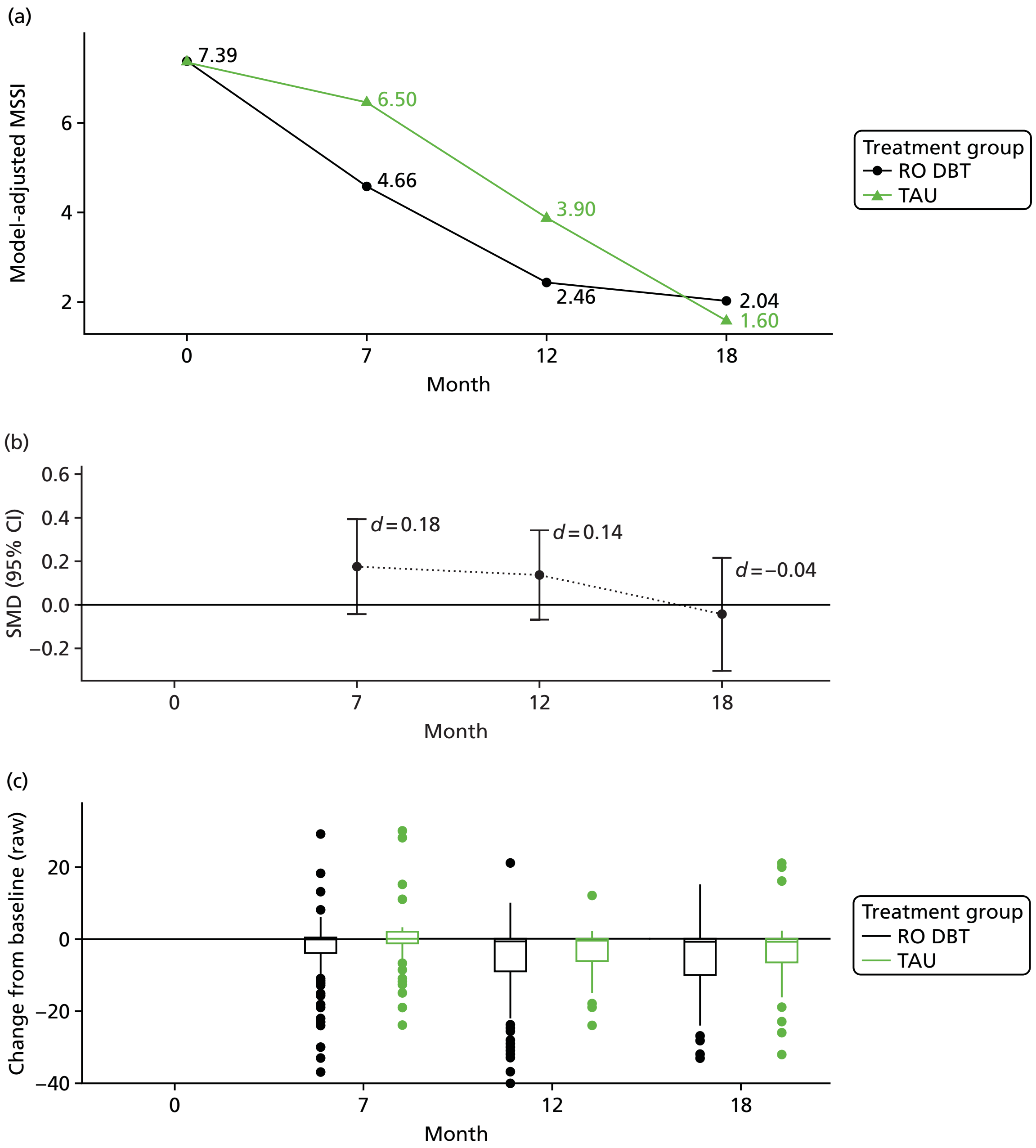

FIGURE 15 The MSSI. (a) Adjusted predictions; (b) standardised mean differences between groups; and (c) change from baseline. SMD, standardised mean difference.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

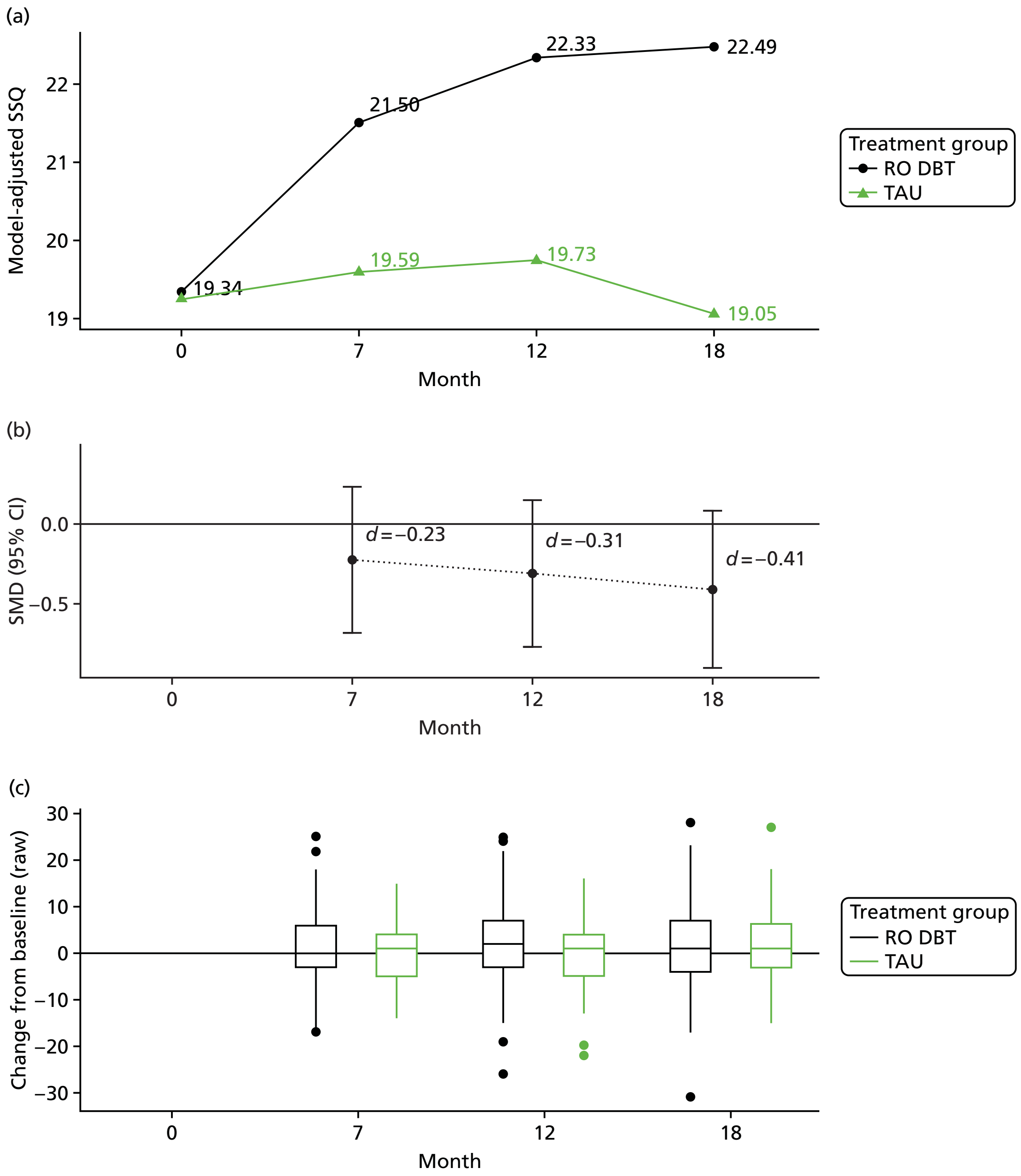

FIGURE 16 The SSQ. (a) Adjusted predictions; (b) standardised mean differences between groups; and (c) change from baseline. SMD, standardised mean difference.

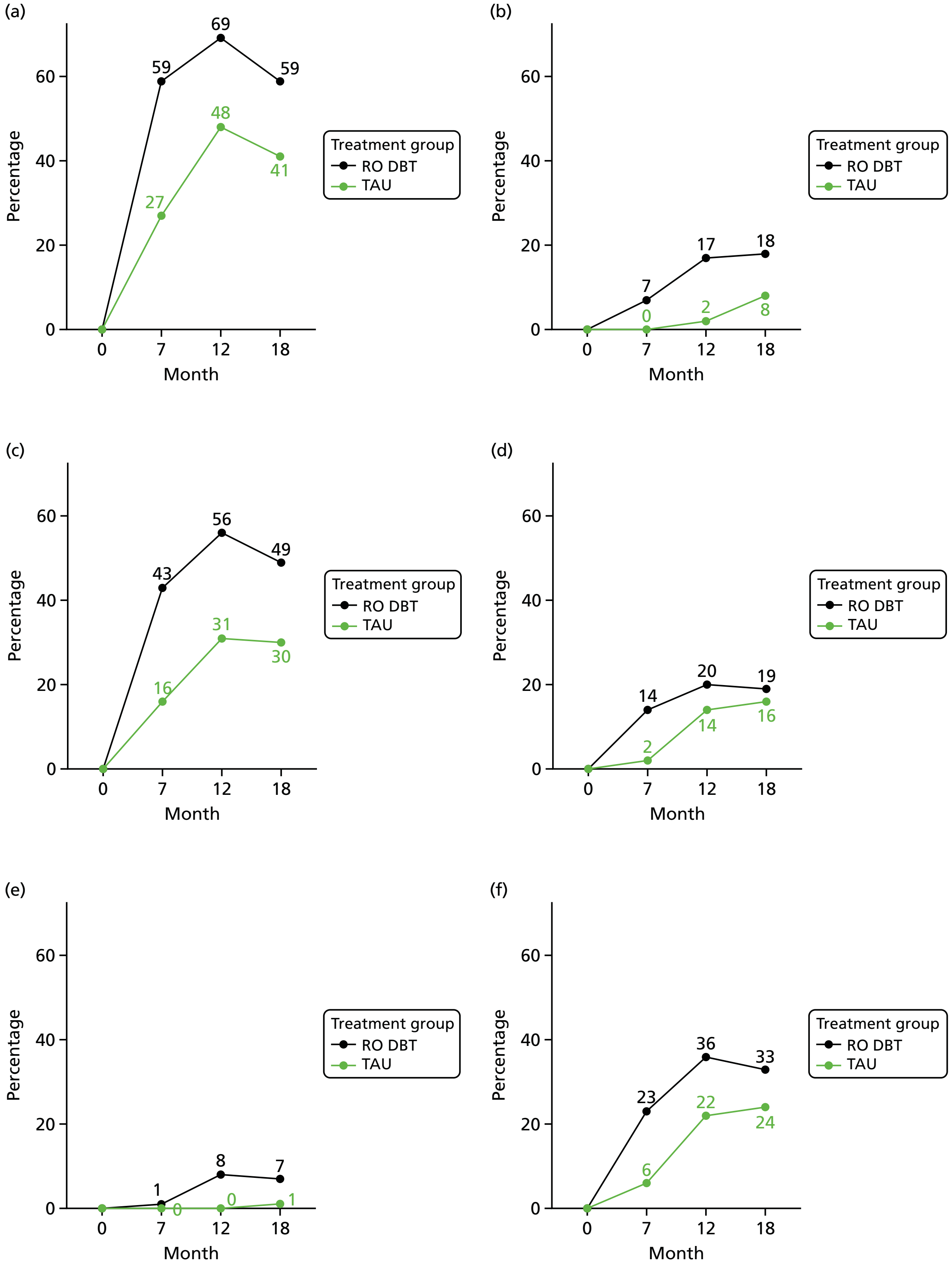

FIGURE 17 Adjusted (a–f) and raw (g–l) data remission rates by group. Presented are the percentages of patients who achieved a 17.5% 'worthwhile' reduction in HRSD score (a and g), a 50% reduction in HRSD score (b and h), a reliable change (Reliable Change Index) (c and i), a reduction of 2 SDs (d and j), full remission (e and k) and full or partial remission (f and l). (*continued*)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

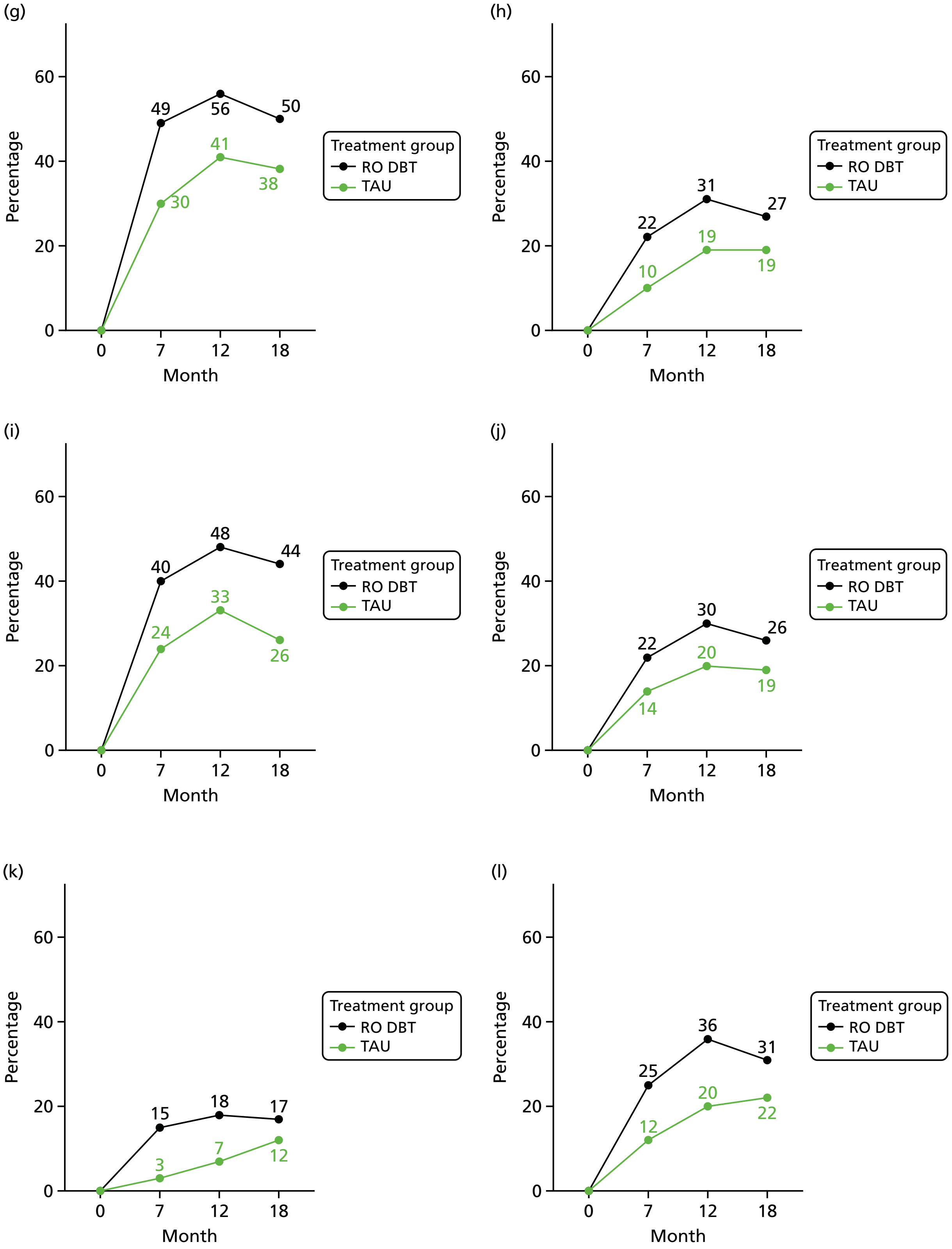

FIGURE 17 Adjusted (a–f) and raw (g–l) data remission rates by group. Presented are the percentages of patients who achieved a 17.5% 'worthwhile' reduction in HRSD score (a and g), a 50% reduction in HRSD score (b and h), a reliable change (Reliable Change Index) (c and i), a reduction of 2 SDs (d and j), full remission (e and k) and full or partial remission (f and l).

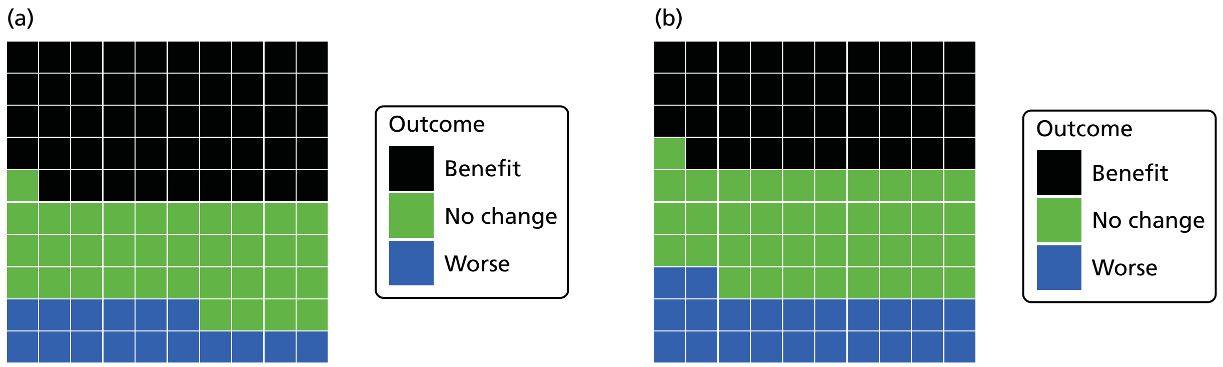

FIGURE 18 The expected outcome for patients choosing RO DBT vs. TAU. (a) Month 7; and (b) month 12. A 'better' outcome indicates that a patient who would have a 17.5% greater reduction in symptoms when choosing RO DBT than they would when choosing TAU.

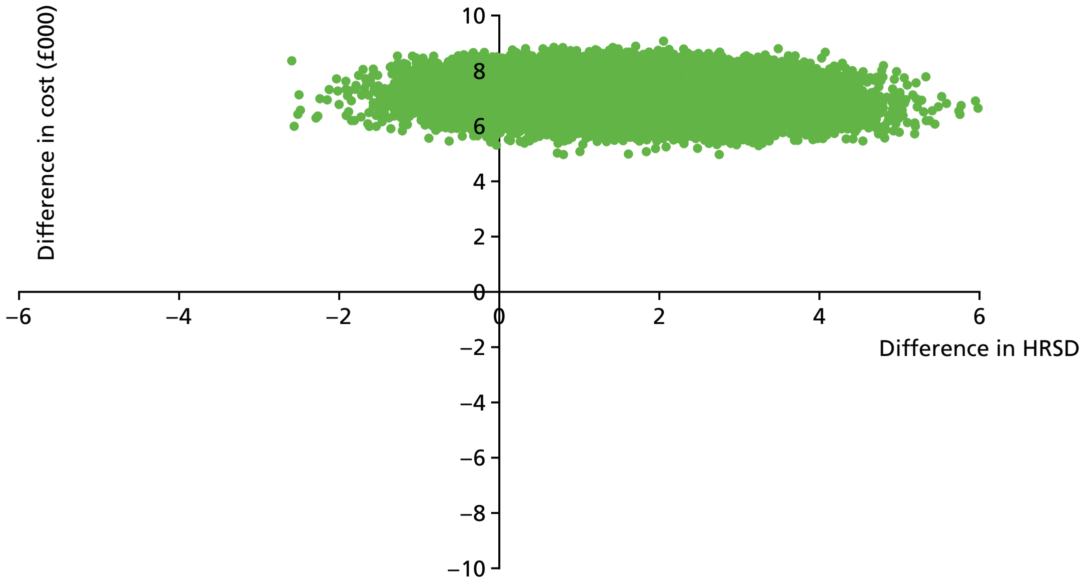

FIGURE 19 A scatterplot of differences in costs vs. differences in HRSD scores for RO DBT vs. TAU at 12 months' follow-up taking the NHS/personal social services perspective.

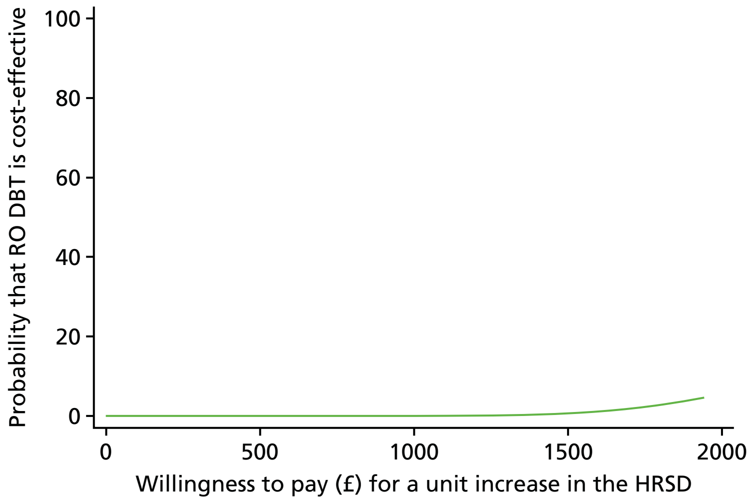

FIGURE 20 Cost-effectiveness acceptability curve showing the probability that RO DBT is cost-effective compared with TAU for different values that a decision-maker might be willing to pay for a point of improvement in HRSD scores.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Lynch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care

Published by the NIHR Journals Library