4508 words

4 tables

1 figure

Dismantling, optimising and personalising internet cognitive-behavioural therapy for depression: A systematic review and individual participant data component network meta-analysis

|  |  |  |
| --- | --- | --- |
| Toshi A. Furukawa, MD PhD\*† | Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health, Kyoto Japan | furukawa@kuhp.kyoto-u.ac.jp |
| Aya Suganuma, Mchem, MD\* | Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine / School of Public Health, Kyoto Japan | chatblanc2011@gmail.com |
| Edoardo G. Ostinelli, MD\* | Department of Psychiatry, Warneford Hospital, University of Oxford, Oxford, United Kingdom | eg.ostinelli@gmail.com |
| Gerhard Andersson, PhD | Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institutet and Stockholm Health Care Services, Stockholm County Council, Sweden  Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden | gerhard.andersson@liu.se |
| Christopher G. Beevers, PhD | Department of Psychology and Institute for Mental Health Research, University of Texas At Austin, Austin, USA | beevers@utexas.edu |
| Jason Shumake, PhD | Department of Psychology and Institute for Mental Health Research, University of Texas At Austin, Austin, USA | shumake@utexas.edu |
| Thomas Berger, PhD | Department of Clinical Psychology and Psychotherapy, University of Bern, Bern, Switzerland | thomas.berger@ptp.unibe.ch |
| Florien Willemijn Boele, PhD | Patient Centred Outcomes Group, Leeds Institute of Medical Research at St James’s, University of Leeds, United Kingdom | F.Boele@leeds.ac.uk |
| Claudia Buntrock, PhD | Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander-University Erlangen-Nuremberg, Germany | claudia.buntrock@fau.de |
| Per Carlbring, PhD | Department of Psychology, Stockholm University, Stockholm, Sweden | per@carlbring.se |
| Isabella Choi, PhD | Central Clinical School, Brain and Mind Centre, Faculty of Medicine and Health, University of Sydney, Sydney, Australia | isabella.choi@sydney.edu.au |
| Heleen Christensen, PhD | Black Dog Institute and University of New South Wales, Prince of Wales Hospital, Sydney, Australia | h.christensen@blackdog.org.au |
| Andrew Mackinnon PhD | Black Dog Institute and University of New South Wales, Prince of Wales Hospital, Sydney, Australia | andrew.mackinnon@biostats.com.au |
| Jennifer Dahne, PhD | Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina (MUSC), Charleston, South Carolina, USA | dahne@musc.edu |
| Marcus J.H. Huibers | Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit Amsterdam | m.j.h.huibers@vu.nl |
| David D. Ebert, PhD | Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany | david.ebert@fau.de |
| Louise Farrer, PhD | Centre for Mental Health Research, The Australian National University, Canberra, Australia | Louise.Farrer@anu.edu.au |
| Nicholas R. Forand, PhD | Department of Psychiatry, The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, New York, USA | nforand@northwell.edu |
| Daniel R. Strunk, PhD | Department of Psychology, The Ohio State University, Columbus, OH, USA | strunk.20@osu.edu; |
| Iony D. Ezawa, PhD | Department of Psychology, The Ohio State University, Columbus, OH, USA | ezawa.6@buckeyemail.osu.edu |
| Erik Forsell, PhD | Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, & Stockholm Health Care Services, Region Stockholm, Sweden. | erik.forsell@ki.se |
| Viktor Kaldo,  PhD | Department of Psychology, Faculty of Health and Life Sciences, Linnaeus University, Växjö, Sweden.  Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, & Stockholm Health Care Services, Region Stockholm, Sweden | viktor.kaldo@ki.se |
| Anna Geraedts, PhD | Soulve Innovations, Utrecht, The Netherlands | asgeraedts@icloud.com |
| Simon Gilbody, PhD | Department of Health Sciences, University of York, York, UK | simon.gilbody@york.ac.uk |
| Elizabeth Littlewood, PhD | Department of Health Sciences, University of York, York, UK | liz.littlewood@york.ac.uk |
| Sally Brabyn, MSc | Department of Health Sciences, University of York, York, UK | sally.brabyn@york.ac.uk |
| Heather D. Hadjistavropoulos, PhD | Department of Psychology, University of Regina, Regina, Saskatchewan, Canada | Heather.Hadjistavropoulos@uregina.ca |
| Luke H. Schneider, PhD | Anxiety Treatment and Research Clinic, St. Joseph’s Healthcare Hamilton, Hamilton, Ontario, Canada | Luke.Schneider@uregina.ca |
| Robert Johansson, PhD | Department of Psychology, Stockholm University, Stockholm, Sweden | robert.johansson@psychology.su.se |
| Robin Kenter, PhD | Department of Clinical Psychology, Faculty of Psychology, University of Bergen | Robin.Kenter@uib.no |
| Marie Kivi, PhD | Department of Psychology, University of Gothenburg, Gothenburg, Sweden | marie.kivi@psy.gu.se |
| Cecilia Björkelund, MD, PhD | Primary Health Care, School of Public Health and Community Medicine, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden | cecilia.bjorkelund@allmed.gu.se |
| Annet Kleiboer, PhD | Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, Netherlands | a.m.kleiboer@vu.nl |
| Heleen Riper, PhD | Department of Clinical, Neuro- and Developmental Psychology, VU University, Amsterdam, The Netherlands. | h.riper@vu.nl |
| Jan Philipp Klein, MD | Department of Psychiatry and Psychotherapy, Luebeck University, Luebeck, Germany | Philipp.Klein@uksh.de |
| Johanna Schröder, PhD | Institute for Sex Research, Sexual Medicine and Forensic Psychiatry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany | jo.schroeder@uke.de |
| Björn Meyer, PhD | Research Department, GAIA AG , Hamburg, Germany | bjoern.meyer@gaia-group.com |
| Steffen Moritz | Department of Psychiatry and Psychotherapy, University Medical Center Hamburg, Hamburg, Germany | moritz@uke.de |
| Lara Bücker | Department of Psychiatry and Psychotherapy, University Medical Center Hamburg, Hamburg, Germany | lbuecker@uke.de |
| Ove Lintvedt, PhD | Norwegian Center for E-health research, Tromsø, Norway | Ove.Lintvedt@unn.no |
| Johan Lundgren, RN PhD | Department om Health, Medicine and Caring Sciences, Linköping University, Norrköping, Sweden | johan.lundgren@liu.se |
| Jeannette Milgrom, PhD | Parent-Infant Research Institute and Austin Health,  Melbourne School of Psychological Sciences, University of Melbourne, Victoria, Australia | jeannette.milgrom@austin.org.au |
| Alan W. Gemmill, PhD | Parent-Infant Research Institute, Department of Clinical and Health Psychology, Austin Health, Victoria, Australia | alan.gemmill@austin.org.au |
| David C. Mohr, PhD | Center for Behavioral Intervention Technologies, Department of Preventive Medicine, Northwestern University, Chicago, IL, USA | d-mohr@northwestern.edu |
| Jesus Montero-Marin, PhD | Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK | jesus.monteromarin@psych.ox.ac.uk |
| Javier Garcia- Campayo, MD, PhD | Aragon Institute for Health Research (IIS Aragón), Miguel Servet University Hospital, 50009 Zaragoza, Spain;  Primary Care Prevention and Health Promotion Research Network, RedIAPP, 28029 Madrid, Spain | jgarcamp@gmail.com |
| Stephanie Nobis | GET.ON Institut, Hamburg, Germany | s.nobis@geton-institut.de |
| Anna-Carlotta Zarski, PhD | Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander University Erlangen-Nürnberg | anna-carlotta.zarski@fau.de |
| Kathleen O'Moore, DCP/MSc | Black Dog Institute, University of New South Wales, Sydney, NSW, Australia | k.omoore@blackdog.org.au |
| Alishia D. Williams, PhD | Department of Psychology, Faculty of Science, The University of New South Wales, Sydney, Australia | alisha.williams@unsw.edu.au |
| Jill M. Newby, PhD | School of Psychology, University of New South Wales at the Black Dog Institute, Sydney, Australia | j.newby@unsw.edu.au |
| Sarah Perini, PhD | Clinical Research Unit for Anxiety and Depression, St Vincent’s Hospital, Sydney, Australia | sarahperini@yahoo.com |
| Rachel Phillips, MSc | Faculty of Medicine, School of Public Health, Imperial College London, London, UK | r.phillips@imperial.ac.uk |
| Justine Schneider, PhD | School of Sociology & Social Policy and Institute of Mental Health University of Nottingham, UK | Justine.Schneider@nottingham.ac.uk |
| Wendy Pots, LCP, PhD | Department of Psychology, Health & Technology, University of Twente, Enschede, The Netherlands | w.t.m.pots@utwente.nl |
| Nicole E Pugh, PhD | Private practice, Vancouver, Canada | Nicky.e.pugh@gmail.com |
| Derek Richards, PhD | University of Dublin, Trinity College, School of Psychology, E-mental Health Research Group, Dublin, Ireland  SilverCloud Health, Clinical Research & Innovation, Dublin, Ireland | DRICHARD@tcd.ie |
| Isabelle M. Rosso, PhD | McLean Hospital, Belmont, Massachusetts, USA | isabelle\_rosso@hms.harvard.edu |
| Scott L. Rauch, MD | McLean Hospital, Belmont, Massachusetts, USA | srauch@partners.org |
| Lisa B. Sheeber, PhD | Oregon Research Institute, Eugene, OR, USA | lsheeber@ori.org |
| Jessica Smith, MSc | Imperial Clinical Trials Unit, Imperial College London, United Kingdom | jessicasmith1041@gmail.com |
| Viola Spek, PhD | School of Applied Psychology, Fontys University of Applied Science, Eindhoven, The Netherlands | v.spek@fontys.nl |
| Viktor J. Pop, MD, PhD | Department of Medical & Clinical Psychology, Tilburg University, The Netherlands | v.j.m.pop@uvt.nl |
| Burçin Ünlü , PhD | PsyQ Online, Haarlem, The Netherlands | unlu\_burcin@outlook.com |
| Kim M. P. van Bastelaar, PhD | Amsterdam University Medical Centre, Amsterdam, The Netherlands | kmpvanbastelaar@gmail.com |
| Sanne van Luenen, PhD | Department of Clinical Psychology, Leiden University, Leiden, The Netherlands | s.van.luenen@fsw.leidenuniv.nl; |
| Nadia Garnefski, PhD | Department of Clinical Psychology, Leiden University, The Netherlands | garnefski@fsw.leidenuniv.nl |
| Vivian Kraaij, PhD | Department of Clinical Psychology, Leiden University, Leiden, The Netherlands | kraaij@fsw.leidenuniv.nl |
| Kristofer Vernmark, PhD | Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden  Psykologpartners, Sweden | kristofer.vernmark@psykologpartners.se |
| Lisanne Warmerdam | National Health Care Institute, The Netherlands | lisannew@hotmail.com |
| Annemieke van Straten, PhD | Vrije Universiteit Amsterdam  Dept. of Clinical- Neuro- and Developmental psychology & Amsterdam Public Health Research Institute | a.van.straten@vu.nl |
| Pavle Zagorscak, PhD | Department for Clinical Psychological Intervention, Freie Universität Berlin, Berlin, Germany | zagorscak@zedat.fu-berlin.de |
| Christine Knaevelsrud, PhD | Department for Clinical Psychological Intervention, Freie Universität Berlin, Berlin, Germany | christine.knaevelsrud@fu-berlin.de |
| Manuel Heinrich, MSc | Department for Clinical Psychological Intervention, Freie Universität Berlin, Berlin, Germany | manuel.heinrich@fu-berlin.de |
| Clara Miguel, MSc | Department of Clinical, Neuro- and Developmental Psychology, VU University, Amsterdam, The Netherlands. | c.miguelsanz@student.vu.nl |
| Andrea Cipriani, MD, PhD | Department of Psychiatry, Warneford Hospital, University of Oxford, Oxford, UK  Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK | andrea.cipriani@psych.ox.ac.uk |
| Orestis Efthimiou, PhD‡ | Institute of Social and Preventive Medicine, University of Bern, Switzerland | oremiou@gmail.com |
| Eirini Karyotaki, PhD‡ | Department of Global Health and Social Medicine, Harvard Medical School, Boston, USA | eirini\_karyotaki@hms.harvard.edu |
| Pim Cuijpers, PhD‡ | Department of Clinical, Neuro- and Developmental Psychology, VU University, Amsterdam, The Netherlands. | p.cuijpers@vu.nl |

\* co-first authors

‡ co-last authors

† corresponding author

# Summary

**Background**: Internet cognitive behavioural therapy (iCBT) is a viable delivery format of CBT for depression. However, it includes a wide array of cognitive and behavioural skills in different delivery methods, and it remains unclear which of these components are more efficacious and for whom.

**Methods**: We did a systematic review and individual patient data (IPD) component network meta-analysis (cNMA) of iCBT trials for depression. We searched PubMed, PsycINFO, Embase and the Cochrane Library up to January 2019 for randomised controlled trials (RCTs) that compared any form of iCBT against another or a control condition in the acute treatment of adults with depression. We sought IPD from the original authors. When IPD were unavailable, we used aggregate data (AD). Two independent researchers identified the included components. The primary outcome was depression severity, expressed as incremental mean difference (iMD) in the Patient Health Questionnaire-9 (PHQ-9) scores when a component is added to a treatment. The secondary outcome was dropout from treatment.

**Findings**: We identified 76 RCTs, including 48 trials contributing IPD (11704 participants) and 28 trials with AD (6474 participants). The participants’ mean age was 42.0 years and 71% (12406 out of 17521 reported) were women. There was suggestive evidence that behavioural activation may be beneficial (iMD -1.83, 95% Credible Interval: -2.9 to -0.8), while relaxation may be harmful (1.20, 0.17 to 2.27). Baseline severity emerged as the strongest prognostic factor. Combining human and automated encouragement reduced dropouts from treatment (incremental odds ratio, 0.32, 0.13 to 0.93). We constructed a web app that estimates relative efficacies between any two combinations, given baseline patient characteristics (<https://esm.ispm.unibe.ch/shinies/cNMA_iCBT/>).

**Interpretation**: IPD cNMA revealed potentially helpful, less helpful or harmful components and delivery formats for iCBT packages. The future iCBT packages aiming to be effective and efficient may choose to include beneficial components and drop ones that are potentially detrimental. The web app can facilitate shared decision making by therapist and patient in choosing their preferred iCBT package.

**Funding**: Japan Society for the Promotion of Science

# Research in context

**Evidence before this study:**

Psychotherapy is a complex intervention, comprising multiple components in various combinations. There are two types of research to disentangle specific contributions of psychotherapy components: so-called dismantling studies and, more recently, component network meta-analyses (cNMA). We searched PubMed from inception to October 9, 2020, for relevant studies with the following terms in titles and abstracts: ("Mental Disorders"[Mesh] "Psychotherapy"[Mesh] dismantl\*) and ("Mental Disorders"[Mesh] "Psychotherapy"[Mesh] component network meta-analysis). We identified three systematic reviews of dismantling studies and two component network meta-analyses. Earlier reviews of dismantling studies found no additive effects of specific components among dismantling studies: however, the most recent review focusing on depression found behavioural activation, but not cognitive restructuring or mindfulness, to have significant additive effects. cNMA of face-to-face cognitive behavioural therapy (CBT) for panic disorder singled out muscle relaxation to be harmful: cNMA of CBT for depression found no evidence of specific effects of any components or their combinations.

**Added value of this study:**

We did a systematic review and individual participant data (IPD) cNMA of internet CBT (iCBT) for depression. There were three new features that strengthened its precision, sensitivity and clinical relevance. First, cNMA increased precision of estimates by including more studies than narrowly defined dismantling studies and by combining direct and indirect estimates while maintaining randomized comparisons. Second, by focusing on iCBT, we could be certain that the components had been offered and made accessible as intended by the researchers, in contrast to face-to-face CBT whose contents may have been modified or skipped when its conduct was not well monitored. Finally use of IPD allowed us to identify prognostic factors and effect modifiers and thus estimate personalised relative effects among different interventions, depending on individual patients’ characteristics. We found suggestive evidence that non-specific treatment effects (including placebo effect and common factors) and behavioural activation may have beneficial effects, while relaxation might be detrimental. There was weaker evidence that behaviour therapy for insomnia may also be helpful. The evidence was inconclusive for some traditional elements of CBT such as cognitive restructuring or homework. Combining human encouragement to proceed with the iCBT program with automated encouragement decreased dropouts from treatment. The web app can estimate relative efficacies between any two combinations based on the baseline patient characteristics (<https://esm.ispm.unibe.ch/shinies/cNMA_iCBT/>).

**Implications of all the available evidence:**

The future iCBT packages aiming to be effective and efficient may choose to include behavioural activation but drop relaxation. They may further choose to include behaviour therapy for insomnia and problem solving, but probably not cognitive restructuring. They may employ automated encouragement and human encouragement to increase adherence. The web app can facilitate shared decision making by therapist and patient in choosing their preferred iCBT package.

# INTRODUCTION

Cognitive-behavioural therapy (CBT) is the most widely studied type of psychotherapy for depression 1,2. CBT encompasses a wide array of cognitive and behavioural skills, which are sometimes administered alone but more commonly in various combinations. Moreover, they are generally administered in a flexible manner, because it is believed that their efficacy is moderated by individual patients’ characteristics. It is then of the utmost importance to know which of these components are more contributory to its effectiveness, which combinations of them are optimal, and for whom 3.

Traditional approaches to examine components involved so-called dismantling studies, in which the whole package is compared against a package that omits one component4,5. However, such studies have been typically underpowered and of poor methodological quality. Moreover, they have proved difficult to synthesise as each study examined very diverse components and covered heterogeneous conditions 6. With new advances in the science of evidence synthesis, however, we can now dismantle complex interventions through component network meta-analysis (cNMA) by estimating the individual efficacies of the various components contained in a network of randomised controlled trials (RCTs) 7,8. cNMA increases statistical power by combining direct and indirect comparisons while fully respecting the randomised structure of the evidence, i.e. treatment effects are estimated separately in each study, and then study-specific estimates are synthesised on a second level 9.

One fundamental limitation of all previous dismantling studies and cNMAs, however, is that they dealt mainly with face-to-face CBT, in which it is difficult to be sure that the claimed components have actually been administered as intended and that no other elements were introduced, unless treatment fidelity was monitored. In the past two decades computerised or internet CBT (iCBT) has been introduced and widely tested in trials. It is now well established that guided iCBT can be as effective as face-to-face individual, group or other delivery formats of CBT of similar length 10-12. iCBT provides a unique platform where each cognitive and behavioural skill is offered uniformly and as intended by the program developers. Moreover iCBT brings with it new clinical questions regarding delivery methods to improve its adherence.

In this study we first conducted the NMA at the treatment level to examine the suitability of the available evidence for NMA, then conducted the cNMA. For cNMA we included not only the aggregate data (AD) of trial summary statistics but also individual participant data (IPD). IPD allow us to examine prognostic factors (baseline characteristics which predict the outcome regardless of the intervention) and effect modifiers (those which predict differential response to treatments) and to estimate individual relative treatment effects based on these factors. The aim of this IPD cNMA was, therefore, to elucidate which of the skills and delivery methods commonly included in the broadly conceived iCBT packages are efficacious and for whom.

# METHODS

This study is registered in PROSPERO (CRD42018104683) and the protocol has been published 13, This report follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension statements for network meta-analysis 14 and for individual participant data meta-analysis 15. The PRISMA-IPD checklist is provided in the appendix (pp 84-87).

## Search strategy, selection criteria and data extraction

We used an existing database of psychological treatments for depression which is updated annually through comprehensive literature searches in PubMed, PsycINFO, Embase and the Cochrane Library 16. Searches combine terms for depression and psychotherapies with filters for randomised controlled trials (See the appendix (pp 4-8) for full search strings). Two researchers independently checked this database for eligible studies; disagreements were resolved by discussion. The last update search was conducted in January 2019.

Table 1. Components, their definitions and the number of trial arms including each component

|  |  |  |
| --- | --- | --- |
| **Components (abbreviation)** | **Definition** | **No of arms** |
| Waiting component (*w*) | Participants know that they can receive an active treatment, after a waiting phase. Usually patients on a wait list do not receive any treatment during the waiting phase. However, in some trials patients allocated to waiting list received some non-specific therapeutic components, such as psychological placebo, psychoeducation or treatment as usual. In such cases, we assumed that the “waiting component” (*w*) was present, recorded the interventions provided while waiting, and classified such an arm as WL. | 52 |
| Conventional drug treatment (*dt*) | Treatment as usual or care as usual can denote many different conditions in the literature 17,18. In this study we focused on the use of “conventional drug treatment” and extracted the data if it was present (drug treatment was part of the protocol treatment), allowed or absent. When the drug was used to the same extent in both arms, the component *dt* was assumed present. | 143 |
| Non-specific treatment effects (ns) | Effects of an intervention due to the patients’ belief that they are receiving some form of treatment (placebo effect) and to the common or non-specific factors of psychotherapies. These two elements were indistinguishable in the current network of psychotherapies. | 142 |
| Psychoeducation about depression (*pe*) | Provision of information about the cause and nature of depression. Patients are taught their symptoms can be interpreted under a certain psychopathological model. For example, if cognitive distortion is cited as the cause of depression, such explanation counted towards *pe* as defined here. We considered psychoeducation present only if there was a dedicated module (psychoeducation or introductory).  Advice about lifestyle modification (e.g. exercise, food, sleep hygiene (as opposed to CBT for insomnia)) or provision of information about depression in informational websites were regarded as a form of psychoeducation. | 111 |
| Cognitive restructuring (*cr*) | This component teaches the patient to evaluate and modify their own irrational, maladaptive or dysfunctional thoughts using strategies such as Socratic questioning and guided imagery. | 74 |
| Behavioural activation (*ba*) | This component aims at helping people increase potentially reinforcing experiences through activity scheduling and increased engagement in pleasant activities. | 84 |
| Interpersonal skills training (*is*) | Training in appropriate social behaviours. Includes assertiveness training which teaches the patient to stand up to their own rights by expressing their feelings and wishes in an honest and respectful manner that does not insult or hurt the others. | 31 |
| Problem-solving (*ps*) | This skill includes the following step-by-step approach to personal problems: defining personal problems, generating multiple solutions, selecting the best solution, working out a systematic plan for this solution, and evaluating whether the solution has resolved the problem. | 55 |
| Relaxation (*re*) | This skill is aimed at reducing general tension through induction of a relaxed body state. The most common technique is Jacobson’s progressive muscle relaxation or applied relaxation. | 36 |
| Third-wave components (*3w*) | Various techniques aimed at helping patients to develop more adaptive emotional responses to situations, such as the ability to observe symptomatic processes without overly identifying with them or without reacting to them in ways that cause further distress 19. Some typical examples include training in mindfulness, self-compassion, or acceptance. | 14 |
| Behaviour therapy for insomnia (*bi*) | This skill aims at treating chronic insomnia based on the principles of sleep restriction and stimulus control. It may also involve cognitive restructuring around maladaptive beliefs for sleep. It may also involve teaching sleep hygiene: however, sleep hygiene only would count towards lifestyle modification and hence, *pe*. | 4 |
| Relapse prevention (*rp*) | Review of learned skills and listing action plans for the foreseeable future problems based on the skills learned. A mere explanation of relapse in depression counted towards *pe*; in order to qualify for relapse prevention component, it needs more participation from the patient. | 62 |
| Homework required (*hw*) | When completion of some homework assignment is required (or explicitly encouraged repeatedly) before proceeding with the program, either checked by humans or mandated by the program. The homework must pertain to exercise in applying the learned CBT or other skills in one’s own situations and must require some active participation from the participant. Simple reviewing of the materials or further reading were not regarded as homework. | 68 |
| Initial face-to-face contact (*ff*) | Initial face-to-face human contact, such as the initial evaluation session or the initial orientation session, is present. We also considered this component to be present when the patients were receiving conventional drug treatment and the doctors were aware that the patients were in the trial, or when they were referred from their general practitioner for the trial. | 57 |
| Automated encouragement to proceed with iCBT (*ae*) | Provision of automated, fixed prompts/encouragements to proceed with the treatment program. Such prompts should not contain any support related to the therapeutic contents. | 48 |
| Human encouragement to proceed with iCBT (*he*) | Prompts/encouragements are prepared and provided by human beings to proceed with the treatment program via telephone or email. Such prompts should not contain any support related to the therapeutic contents.  Peer support such as discussion group counted towards this component. | 73 |
| Therapeutic guidance for iCBT (*tg*) | Guidance regarding the contents of iCBT. Therapeutic guidance related to the treatment content could be provided on a scheduled basis or as-needed basis. Provision of technical support only did not count toward this component. | 44 |

We included all RCTs that compared any form of iCBT against another form of iCBT or a control condition in the acute phase treatment of adults (≥18 years old) of both sexes with depression, either diagnosed as unipolar major or minor depression according to operationalised diagnostic criteria or judged so by elevated scores on self-report measures with satisfactory reliability and validity. We excluded studies with inpatients or patients with bipolar depression.

We conceptualised CBT broadly as psychotherapy involving training in any of the cognitive or behavioural skills including cognitive restructuring, behavioural activation, interpersonal skills training, structured problem solving or third wave components. iCBT must be a web-based or app-based program using the internet to deliver the CBT contents. We excluded telephone CBT, computerised CBT available only on a clinic-based computer, and when the use of the internet was limited to teleconferencing or emails.

Control conditions of interest included waiting list (WL), no treatment (NT), attention/psychological placebo (APP), and treatment as usual (TAU). Studies have defined different conditions as TAU 17,18. In our study, TAU was defined as including conventional drug treatment either as part of the general practitioners’ care or as part of the study protocol. Watchful waiting or follow-up by community nurses were classified as APP, even when it was regarded as ‘treatment as usual’ in some settings.

Table 1 provides the definitions of our components of interest for the active as well as control arms, as conceptualized and defined by TAF and PC who are expert clinicians and researchers in CBT and iCBT. Table 2 presents various forms of iCBT as well as the control conditions as conceptualised from the component perspective. The intervention could be of any duration.

Table 2. Conceptualization of iCBT or control conditions from the component perspective

|  |  |  |
| --- | --- | --- |
| **iCBT or control conditions** | **Possible decompositions into components** | **No of arms** |
| Cognitive-behavioural therapy (CBT) | *ns ± pe ± re + cr + (ba ± ps ± at ± bi) ± rp ± dt ± ae ± he ± tg ± ff ± hw* | 72 |
| Cognitive therapy (CT) | *ns ± pe ± re + cr ± rp ± dt ± ae ± he ± tg ± ff ± hw* | 2 |
| Behavioural activation (BA) | *ns ± pe ± re + ba ± rp ± dt ± ae ± he ± tg ± ff ± hw* | 11 |
| Problem-solving therapy (PST) | *ns ± pe ± re + ps ± rp ± dt ± ae ± he ± tg ± ff ± hw* | 11 |
| Third-wave CBT (3W) | *ns ± pe ± re ± cr ± ba ± ps ± at ± bi + 3w ± rp ± dt ± ae ± he ± tg ± ff ± hw* | 6 |
| Psychoeducation (PE) | *ns + pe ± rp ± dt ± ae ± he ± tg ± ff* | 6 |
| Waiting list (WL) | *w ± ns ± pe ± dt ± ff* | 51 |
| Treatment as usual (TAU) | *ns + dt + ff* | 10 |
| Attention/psychological placebo (APP) | *ns ± ff* | 9 |
| No treatment (NT) | *± ff* | 1 |

“*x*” at the beginning or “+ *x*” denote that *x* is required: “± *x*” denotes that *x* is optional. At least one of the components in ( ) is required.

Abbreviations for components as in Table 1.

Pairs of two independent reviewers (TAF, AS, EGO) classified all identified treatment arms and their constituent components according to the definitions in Tables 1 and 2, using all available information from the publications, the iCBT programs if accessible, and inquiries with the original investigators. The component could be of any length. A CBT skill component was judged present when it was mentioned as such in the publication and/or was allocated a session/lesson in the program. When skills not covered in this classification (e.g. “expressive writing”, “dreamwork”) were included, we assumed such interventions to have some non-specific treatment effects only. The impact of this assumption was tested in a sensitivity analysis excluding such trials.

Pairs of two independent reviewers (TAF, AS, EGO) independently assessed the validity of the included studies using the revised risk of bias tool by Cochrane 20. The assessment was strictly about the primary outcome used in this review, namely depression severity derived either from the IPD or the AD depending on data availability.

Any disagreement was resolved through discussion or through consultation with a third reviewer.

## Collection of IPD and integrity checks

Authors of the identified studies were contacted via email and requested to contribute individual-level data. When data provision was explicitly declined or none of the authors responded after three attempts, the IPD was deemed unavailable. After collecting IPD, two independent reviewers cross-examined the provided data against the original publications. When the numbers did not match, we contacted the authors for clarification.

## Statistical analyses

Our primary outcome was depression severity as measured on a continuous scale for depression at the end of the acute phase treatment. We accepted any depression measures with established reliability and validity. The scale scores were converted into the most frequently used scale, namely the Patient Health Questionnaire-9 (PHQ-9) 21, when the established conversion algorithms were available 22,23.

Our secondary outcomes were (i) dropout from treatment, defined as completing less than 80% of the contents of the program in IPD studies or according to the original authors’ definition in AD studies, and (ii) dropout from the end-of-treatment assessment for any reason.

We first conducted a NMA at the treatment level to examine if the network of the identified trials was amenable to NMA. We assessed the transitivity of the network, a fundamental prerequisite for NMA to have effect modifiers evenly distributed across comparisons, by examining the distribution of potential effect modifiers in studies grouped by treatment comparison. We performed a two-step random effects NMA at the treatment level, using both AD and IPD studies; for the former we used the published data, and for the latter we used multiple imputations based on IPD to impute missing data. We performed the NMA in a frequentist setting in R using *netmeta* 24, assuming common heterogeneity for all treatment comparisons 25. We checked network inconsistency, a statistical expression of intransitivity, using the back-calculation 26 and the design-by-treatment method 27.

Subsequently, we employed cNMA models that jointly synthesised AD and IPD studies. We conducted both a two-step and a one-step cNMA. In the two-step approach, we synthesised trial-level estimates of treatment effects from studies whose IPD were available and therefore could be re-analysed, and trial-level estimates from studies whose IPD were not available and whose published results were used. In the one-step approach, we used the full IPD including patient-level covariates when available and trial-level estimates of treatment effects when IPD were not available. The models disentangled the effects of components assuming additivity 9, and examined component-covariate interactions utilising shrinkage methods 28. We estimated component-specific ‘incremental mean differences’ (, the added benefit of adding component to a treatment). The component-covariate interactions were modelled assuming linearity. We examined the robustness of the assumption of additive component effects by conducting a sensitivity analysis including two-way interactions. We used the parameter estimates to develop a web-application where the inputs are patient characteristics and two combinations of components, and the output is the estimated relative treatment effects between the two combinations.

We repeated the procedure for the two secondary outcomes, employing a binomial likelihood, to estimate ‘incremental odds ratios’ ( for each component) 7. For the outcome “dropout from treatment,” we only used studies comparing active treatments, since this outcome cannot be measured for controls (WL, TAU, NT, APP).

We conducted four pre-specified sensitivity analyses i) excluding studies without formal diagnosis of major depression, ii) excluding studies with patients from special populations, iii) excluding arms that used skills not covered in our classifications, and iv) limiting to studies with high adherence, for the primary outcome. In addition we did two post-hoc analyses v) by excluding arms which taught more than four CBT components, and vi) by examining a model that assumed two-way interactions between the components.

Details of statistical models used and of fitting procedures are given in the appendix (pp 9-15). The appendix (p 16) lists changes from the protocol.

## Data availability bias, Small study effects and Publication bias

We examined if studies providing IPD were systematically different from studies not providing IPD by comparing the baseline characteristics and symptom changes of included participants.

We examined possible small study effects and publication bias by visually inspecting contour-enhanced funnel plots of pairwise meta-analyses for efficacy between all active arms versus WL.

## Data sharing

The full aggregate-level data and the analysis R codes are available online at GitHub (<https://github.com/esm-ispm-unibe-ch-REPRODUCIBLE/iCBT-cNMA>). Individual-level data cannot be made available due to confidentiality agreements in the original studies.

## Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

# RESULTS

## Selected studies

Up to January 2019, we screened 21162 citations, of which we examined 131 potentially eligible studies in full text. We finally included 76 trials and sought IPD from the original authors. Authors of forty-eight (63%) studies agreed to provide IPD representing 11704 randomised participants; the remaining 28 studies were included as aggregate data, representing 6474 randomised participants (See PRISMA flowchart in the appendix (p 17). See the appendix (pp 18-41) for the list of the included studies (with responses to author queries), excluded studies (with reasons) and studies awaiting assessment).

The participants’ weighted mean age was 42.0 years, while 71% (12406 out of 17521 reported) were women. Operationalised diagnostic criteria were used in 27 studies.

These 76 studies included 179 arms. They included all of our pre-specified components of interest, ranging from *bi* (behaviour therapy for insomnia) in 4 arms to *dt* (conventional drug treatment) in 143 arms. The interrater reliability of judgements for components was excellent, with an average percentage agreement of 93.3% and an average kappa of 0.76 for the 17 components (Table 1, appendix pp 42-51). We were able to have access to five iCBT programs ourselves (Sadness Program, Wellbeing Course, Kokoro-app, BeatingTheBlues and MoodGym) which were used in 20 of the included studies (appendix pp 42-51). The identified components were confirmed by the coauthors who contributed the IPD. Using the conceptualisations in Table 2, each arm represented CBT (n=72), CT (n=2), BA (n=11), PST (n=11), 3W (n=6), PE (n=6), WL (n=51), TAU (n=10), APP (n=9), or NT (n=1). The treatment duration ranged from 3 to 24 weeks (median 8). The median completion rate of program lessons was 72% (range: 25% to 95%) in 46 studies that reported the values. Antidepressant use was reported by 40% of the participants (4031 out of 10041 reported). (appendix pp 52-57)

The risk of bias according to the Cochrane’s revised risk of bias tool 20 was low in 66 studies (86%) for the randomisation process, in 6 studies (8%) for the deviation from intended interventions, in 64 studies (83%) for missing outcome data, in 2 studies (3%) for measurement of the outcome, and 53 studies (69%) for selection of the reported results. The inter-rater reliability of risk of bias assessment was satisfactory, with an average percentage agreement of 93.9% and an average weighted kappa of 0.52 for the five domains. (appendix pp 58-59)

## NMA of treatments for depression severity

In the IPD studies a small minority of patients (387 out of 11122, 3.5%) had scores less than 5 points on PHQ-9 at baseline and would be classified as euthymic according to the commonly accepted cut-off 21. We excluded such patients in the following analyses because they would not normally be targets for treatment for acute depression. Patients missing both baseline and endpoint scores (390 out of 11122, 3.5%) were also excluded. Two studies 29,30 used scales that could not be converted into PHQ-9 and were therefore excluded from the following analyses.

The mean age of the participants was 42.9 in the 48 trials with IPD and 39.5 in the 28 studies with AD. The proportion of women were 65.5% and 71.1%: the average baseline and endpoint PHQ-9 scores were13.3 and 9.1, and 13.0 and 8.9, respectively. Overall, there was no indication of systematic differences between IPD and AD studies.

Figure 1 shows the network geometry for the primary outcome. Table 3 shows the results from the pairwise meta-analyses and NMA at the treatment level, using the two-step approach in which we synthesised trial-level results based on the IPD as described above when IPD were available and could be re-analysed, and trial-level results as reported in the original publications when IPD were not available. CT, BA, 3W, CBT, PE and PST were all shown to be superior to WL. The appendix (pp 61-63) shows the forest plots of pairwise meta-analyses with more than five studies. Transitivity was assessed by investigating the distribution of potential effect modifiers, including age, gender and baseline severity (appendix p 64): one comparison, namely CT vs APP, had much higher baseline severity but otherwise these effect modifiers appeared to be evenly distributed across comparisons. The design-by-treatment test for global inconsistency showed strong evidence of global inconsistency (Q=37.8, 17 degrees of freedom, p=0.003); however, the back-calculation method identified only one comparison out of 18 showed evidence of inconsistency, a proportion that would be empirically expected 31. (appendix p 65) The contour-enhanced funnel plot comparing all active treatments vs WL showed no evidence of publication bias (appendix p 66).

Table 3. Relative effects (differences in PHQ-9 scores) of iCBT of depression

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3W | -1.7  (-4.3 to 0.9) | 1.6  (-1.5 to 4.7) | -2.0  (-4.5 to 0.6) | . | . | . | . | -0.3  (-3.8 to 3.1) | -3.3  (-4.6 to -1.9) |
| -1.9  (-3.4 to -0.5) | APP | 2.1  (-0.3 to 4.4) | 1.4  (-0.2 to 3.1) | 4.2  ( 1.4 to 7.1) | . | 2.2  (-0.3 to 4.7) | . | . | -1.5  (-3.4 to 0.4) |
| 0.5  (-0.9 to 1.8) | 2.4  ( 1.1 to 3.7) | BA | 0.3  (-4.7 to 5.3) | . | . | -3.3  (-4.9 to -1.7) | . | . | -3.0  (-4.1 to -1.9) |
| -0.4  (-1.6 to 0.7) | 1.5  ( 0.3 to 2.7) | -0.9  (-1.9 to 0.1) | CBT | . | -0.7  (-3.3 to 1.8) | -0.7  (-2.3 to 0.8) | -0.3  (-3.0 to 2.4) | -1.5  (-2.5 to -0.4) | -2.7  (-3.2 to -2.2) |
| 2.3  (-0.9 to 5.5) | 4.2  ( 1.4 to 7.1) | 1.8  (-1.3 to 4.9) | 2.7  (-0.3 to 5.8) | CT | . | . | . | . | . |
| -1.2  (-3.9 to 1.6) | 0.8  (-2.0 to 3.6) | -1.6  (-4.4 to 1.1) | -0.7  (-3.3 to 1.8) | -3.4  (-7.4 to 0.5) | NT | . | . | . | . |
| -1.6  (-3.2 to 0.0) | 0.3  (-1.1 to 1.8) | -2.1  (-3.3 to -0.9) | -1.2  (-2.4 to 0.0) | -3.9  (-7.1 to -0.7) | -0.5  (-3.3 to 2.3) | PE | . | . | -1.7  (-5.7 to 2.3) |
| -1.9  (-3.4 to -0.5) | -0.0  (-1.5 to 1.5) | -2.4  (-3.8 to -1.1) | -1.5  (-2.6 to -0.4) | -4.2  (-7.5 to -1.0) | -0.8  (-3.6 to 2.0) | -0.3  (-1.9 to 1.2) | PST | . | -1.0  (-2.0 to 0.0) |
| -1.8  (-3.2 to -0.3) | 0.2  (-1.3 to 1.7) | -2.3  (-3.6 to -0.9) | -1.3  (-2.3 to -0.3) | -4.1  (-7.3 to -0.9) | -0.6  (-3.3 to 2.1) | -0.2  (-1.7 to 1.4) | 0.2  (-1.3 to 1.6) | TAU | . |
| -3.0  (-4.2 to -1.9) | -1.1  (-2.3 to 0.1) | -3.5  (-4.4 to -2.6) | -2.6  (-3.1 to -2.1) | -5.3  (-8.4 to -2.3) | -1.9  (-4.5 to 0.7) | -1.4  (-2.6 to -0.2) | -1.1  (-2.1 to -0.1) | -1.3  (-2.4 to -0.2) | WL |

3W: Third-wave CBT, APP: attention/psychological placebo, BA: behavioural activation, CBT: cognitive-behavioural therapy, CT: cognitive therapy, NT: no treatment, PE: psychoeducation, PHQ-9: Personal Health Questionnaire-9, PST: problem-solving therapy, TAU: treatment as usual, WL: waiting list

The right-upper half of this table shows the direct effects (pairwise meta-analyses). The left-lower half shows the NMA results. Common heterogeneity τ was estimated to be 1.1 in terms of the PHQ-9 scores. An effect size <0 in the lower half shows that the treatment in the column is favoured (i.e. lower PHQ-9 scores) versus the treatment in the row.

## Component NMA for depression severity

Figure 2 shows the main effects from the IPD cNMA, using the one-step approach in which we analysed IPD if available but also used trial-level results when IPD were not available. There was stronger evidence that *behavioural activation (ba)* and *non-specific treatment effects (ns)* had beneficial effects, while *relaxation (re)* was detrimental. There was weaker evidence that *behaviour therapy for insomnia (bi)* was also helpful, showing large point estimates but with wider confidence intervals.

Baseline severity was the strongest prognostic factor but the estimated effect modifications (i.e. interactions with treatment components) were relatively small. Full results for the interaction terms, the network diagrams at the component level and results from cNMA at the AD level are given in the appendix (pp 67-72).

Based on these results we can estimate patient-specific relative effects between any combinations. For example, for a 45-year-old female patient, in a relationship and with a baseline PHQ-9 score of 12, the relative effect of an efficient iCBT package consisting of psychoeducation, behavioural intervention and problem solving using automated and human encouragement versus waiting list control can be estimated as (*ns* + *pe* + *ba* + *ps* + *ae* + *he*) versus *w* to amount to -4.9 (95%Credible Interval [CrI]: -6.8, -3.1). The web app at <https://esm.ispm.unibe.ch/shinies/cNMA_iCBT/> provides estimates of relative efficacies between any two combinations of components, for any given patient characteristics.

## Secondary outcomes

For dropouts from treatment, we conducted cNMA of studies comparing active treatments from five trials providing AD and 10 providing IPD. Due to the limited sample size, we only conducted a two-step cNMA. The obtained estimates for the components were imprecise but there was some suggestion that *automated encouragement (ae)* and *human encouragement (he)* reduced dropout from treatment (appendix p 73). For dropouts from assessment, both the two-step and one-step component NMAs were conducted (appendix pp 74-76).

## Sensitivity analyses

The appendix (pp 77-83) shows the results of the pre-specified and post hoc sensitivity analyses for the primary outcome. Results were essentially concordant with the primary analyses.

# DISCUSSION

We conducted a NMA of the broadly conceived iCBT treatments for depression and their control conditions, and a component NMA of 17 components variably included in these packages based on 48 trials with IPD and 28 trials with AD. All iCBT treatment packages were found superior to the waiting list control. There was evidence that *non-specific treatment effects (ns)* and *behavioural activation* (*ba*) may have beneficial effects while *relaxation* (*re*) may have negative effects, and weaker evidence that *behaviour therapy for insomnia* (*bi*) may be helpful. Having *automated encouragement via emails or text messaging (ae)* and *human encouragement by telephone or email without reference to the therapeutic contents* (*he*) may enhance adherence to the treatment. Baseline severity emerged as the strongest prognostic factor: given the baseline severity, age, gender and relationship status, our web app can estimate the relative effects of any combination treatment over another.

López-López et al’s 8 recent component NMA of CBT for depression found no evidence of specific effects for any skills or delivery formats. These negative findings may be due to the fact that they included mostly face-to-face CBT interventions, which generally allowed broad discretion to therapists who may have reduced, modified or even left out some contents while introducing new contents outside the therapy manual, unless the trials were rigorously monitored for fidelity. Their decomposition of the CBT contents also raised some concerns because some arms had only homework as their content while many arms did not have psychoeducation (few CBT programs would take place without psychoeducation about depression and its cognitive-behavioural model).

In our study *behavioural activation (ba)* emerged as a beneficial component. This is in line with the recent systematic review of dismantling studies, which estimated a pooled standardised mean difference (SMD) of -0.46 (95%CI: -0.91 to -0.01) for the additive effect of behavioural activation 6. Given that the typical standard deviation of PHQ-9 is approximately 6, the iMD of -1.83 (-2.90 to -0.80) for *behavioural activation (ba)* in our cNMA corresponds to a standardised effect size of -0.31 (-0.48 to -0.13) and is largely consistent with the estimate of the systematic review.

The same systematic review found no additive effect of cognitive restructuring: the SMD based on three dismantling studies was -0.06 (-0.68 to 0.55) 6. This is again concordant with our current findings which estimated iMD for *cognitive restructuring (cr)* to be 0.30 (-0.87 to 1.41), corresponding with standardised effect size of 0.05 (-0.15 to 0.24).

The component *non-specific treatment effects (ns)* had a robust additive effect of an iMD of -1.41 (-2.52 to -0.30), corresponding with a standardised effect size of -0.24 (-0.42 to -0.05). In our decomposition model, *ns* comprehends both the placebo effect and the common or non-specific psychotherapy factor 33,34, and can contribute to the efficacy of many bona fide treatments in comparison with non-active controls.

There was suggestive evidence that *behaviour therapy for insomnia (bi)* could be beneficial. However, it was included as a component in only four studies and the estimates for its efficacy were imprecise. It must be pointed out, however, that these four studies did not limit their participants to those having depression and insomnia. The effects of *bi* warrants further research.

In contrast, *relaxation (re)* emerged as potentially harmful in our cNMA, with iMD of 1.20 (0.17 to 2.27). We are aware of no dismantling study for relaxation in depression treatments. However, the cNMA of CBT for panic disorder has also identified muscle relaxation as detrimental 7. There may be various potential explanations for this. Relaxation tended to be included in CBT arms which included a greater number of skills, which may have allowed less time to learn and practice those particular skills and have reduced their efficacy. However, when we ran a sensitivity analysis limiting the included number of CBT skills per arm to four or less, the estimates were essentially similar to those of the primary analysis. Relaxation may be conceived as working in the opposite direction of exposure or behavioural activation, which may be principal therapeutic mechanisms in anxiety and depression, respectively. Relaxation may also plausibly exert its effects by reducing hyperarousal symptoms but this may have been unpleasant for depressed patients who already feel flat and under-aroused. In our network, relaxation exercises were included in 36 arms but the reports generally did not specify what patients actually did in each program.

With regard to delivery formats of iCBT, we studied the effects of *initial face-to-face contact (ff)*, *automated encouragement (ae)*, *human encouragement (he)* and *therapeutic guidance (tg)*. We found that *human encouragement (he)* in conjunction with *automated encouragement (ae)* decreases dropout from treatment (combined iOR=0.32, 95%CI: 0.13 to 0.93) (appendix p 74) and might be able to promote therapy efficacy (combined iMD= -0.55, 95%CI: -1.75 to 0.65) (Figure 2). Because human encouragement without reference to the therapeutic contents can be provided by trained laypersons, these findings if replicated could increase scalability of iCBT.

The findings about *therapeutic guidance (tg)* may appear surprising given that guided iCBT has been shown to be superior to unguided iCBT 10, especially among patients with higher baseline severity 35. In our decomposition, guided iCBT could involve *human encouragement (he)* only or both *human encouragement (he)* and *therapeutic guidance (tg)*: the additive effect of guidance could then be due to *human encouragement* (iMD=-0.5, -1.5 to 0.5) or due to the combination of *human encouragement* plus *therapeutic guidance* (combined iMD=-0.6, -1.7 to 0.4) when the baseline severity was high at 25 points; the incremental effect grew smaller when the baseline severity was lower. The component perspective also brought some insight into the nature of the current unguided iCBT programs. They often included *relaxation (re)* as component: 48% (14/29) when neither *he* nor *tg* was used, 45% (13/29) when only *he* was used, but only 20% (9/44) when *he+tg* were used (Fisher’s exact p=0.02). There may therefore be room for unguided iCBT itself to improve its efficacy by appropriately choosing the included components.

There are several important limitations to this study. First, the included studies were limited to iCBT. Any conclusions about specific efficacy of CBT skills therefore pertain to iCBT and will inform which components to include in the efficacious and efficient iCBT package, but such findings may not readily generalise to face-to-face CBT. For example, behavioural activation emerged as beneficial while cognitive restructuring did not in the current analyses not because of these skills’ intrinsic efficacy but simply because of their ease of learning within iCBT. However, the concordance between our findings and those of face-to-face dismantling studies 6 is encouraging. In a similar vein, our conclusions are applicable to the components and delivery modes as implemented in the included studies. There are potentially different ways for a certain component to be delivered and it is always possible that a specific form of the same component may prove to be particularly efficacious. Second, although the median completion rate of the program lessons was 72% and thus we can assume that most participants had actually conducted a majority of the included components, no data were available as to exactly which components were conducted and by whom. Our analysis considered components as being simply present or absent based on the descriptions of each program. In reality, components may have varied between programs in terms of their length, depth and content. Even when present, some components may have been less comprehensively executed than others, for example when they were offered later in the iCBT package than earlier or when they were presented along with many other components. Neither is there any guarantee that the completion of a component means that the participants had acquired the corresponding skill. The present analyses, therefore, can only speak to the effects of the decision to include certain components as their principal ones in the iCBT program at the start of the treatment, as per the intention-to-treat principle of interpreting RCTs. In order to appreciate the effects of actual execution of the treatment, we would need more detailed data of individual participants’ adherence, their learned contents and different analytical approaches 36. Third, several programs had components not covered in the current scheme, including expressive writing, dreamwork, positive psychology, graded exposure, worry time or physical exercises. Such trials may have had a slightly different focus, such as the transdiagnostic program including graded exposure where the same cognitive restructuring technique may have been presented slightly differently than when it is targeted mainly at depression. A sensitivity analysis excluding trials that used such miscellaneous components, however, corroborated the primary analysis results. Fourth, there was indication of intransitivity for a comparison between CT and APP and of global inconsistency in the network. However, this is unlikely to have affected the estimates for the components to a great extent, because the components included in the CT-APP comparison were also included in other comparisons. It remains the case, however, that we only assessed the transitivity assumption for a small number of possible effect modifiers. Thus, our network may have been confounded by unobserved imbalances across comparisons, such as antidepressant use and treatment duration. Fifth, when the IPD were pooled, there were only four commonly reported patient characteristics that we could analyse as prognostic factors or effect modifiers. Important potential effect modifiers, such as childhood adversities 37 or baseline cognitive or behavioural skills 38, were therefore not included in our model. Researchers are encouraged to agree on essential measurements to be taken in future iCBT trials. Lastly, we assumed additivity of the component effects, i.e. that for any given component c, the relative effect of (c+X) vs. X is the same for any combination of components X (not including c). Thus, our results assumed no interactions between components. Although a post hoc sensitivity analysis examining potential interactions among components provided similar results to the additive model, the study was possibly underpowered to detect interactions among components. It needs also be emphasized that combinations of components can be justified only so far as they are clinically sensible. We must also be careful when extrapolating the combinations beyond those examined in the current dataset: as we combine more and more components to build packages not explored in the trials, the uncertainty around the estimates increases, while possible deviations from our assumptions (e.g. additivity of component effects) may have a bigger impact on the validity of our results. More and larger studies varying in included components, for example in the form of fully factorial trials 39,40, are necessary to extend and consolidate the estimates.

The major strengths of the study are as follows. We employed the state-of-the-art evidence synthesis methods to elucidate specific efficacies of various skills and delivery methods of iCBT. The cNMA increased precision by including more studies than narrowly defined dismantling studies and by combining direct and indirect estimates. The included studies were generally of adequate quality except for the domains where psychotherapy trials cannot escape possible biases due to unfeasibility of blinding. The decomposition of treatments sometimes called for subtle judgements and publications may have failed to provide enough information: with detailed definitions as in Table 1 and inquiries of the authors where necessary (see the appendix pp 18-41 for results of our communications), we were able to achieve satisfactory to excellent inter-rater agreement. In the current network, there was no evidence of data availability bias for IPD or of publication bias of active treatments over the waiting list, the most frequent control condition. Lastly, by using IPD we were able to estimate relative efficacies of any combination of components based on patients’ baseline characteristics.

In conclusion, this IPD cNMA of iCBT has identified potentially helpful and less helpful components and delivery formats for achieving depression reduction and enhancing adherence to the program with suggestive evidence. Future iCBT packages aiming to be effective and efficient may include *behavioural activation* (*ba)* but not *relaxation* (*re)*. They may further choose to include *behavioural therapy for insomnia (bi)* and *problem solving (ps)* but probably not *cognitive restructuring (cr)*. To boost adherence, they may employ *automated encouragement (ae)* and *human encouragement (he)*. Such packages have potentials to be more efficacious, less burdensome for users and less demanding on provider resources, with the net effect of rendering iCBT even more scalable. However, readers should note that our analyses are limited by several important factors outlined above; moreover, the evidence supporting some of these recommendations, especially for *cognitive restructuring*, *problem solving* or *therapeutic guidance* is still relatively imprecise, warranting further experimentations to refine iCBT packages.

## Contributors

TAF and OE conceived the study. TAF, AS, OE, PC, and EK designed the study. PC, EK, AS, EGO and TAF selected the studies and extracted data. GA, CGB, JS, TB, FWB, CB, PC, IC, HC, AM, JD, MH, DDE, LF, NRF, DRS, IDE, EF, VK, AG, SG, LL, SB, HDH, LHS, RJ, RK, MK, CB, AK, HR, JPK, JS, BM, SM, LB, OL, JL, JM, AWG, DCM, JMM, JGC, SN, A-CZ, KO, ADW, JMN, SP, RP, JS, WP, NEP, DR, IMR, SLR, LBS, JS, VS, VJP, BU, KMPB, SL, NG, VK, KV, LW, AS, PZ, CK, and MH contributed the IPD. OE, EK and PC verified the data. OE analysed the data. TAF, AS, EGO, OE, EK and PC interpreted the results. TAF and OE wrote the first draft of the manuscript. All authors had access to all the data and provided critical input and revisions to the draft manuscripts and approved the final manuscript. TAF, AS, EGO, OE, EK and PC had final responsibility for the decision to submit for publication.

## Declaration of interests

TAF reports grants and personal fees from Mitsubishi-Tanabe, personal fees from MSD, personal fees from Shionogi, outside the submitted work; In addition, TAF has a patent 2018-177688 concerning smartphone CBT apps pending, and intellectual properties for Kokoro-app licensed to Tanabe-Mitsubishi.

JD is co-owner of Behavioral Activation Tech LLC, a small business that develops and evaluates mobile application-based treatments for depression and co-occurring disorders.

AG is currently employee of Soulve Innovations.

BM is currently employee of GAIA AG.

DCM has received consulting fees from Apple, Inc, Otsuka Pharmaceuticals, One Mind Foundation, and Pear Therapeutics, and has an ownership interest in Actualize Health, Inc.

JMM is supported by the Wellcome Trust Grant (104908/Z/14/Z).

SN is currently employee of GET.ON Institut.

DR is currently employee of SilverCloud Health.

## Acknowledgments

This study was supported, in part, by the Japan Society for the Promotion of Science grant-in-aid for scientific research (grant number 17K19808) and by a grant-in-aid from Japan Agency for Medical Research and Development (AMED) to TAF (grant number 19dk0307085), and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005) to AC.

EGO is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215-20005).

JD is supported by the National Institute on Drug Abuse (NIDA K23 DA045766).

HH is funded by the Canadian Institutes of Health Research (152917).

JPK received funding for clinical trials (German Federal Ministry of Health, Servier), payments for presentations on internet interventions (Servier), payments for workshops and books (Beltz, Elsevier, Hogrefe and Springer) on psychotherapy for chronic depression and on psychiatric emergencies.

AC is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017–08ST2–006), the NIHR Oxford and Thames Valley Applied Research Collaboration and the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215–20005).

OE was supported by project grant No. 180083 from the Swiss National Science Foundation (SNSF).

EK was supported by the Netherlands Organization for Health Research and Development (NWO; project number 019.182SG.001).

The views expressed are those of the authors and not necessarily those of any of the funding agencies listed above.

# References

1. Barth J, Munder T, Gerger H, et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Med* 2013; **10**(5): e1001454.

2. Cuijpers P, Karyotaki E, de Wit L, Ebert DD. The effects of fifteen evidence-supported therapies for adult depression: A meta-analytic review. *Psychother Res* 2019: 1-15.

3. Institute of Medicine. Psychosocial Interventions for Mental and Substance Use Disorders: A Framework for Establishing Evidence-Based Standards. Washington, DC: The National Academies Press; 2015.

4. Ahn H, Wampold BE. Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology* 2001; **48**: 251-7.

5. Bell EC, Marcus DK, Goodlad JK. Are the parts as good as the whole? A meta-analysis of component treatment studies. *J Consult Clin Psychol* 2013; **81**(4): 722-36.

6. Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Hollon SD. Component studies of psychological treatments of adult depression: A systematic review and meta-analysis. *Psychother Res* 2017: 1-15.

7. Pompoli A, Furukawa TA, Efthimiou O, Imai H, Tajika A, Salanti G. Dismantling cognitive-behaviour therapy for panic disorder: a systematic review and component network meta-analysis. *Psychol Med* 2018; **48**(12): 1945-53.

8. López-López JA, Davies SR, Caldwell DM, et al. The process and delivery of CBT for depression in adults: a systematic review and network meta-analysis. *Psychol Med* 2019; **49**(12): 1937-47.

9. Welton NJ, Caldwell DM, Adamopoulos E, Vedhara K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. *Am J Epidemiol* 2009; **169**(9): 1158-65.

10. Cuijpers P, Noma H, Karyotaki E, Cipriani A, Furukawa TA. Effectiveness and acceptability of cognitive behavior therapy delivery formats in adults with depression: A network meta-analysis. *JAMA psychiatry* 2019; **76**(7): 700-7.

11. Wagner B, Horn AB, Maercker A. Internet-based versus face-to-face cognitive-behavioral intervention for depression: a randomized controlled non-inferiority trial. *J Affect Disord* 2014; **152-154**: 113-21.

12. Karyotaki E, Ebert DD, Donkin L, et al. Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. *Clin Psychol Rev* 2018; **63**: 80-92.

13. Furukawa TA, Karyotaki E, Suganuma A, et al. Dismantling, personalising and optimising internet cognitive-behavioural therapy for depression: a study protocol for individual participant data component network meta-analysis. *BMJ open* 2019; **8**(11): e026137.

14. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions: Checklist and Explanations. *Ann Intern Med* 2015; **162**(11): 777-84.

15. Stewart LA, Clarke M, Rovers M, et al. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA* 2015; **313**(16): 1657-65.

16. Cuijpers P, Karyotaki E, Ciharova M. A meta-analytic database of randomised trials on psychotherapies for depression. 2020. osf.io/825c6.

17. 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.

18. Cuijpers P, Quero S, Papola D, Cristea IA, Karyotaki E. Care-as-usual control groups across different settings in randomized trials on psychotherapy for adult depression: a meta-analysis. *Psychol Med* 2019: 1-11.

19. Roemer L, Erisman SM, Orsillo SM. Mindfulness and acceptance-based treatments for anxiety disorders. In: Antony MM, Stein MB, eds. Oxford Handbook of Anxiety and Related Disorders. Oxford: Oxford University Press; 2008: 476-87.

20. Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj* 2019; **366**: l4898.

21. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; **16**(9): 606-13.

22. Wahl I, Lowe B, Bjorner JB, et al. Standardization of depression measurement: a common metric was developed for 11 self-report depression measures. *J Clin Epidemiol* 2014; **67**(1): 73-86.

23. Furukawa TA, Reijnders M, Kishimoto S, et al. Translating the BDI and BDI-II into the HAMD and vice versa with equipercentile linking. *Epidemiol Psychiatr Sci* 2019: 1-13.

24. Rücker G, Krahn U, König J, Efthimiou O, Schwarzer G. netmeta: Network Meta-Analysis using Frequentist Methods. 2019. <https://github.com/guido-s/netmeta>

<http://meta-analysis-with-r.org>.

25. Efthimiou O, Debray TP, van Valkenhoef G, et al. GetReal in network meta-analysis: a review of the methodology. *Res Synth Methods* 2016.

26. Konig J, Krahn U, Binder H. Visualizing the flow of evidence in network meta-analysis and characterizing mixed treatment comparisons. *Stat Med* 2013; **32**(30): 5414-29.

27. White IR, Barrett JK, Jackson D, Higgins JP. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Research Synthesis Methods* 2012; **3**: 111-25.

28. Seo M, White IR, Furukawa TA, et al. Comparing methods for estimating patient-specific treatment effects in individual patient data meta-analysis. *Stat Med* 2020.

29. Beevers CG, Pearson R, Hoffman JS, Foulser AA, Shumake J, Meyer B. Effectiveness of an internet intervention (Deprexis) for depression in a united states adult sample: A parallel-group pragmatic randomized controlled trial. *J Consult Clin Psychol* 2017; **85**(4): 367-80.

30. Day V, McGrath PJ, Wojtowicz M. Internet-based guided self-help for university students with anxiety, depression and stress: a randomized controlled clinical trial. *Behav Res Ther* 2013; **51**(7): 344-51.

31. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol* 2013; **42**(1): 332-45.

32. Seo M, Furukawa TA, Veroniki AA, et al. The Kilim plot: A tool for visualizing network meta-analysis results for multiple outcomes. *Res Synth Methods* 2021; **12**(1): 86-95.

33. Wampold BE. How important are the common factors in psychotherapy? An update. *World Psychiatry* 2015; **14**(3): 270-7.

34. Cuijpers P, Reijnders M, Huibers MJH. The Role of Common Factors in Psychotherapy Outcomes. *Annual review of clinical psychology* 2019; **15**: 207-31.

35. Karyotaki E, Efthimiou O, Miguel C, et al. Internet-Based Cognitive Behavioral Therapy for Depression: A Systematic Review and Individual Patient Data Network Meta-analysis. *JAMA psychiatry* 2021.

36. Hernán MA, Hernández-Díaz S. Beyond the intention-to-treat in comparative effectiveness research. *Clin Trials* 2012; **9**(1): 48-55.

37. Nemeroff CB, Heim CM, Thase ME, et al. Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. *Proc Natl Acad Sci U S A* 2003; **100**(24): 14293-6.

38. Cheavens JS, Strunk DR, Lazarus SA, Goldstein LA. The compensation and capitalization models: a test of two approaches to individualizing the treatment of depression. *Behav Res Ther* 2012; **50**(11): 699-706.

39. Uwatoko T, Luo Y, Sakata M, et al. Healthy Campus Trial: a multiphase optimization strategy (MOST) fully factorial trial to optimize the smartphone cognitive behavioral therapy (CBT) app for mental health promotion among university students: study protocol for a randomized controlled trial. *Trials* 2018; **19**(1): 353.

40. Watkins E, Newbold A, Tester-Jones M, et al. Implementing multifactorial psychotherapy research in online virtual environments (IMPROVE-2): study protocol for a phase III trial of the MOST randomized component selection method for internet cognitive-behavioural therapy for depression. *BMC Psychiatry* 2016; **16**(1): 345.

Figure 1. Network diagram

3W: Third-wave CBT, APP: attention/psychological placebo, BA: behavioural activation, CBT: cognitive-behavioural therapy, CT: cognitive therapy, NT: no treatment, PE: psychoeducation, PST: problem-solving therapy, TAU: treatment as usual, WL: waiting list

Figure 2. Individual participant data component NMA for depression severity

iMD: incremental mean difference, CrI: credible interval, PHQ-9: Patient Health Questionnaire-9, in relationship=1 when married or having a stable partner; =0 when single, separated/divorced, or widowed.

Potentially beneficial components are shown in green (darker green for stronger statistical evidence), while potentially harmful components are shown in red according to an index similar to the Z-score (median of the posterior distribution divided by the corresponding standard deviation for Bayesian analyses), thus taking account of the magnitude of the effect estimates and their uncertainty 32. More details about the colouring scheme are provided in the appendix (p 68).

The specific efficacy for *conventional drug treatment (dt)* could not be estimated because *dt* was either present or absent in all comparisons in the network.

Common heterogeneity τ was estimated to be 1.20 (95%CrI: 0.89 to 1.57) in terms of the PHQ-9 scores.