A meta-analysis of CBT efficacy for depression comparing adults and older adults

Alessa D. Werson a,b,*, Richard Meiser-Stedman c, Ken Laidlaw b

a Norwich Medical School, University of East Anglia, United Kingdom
b Department of Psychology: College of Life and Environmental Sciences, University of Exeter, United Kingdom
c Department of Clinical Psychology and Psychological Therapies, Norwich Medical School, University of East Anglia, United Kingdom

ARTICLE INFO

Keywords: Meta-analysis Cognitive behavioral therapy CBT Major depressive disorder Depression Young adults Middle aged adults Older adults Older people

ABSTRACT

Introduction: This meta-analysis investigates CBT treatment efficacy for depression, and compares outcomes between adults (young and middle aged) and older adults (OA). Methodology. Effect sizes (Hedges’ g) were obtained from 37 peer-reviewed RCTs, 25 adult papers (participant n = 2948) and 12 OA papers (participant n = 551), and analysed with the random effects model. Results. No significant difference between age groups is reported in terms of CBT efficacy for depression compared to other treatments (Qbetween (1) = 0.06, p = .89), with the overall effect favouring CBT over any other treatments (g = 0.48, 95% CI = 0.29–0.68). The same pattern of results was found when restricting studies to those which used active control conditions (Qbetween (1) = 0.03, p = .86) or passive control conditions (Q (1) = 2.45, p = .12). Discussion. No significant differences in efficacy for CBT treatment for depression are found when comparing adults and OA. CBT is as efficacious with OA as with adults.

Depression is a significant problem affecting the quality of life of older adults (OA; RCP, 2019). Examining depression treatment efficacy for OA is of particular importance due to the ageing population and depression being reported to be the most common mental health problem in OA (Laidlaw et al., 2008). Additionally, therapists may still be reluctant to work with OA due to stigma attached with psychological treatment for OA (Laidlaw, 2019) or uncertainty of life experiences and life events associated with late life depression influencing on treatment (Werson et al., 2020). It is imperative that OA receive efficacious evidenced-based treatment for depression as left untreated serious consequences may ensue, such as an increased mortality rate (Pocklington, 2017). Additionally, as people are living longer and there are increased demands for access to psychotherapy for depression among OA, barriers such as erroneous negative age-stereotypes related to efficacy of CBT urgently need to be addressed. Cognitive behavioral therapy (CBT) is efficacious as a treatment for OA (Bilbrey et al., 2020) and is now accepted as a first-rank treatment for depression (NICE, 2019). One of the most systematically researched psychotherapies for OA (Laidlaw, 2015), CBT efficacy in OA is supported by numerous sources (e.g. Gould et al., 2012). Clinical data from Improving Access to Psychological Therapy (IAPT) services indicate OA report enhanced recovery rates and reduced attrition for CBT treatment for depression (Chaplin et al., 2015; Pettit et al., 2017).

Despite efficacy data there remains a persistent belief among many healthcare professionals that CBT is less efficacious with OA compared to outcomes achieved by young and middle aged adults (hereinafter collectively referred to as “adults”; Collins and Corna, 2018; Frost et al., 2019). This may be, in part, due to a concern that comorbidity and advanced age act as complicating factors for psychological treatment.

As a consequence of an erroneous belief in reduced efficacy of CBT with OA, a number of potentially negative consequences ensue. This includes a pervasive, and persistent erroneous belief among healthcare professionals, including referrers and psychological therapists, is that OA would not want access to psychological therapy (Collins and Corna, 2018). Research also suggests that mental health problems in later life are poorly treated and are seen as secondary to physical health concerns (Frost et al., 2019). Additionally, an erroneous belief exists that depression in later life is either a normal part of ageing or that it is justifiable and is therefore an ‘understandable’ consequence of the challenges of ageing (Frost et al., 2019). As such, OA with depression can be marginalised and their mental health needs may go unmet.

The world is experiencing a profound and irreversible demographic...
change with increasing numbers and relative proportions of OA active and present in society (United Nations, 2019) and within the psychotherapy community an adequate response will be to ensure those OA who seek psychotherapy have access to evidence-based interventions, such as CBT, delivered by practitioners who do not discriminate based on age. Currently, evidence suggests that OA receive the least psychological support out of all age groups (Pettit et al., 2017).

Up till now when comparing efficacy for CBT between age groups, research has tended to report meta-analyses separately for adults and OA (e.g. Gould et al., 2012; Holvast et al., 2017; Cuijpers et al., 2020a). With a few notable exceptions the approach has been to compare separate meta-analysis outcomes from different studies when considering the question of whether CBT is equivalently efficacious between age groups. As this approach often compares data with different methodological approaches, this is manifestly unsuitable and unsatisfactory for addressing this important question. Where researchers have attempted an integrative review this has been conducted in a wider review of psychotherapy outcome (Cuijpers et al., 2009) or where age-comparisons are a minor aspect of multiple comparisons (Cuijpers et al., 2013b). Similarly, Cuijpers et al. (2020b) examined the outcome of a range of psychological treatments for different age cohorts, but this was examined within the broader category of psychotherapy and therefore does not isolate and address the specific question considering efficacy of CBT exclusively, the NICE recommended guideline for treatment of depression, as investigated in the current meta-analysis.

In a first in-depth comparison, this current paper will systematically review the evidence for the efficacy of CBT for depression and examine whether there are age differences between adults and OA in terms of treatment outcome. Secondly, this meta-analysis examines treatment effect for depression comparing CBT to other treatments, including active and passive treatment. Moderator analyses investigate differences between the two age groups within these analyses.

1. Methods

1.1. Literature search and study selection

This meta-analysis was registered with PROSPERO (CRD42018094089). Studies were searched for on 31st August 2020. Where possible, search terms included Medical Subject Headings (in short called MeSH terms). MeSH terms are used to provide uniformity and consistency in literature searches by indexing and categorising literature. Databases searched were APA PsycInfo, APA PsycArticles, APA PsycExtra, Cochrane Library, Embase, Medline (EBSCO), PubMed Central, Scopus, and Web of Science. PROSPERO and Ethos were also searched to identify any unpublished studies. Reference lists of papers relevant for this meta-analysis were checked to identify additional studies to complete the literature search.

1.2. Inclusion and exclusion criteria

This meta-analysis included peer reviewed, published research studies written in English. Two literature searches were completed, one for studies recruiting adults and one for studies recruiting OA, restricted to Western countries with the aim to preserve homogeneity as different sociocultural factors may influence on treatments and outcomes, thus potentially resulting in unequal comparisons.

So as to ensure a more complete sampling of OA outcome studies, and due to the contemporary ageing population and timeframe research is conducted in, the adult age includes young and middle aged adults which was set at 18–65 years old and OA age at 60+ years old. Studies where recruitment did not differentiate between the two age groups as set for this meta-analysis (e.g. participants included were 18–75 year old), and an overlap was created where these two age groups could not be distinguished from one another were excluded. MeSH search terms used in the literature search for adults were “young adult” and “middle aged”, and for OA “aged” (“Adult MeSH Descriptor Data 2021”, 2021). In databases where use of MeSH terms for age group specification was not possible the terms “young adult”, “adult””, “middle age””, and “18–65” were used for adults, and the terms “older people”, “older adult””, “retired”, “over 65””, “aged”, and “geriat” for OA.

For inclusion in this current meta-analysis, all research studies had to report psychological treatment for depression as a primary aim, measured using reliable and valid psychometric instruments. Participants were required to meet clinical significant depression symptoms as measured with the use of a reliable and valid clinician administered instrument or self-report questionnaire (for example, the Hamilton Depression Rating Scale or Becks Depression Inventory). Studies where depression was not the primary focus of treatment, involving sub-threshold depression, or using interviewer-administrated measures were excluded to promote a focus on patient-centred depression treatment with a clinical population, and using self-report to follow common practice in psychotherapy studies and reduce heterogeneity. Studies that focused on other conditions, involving comorbidities, were included if the main treatment focused on depression. The clinical population was searched for using the terms “major depressive disorder””, “MDD”, “depression” and “low mood”.

Studies were eligible for inclusion if design involved a randomized controlled trial delivering evidence-based CBT protocols, CBT related- and CBT derived therapies carried out in individual, online, or group settings with support from a trained therapist. Randomisation to a control group could involve either active treatment (other types of psychotherapy or pharmacotherapy) or passive control conditions (wait list, usual care, delayed treatment, or placebo). Open trials, inpatient trials, single case studies, and depression treatment studies without CBT protocols were excluded. The search terms used for intervention were “CBT”, “low intensity CBT”, “cognitive behavioural therap”, “cognitive therap” and “Behavioural activation”.

An Intention to treat (ITT) design was an important inclusion criterion for studies, however, exceptions were made for OA papers due to the timeframe during which research was conducted as a number of included OA studies were published prior to ITT becoming an essential element of RCT trials.

To be eligible for inclusion studies were required to present post-treatment self-report outcome measures, so as to permit the calculation of treatment efficacy. Studies with insufficient data for analysis, literature reviews or meta-analyses, or papers presenting follow up data only were excluded.

1.3. Study quality

Quality of papers included were measured using the RCT of Psychotherapy Quality Rating Scale, a valid and reliable tool to assess quality for randomized controlled trials in psychotherapy (RCT-PQRS: Gerber et al., 2011). The RCT-PQRS consists of 25 items, however item 14 was excluded in quality ratings since this assesses long-term outcomes resulting in papers to be scored out of a maximum 53 points with higher scores reflecting higher study quality. Authors AW & KL established an inter-rater agreement using Cohen’s Kappa (κ). From papers included in this meta-analysis, 4 papers were selected at random and rated independently by both authors. Inter-rater agreement showed κ = 0.97.

1.4. Analyses

Analyses were completed using MAVIS v1.1.3 (Meta-Analysis via Shiny; http://kylehamilton.net/shiny/MAVIS/), which uses the metafor package ( Viechtbauer, 2010) in R, and Fig. 3 was created using metafor in R directly. The effect size metric used is Hedges’ g, as it is robust, and provides a conservative means of adjusting for possible overestimation in effect size where included studies have smaller sample sizes.
Hedges' $g$ effect size is interpreted as small for effect sizes of 0.2, medium for 0.5, and large if 0.8. Data was analysed with random effect models as follows:

1. An estimate of the post-treatment effect size for CBT compared to any other condition.
2. A moderator analysis of post-treatment efficacy of CBT compared to other conditions, considering differences between adults and OA studies.
3. A sensitivity analysis of post-treatment outcomes comparing CBT to active treatment (i.e. psychotherapy or pharmacotherapy). A moderator analysis subsequently investigated outcomes between adults and OA to compare efficacy between age groups.
4. A second sensitivity analysis comparing post-treatment data from CBT to passive control conditions (i.e. treatment as usual, wait list and delayed treatment). This sub-group analysis also followed with a moderator analysis, comparing outcomes between adults and OA age groups.

Analysis included publication bias testing through funnel plot asymmetry, using the weighted regression with multiplicative dispersion model. Missing null studies were estimated through Duval and Tweedie's trim-and-fill method (Duval and Tweedie, 2000). The Fail-safe N was calculated using Rosenthal’s approach to appraise robustness for statistical significance and substantive significance (Borenstein et al., 2011).

2. Results

2.1. Studies included

Literature reviews resulted in 12,831 papers meeting search criteria for adults, and 5223 papers for OA. No relevant ongoing or unpublished papers, or theses were identified. Titles and abstracts were screened for eligibility. In total, 25 adult papers and 12 OA papers were eligible and included in this meta-analysis. An overview of the literature screening process can be viewed in Figs. 1 and 2, for adults and OA respectively.

2.2. Characteristics of studies

Characteristics of all papers included in the current meta-analysis are summarised in Tables 1 and 2.

Adult studies include 13 types of comparator conditions with CBT of which 11 are active treatment and two control conditions (see Table 1).
From these groups resulted a total of 2948 adult participants which were included in analysis (active treatment \( n = 1790 \) and control condition \( n = 1158 \)). Participant age range was 22.4–51.7 years, with a mean age of 37 years (SD = 13.1). Attrition for adult studies ranged between 1.9 % and 45 % with an average of 21.5 %.\(^2\)

Eight comparison groups were included in OA studies, consisting of five active treatment groups and three passive control conditions (see Table 2). These groups totalled 551 participants which were included in analysis (active treatments \( n = 164 \) and control condition \( n = 387 \)). The age range for OA participants was 66.4–77.5 years, averaging 67.5 years (SD = 7.3). Attrition rates varied from 9.1 % to 53.6 %, averaging 28.7 %. Seven out of the 12 papers included for OA used ITT design.

2.3. Study quality

The methodological quality of papers in this meta-analysis were rated using the RCT of Psychotherapy Quality Rating Scale (RCT-PQRS: Gerber et al., 2011, see Tables 3, 4 and 5). Adult papers scored an average quality rating of 41 out of a maximum 53 (95 % CI 39, 44.4). Quality ratings for OA papers also averaged 41 (95 % CI 38.9, 43.7). All papers combined showed quality ratings ranging from 28 to 49, with an average of 41 (95 % CI 39.5, 42.7).

2.4. Analyses

2.4.1. Efficacy of CBT to other conditions in treatment of depression

Data analysis was completed firstly comparing CBT to other conditions (including both active treatments and passive control conditions) using post-treatment data from all 37 papers (adult and OA). A random effects model analysis showed a statistically significant, small effect size \( g = 0.48, 95 \% \text{ CI } = 0.29 \text{ to } 0.68 \), favouring efficacy of CBT treatment \( Z = 4.83, p < .001 \) over other conditions.

Study heterogeneity indicated considerably different findings in treatment effect among studies \((Q (36) = 152.70, P < .001, I^2 = 84 \% \)). Publication bias found a significant result with \( t = 3.12, df = 35, p < .01 \). Duval and Tweedie trim-and-fill method suggested three potentially missing studies. Calculating the effect with the three studies included increased the overall effect size, \( g = 0.58 \) (95 % CI = 0.39 to 0.76), indicating no evidence of an overestimate of the true effect resulting from publication bias. Similarly, Rosenthal’s Fail-safe N showed an observed significance level \( p < .0001 \) (target significance level 0.05), with 1267 failed trials necessary to nullify the observed effect. Therefore, there is little concern about file-drawer effects (Borenstein et al., 2011).

---

1. Excludes SD from Farabaugh et al. (2015), Rush et al. (1977), Sava et al. (2009) and Wilson et al. (1983) due to missing data.
2. Excludes Taylor et al. (1977) & Wilson et al. (1983) due to insufficient data.
3. Excludes SD from Beutler et al. (1987), Scogin et al. (1989), and Thompson et al. (2001) due to missing data.
**Table 1**
Selected characteristics of studies involving adults.

<table>
<thead>
<tr>
<th>Study first-named author (date)</th>
<th>Age mean (SD)</th>
<th>CBT type</th>
<th>Psychometric instrument</th>
<th>N post-treatment</th>
<th>N included in analysis</th>
<th>Format</th>
<th>Number of sessions</th>
<th>Control type</th>
<th>N post-treatment</th>
<th>N included in analysis</th>
<th>Analyses</th>
<th>Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Tjak et al. (2018)</td>
<td>41.49 (12.38)</td>
<td>CBT</td>
<td>HAM-D</td>
<td>38</td>
<td>38</td>
<td>Individual</td>
<td>20</td>
<td>ACT</td>
<td>44</td>
<td>44</td>
<td>ITT</td>
<td>15 %</td>
</tr>
<tr>
<td>Carleborg et al. (2013)</td>
<td>44.4 (13.5)</td>
<td>BA + ACT</td>
<td>BDI</td>
<td>40</td>
<td>40</td>
<td>Online</td>
<td>8</td>
<td>WCG</td>
<td>38</td>
<td>40</td>
<td>ITT</td>
<td>2.5 %</td>
</tr>
<tr>
<td>Carter et al. (2013)</td>
<td>38.38 (11.7)</td>
<td>CBT</td>
<td>BDI</td>
<td>25</td>
<td>50</td>
<td>Individual</td>
<td>30</td>
<td>ST</td>
<td>30</td>
<td>50</td>
<td>ITT</td>
<td>45 %</td>
</tr>
<tr>
<td>Chaves et al. (2017)</td>
<td>51.65 (10.36)</td>
<td>CBT</td>
<td>BDI</td>
<td>39</td>
<td>49</td>
<td>Group</td>
<td>10</td>
<td>PPI</td>
<td>34</td>
<td>47</td>
<td>ITT</td>
<td>23.97 %</td>
</tr>
<tr>
<td>Connolly-Gibbons et al. (2016)</td>
<td>36.2 (12.1)</td>
<td>CT</td>
<td>HAM-D</td>
<td>104</td>
<td>118</td>
<td>Individual</td>
<td>16</td>
<td>DT</td>
<td>105</td>
<td>119</td>
<td>ITT</td>
<td>11.81 %</td>
</tr>
<tr>
<td>Conradi et al. (2007)</td>
<td>42.8 (11.3)</td>
<td>CBT</td>
<td>BDI</td>
<td>38</td>
<td>41</td>
<td>Individual</td>
<td>10–12</td>
<td>TAU</td>
<td>65</td>
<td>72</td>
<td>ITT</td>
<td>8.85 %</td>
</tr>
<tr>
<td>David et al. (2008)</td>
<td>37 (8.33)</td>
<td>CT</td>
<td>BDI</td>
<td>50</td>
<td>56</td>
<td>Individual</td>
<td>20</td>
<td>REBT</td>
<td>52</td>
<td>57</td>
<td>ITT</td>
<td>9.74 %</td>
</tr>
<tr>
<td>Dimidjian et al. (2006)</td>
<td>39.90 (10.97)</td>
<td>BA</td>
<td>BDI</td>
<td>36</td>
<td>43</td>
<td>Individual</td>
<td>BA: 24 max Med:10</td>
<td>Med</td>
<td>56</td>
<td>100</td>
<td>ITT</td>
<td>35.66 %</td>
</tr>
<tr>
<td>Driessen et al. (2013)</td>
<td>38.91 (10.30)</td>
<td>CBT</td>
<td>HAM-D</td>
<td>43</td>
<td>164</td>
<td>Individual</td>
<td>16</td>
<td>PST</td>
<td>45</td>
<td>177</td>
<td>ITT</td>
<td>7.42 %</td>
</tr>
<tr>
<td>Ekeblad et al. (2016)</td>
<td>34.2 (10.8)</td>
<td>CBT</td>
<td>BDI</td>
<td>48</td>
<td>48</td>
<td>Individual</td>
<td>14</td>
<td>IPT</td>
<td>48</td>
<td>48</td>
<td>ITT</td>
<td>28.13 %</td>
</tr>
<tr>
<td>Gilbody et al. (2015)</td>
<td>40.35 (14.31)</td>
<td>CT</td>
<td>PHQ-9</td>
<td>119</td>
<td>119</td>
<td>Individual</td>
<td>16</td>
<td>DT</td>
<td>118</td>
<td>118</td>
<td>ITT</td>
<td>25 %</td>
</tr>
<tr>
<td>Hallgren et al. (2016)</td>
<td>43 (12)</td>
<td>iCBT</td>
<td>MADRS</td>
<td>275</td>
<td>317</td>
<td>Online</td>
<td>–</td>
<td>TAU</td>
<td>256</td>
<td>312</td>
<td>ITT</td>
<td>28.08 %</td>
</tr>
<tr>
<td>Holton et al. (1992)</td>
<td>32.6 (10.8)</td>
<td>CT</td>
<td>BDI</td>
<td>16</td>
<td>25</td>
<td>Individual</td>
<td>20</td>
<td>Med</td>
<td>32</td>
<td>57</td>
<td>ITT</td>
<td>41.46 %</td>
</tr>
<tr>
<td>Jordan et al. (2014)</td>
<td>36.1 (12.85)</td>
<td>CBT</td>
<td>MADRS</td>
<td>23</td>
<td>25</td>
<td>Individual</td>
<td>12</td>
<td>MCT</td>
<td>21</td>
<td>23</td>
<td>ITT</td>
<td>8.33 %</td>
</tr>
<tr>
<td>Lemmens et al. (2015)</td>
<td>41.2 (12.1)</td>
<td>CT</td>
<td>BDI</td>
<td>75</td>
<td>76</td>
<td>Individual</td>
<td>16–24</td>
<td>WCG</td>
<td>30</td>
<td>31</td>
<td>ITT</td>
<td>1.87 %</td>
</tr>
<tr>
<td>Lopes et al. (2014)</td>
<td>35.3 (11.22)</td>
<td>CBT</td>
<td>BDI</td>
<td>20</td>
<td>29</td>
<td>Individual</td>
<td>20</td>
<td>NT</td>
<td>20</td>
<td>34</td>
<td>ITT</td>
<td>36.51 %</td>
</tr>
<tr>
<td>Murphy et al. (1995)</td>
<td>38.7 (12)</td>
<td>CBT</td>
<td>BDI</td>
<td>11</td>
<td>11</td>
<td>Individual</td>
<td>–</td>
<td>Med</td>
<td>10</td>
<td>10</td>
<td>ITT</td>
<td>21 %</td>
</tr>
<tr>
<td>Power and Freeman (2012)</td>
<td>36.1 (11.3)</td>
<td>CBT</td>
<td>BDI</td>
<td>–</td>
<td>22</td>
<td>Individual</td>
<td>12–16</td>
<td>TAU</td>
<td>39</td>
<td>39</td>
<td>ITT</td>
<td>43.2 %</td>
</tr>
<tr>
<td>Ronco et al. (2017)</td>
<td>29 (7.22)</td>
<td>iCBT</td>
<td>PHQ-9</td>
<td>37</td>
<td>37</td>
<td>Online</td>
<td>6</td>
<td>MAC</td>
<td>40</td>
<td>40</td>
<td>ITT</td>
<td>18 %</td>
</tr>
<tr>
<td>Rush et al. (1977)</td>
<td>35.70 (7.22)</td>
<td>CT</td>
<td>BDI</td>
<td>18</td>
<td>19</td>
<td>Individual</td>
<td>CBT: 20 Med: 12</td>
<td>Med</td>
<td>14</td>
<td>22</td>
<td>ITT</td>
<td>21.95 %</td>
</tr>
<tr>
<td>Sava et al. (2009)</td>
<td>37 (10.24)</td>
<td>CT</td>
<td>BDI</td>
<td>49</td>
<td>48</td>
<td>Individual</td>
<td>20</td>
<td>REBT</td>
<td>48</td>
<td>48</td>
<td>ITT</td>
<td>9.5 %</td>
</tr>
<tr>
<td>Soucy et al. (2017)</td>
<td>32.47 (2.6)</td>
<td>BA</td>
<td>BDI</td>
<td>20</td>
<td>20</td>
<td>Guided self-help</td>
<td>–</td>
<td>WCG</td>
<td>20</td>
<td>20</td>
<td>ITT</td>
<td>17.5 %</td>
</tr>
<tr>
<td>Wilson et al. (1983)</td>
<td>39.5 (12.1)</td>
<td>CT</td>
<td>BDI</td>
<td>–</td>
<td>8</td>
<td>Individual</td>
<td>8</td>
<td>WCG</td>
<td>–</td>
<td>9</td>
<td>ITT</td>
<td>–</td>
</tr>
</tbody>
</table>

* ACT = Acceptance and Commitment Therapy; BA = Behavioral Activation; CBT = Cognitive Behavioral Therapy; CT = Cognitive Therapy; DT = Dynamic Psychotherapy; iCBT = internet-based cognitive–behavioral therapy; IPT = Interpersonal Therapy; MAC = monitored attention control; MCT = Metacognitive therapy; Med = medication; PEP = psycho-educational prevention program; NT = Narrative Therapy; PPI = Positive Psychology Intervention; PST = Psychodynamic Supportive Therapy; REBT = Rational-Emotive Behaviour Therapy; ST = Schema Therapy; TAU = Treatment As Usual; WCG = Waitlist Control Group.

BDI = Beck Depression Inventory, HAM-D = Hamilton Depression Rating Scale, MADRS = Montgomery Åsberg Depression Rating Scale, PHQ-9 = Patient Health Questionnaire 9.
Table 2
Selected characteristics of studies involving older adults.

<table>
<thead>
<tr>
<th>Study first-named author (date)</th>
<th>Age Mean (SD)</th>
<th>CBT type(^a)</th>
<th>Psychometric instrument(^b)</th>
<th>N post-treatment</th>
<th>N included in analysis</th>
<th>Format</th>
<th>Number of sessions</th>
<th>Control type(^a)</th>
<th>N post-treatment</th>
<th>N included in analysis</th>
<th>Analyses</th>
<th>Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arean et al. (1993)</td>
<td>66.4 (7.43)</td>
<td>PST</td>
<td>BDI</td>
<td>19</td>
<td>19</td>
<td>Group</td>
<td>12</td>
<td>WCG</td>
<td>20</td>
<td>20</td>
<td>Completer</td>
<td>18.75 %</td>
</tr>
<tr>
<td>Beutler et al. (1987)</td>
<td>70.7 (4.02)</td>
<td>CT + PLA</td>
<td>BDI</td>
<td>–</td>
<td>16</td>
<td>Group</td>
<td>19</td>
<td>PLA</td>
<td>–</td>
<td>15</td>
<td>ITT</td>
<td>53.57 %</td>
</tr>
<tr>
<td>Floyd et al. (2004)</td>
<td>68</td>
<td>CPT</td>
<td>HAM-D</td>
<td>8</td>
<td>8</td>
<td>Individual</td>
<td>12-20</td>
<td>WCG</td>
<td>14</td>
<td>14</td>
<td>Completer</td>
<td>26.67 %</td>
</tr>
<tr>
<td>Gallagher and Thompson (1982)</td>
<td>67.77 (6.07)</td>
<td>CT</td>
<td>BDI</td>
<td>9</td>
<td>9</td>
<td>Individual</td>
<td>16</td>
<td>BT</td>
<td>5</td>
<td>10</td>
<td>ITT</td>
<td>30 %</td>
</tr>
<tr>
<td>Laidlaw et al. (2008)</td>
<td>74.03 (8.01)</td>
<td>CBT alone(^c)</td>
<td>BDI</td>
<td>20</td>
<td>20</td>
<td>Individual</td>
<td>2-17</td>
<td>TAU</td>
<td>20</td>
<td>23</td>
<td>ITT</td>
<td>9.09 %</td>
</tr>
<tr>
<td>Moss et al. (2012)</td>
<td>77.5 (6.72)</td>
<td>BA (Bib)</td>
<td>HAM-D</td>
<td>13</td>
<td>13</td>
<td>Supported self-help</td>
<td>4</td>
<td>DTCG</td>
<td>10</td>
<td>13</td>
<td>ITT</td>
<td>11.54 %</td>
</tr>
<tr>
<td>Scogin, Jamison &amp; Gochnaur (1989)</td>
<td>68.3</td>
<td>CT (Bib)</td>
<td>HAM-D</td>
<td>22</td>
<td>21</td>
<td>Self-help</td>
<td>4</td>
<td>DTCG</td>
<td>22</td>
<td>21</td>
<td>Completer</td>
<td>44 %</td>
</tr>
<tr>
<td>Serfaty et al. (2009)</td>
<td>74.1 (7.0)</td>
<td>TAU + CBT</td>
<td>BDI</td>
<td>64</td>
<td>70</td>
<td>Individual</td>
<td>12</td>
<td>TAU</td>
<td>55</td>
<td>67</td>
<td>ITT</td>
<td>13.14 %</td>
</tr>
<tr>
<td>Steuer et al. (1984)</td>
<td>Age mean and SD not specified, age range = 55 to 78 years; median age = 66</td>
<td>CBT</td>
<td>BDI</td>
<td>10</td>
<td>16</td>
<td>Individual</td>
<td>46</td>
<td>PDT</td>
<td>10</td>
<td>17</td>
<td>ITT</td>
<td>39.39 %</td>
</tr>
<tr>
<td>Thompson and Gallagher (1984)</td>
<td>67</td>
<td>CT</td>
<td>BDI</td>
<td>8</td>
<td>8</td>
<td>Group</td>
<td>10</td>
<td>WCG</td>
<td>8</td>
<td>8</td>
<td>Completer</td>
<td>40.74 %</td>
</tr>
<tr>
<td>Thompson et al. (1987)</td>
<td>66.88 (6.17)</td>
<td>CT</td>
<td>BDI</td>
<td>17</td>
<td>17</td>
<td>Individual</td>
<td>16-20</td>
<td>BPT</td>
<td>17</td>
<td>17</td>
<td>Completer</td>
<td>27.66 %</td>
</tr>
<tr>
<td>Thompson et al. (2001)</td>
<td>66.8 (5.9)</td>
<td>CBT</td>
<td>BDI</td>
<td>24</td>
<td>31</td>
<td>Individual</td>
<td>16-20</td>
<td>Med</td>
<td>21</td>
<td>33</td>
<td>ITT</td>
<td>29.69 %</td>
</tr>
</tbody>
</table>

\(^a\)BA = Behavioral Activation; Bib = bibliography; BPT = Brief Psychodynamic Therapy; BT = behavioral therapy; CBT = Cognitive Behavioral Therapy; CPT = Cognitive Psychotherapy; CT = Cognitive Therapy; DTCG = Delayed Treatment Control Group; Med = pharmacotherapy; PLA = Placebo Control Group; PDT = Psychodynamic group psychotherapy; PST = Problem-Solving Therapy; TAU = Treatment As Usual; WCG = Waitlist Control Group.

\(^b\)BDI = Beck Depression Inventory, HAM-D = Hamilton Depression Rating Scale.

\(^c\)CBT alone is psychological therapy allocation where participants are not receiving medication concurrently.
2.4.2. CBT efficacy compared to other conditions with intention to treat studies only

Analysing CBT efficacy with ITT design studies only included 32 studies (25 adult and 7 OA) and showed Hedges’ $g = 0.45$ (95% CI = 0.24 to 0.67), a small effect favouring CBT treatment ($Z = 4.16, p < .001$). A moderator analysis comparing adults and OA using ITT papers only showed a non-significant difference ($Q(1) = 0.15, p = .70, I^2 = 86\%$) indicating no difference between the two age groups in the favourable effect found for CBT.
A test of heterogeneity between the effect size of all studies and effect size of only ITT studies was non-significant ($Q (1) = 0.04, p = .84, I^2 = 0 \%$). The nonsignificant $p$-value indicates no statistically significant difference is found between the effect found for all studies and the effect for ITT studies only.

### 2.4.3. Adults versus OA in CBT efficacy compared to other conditions

A moderator analysis of adults and OA outcomes of CBT compared to other conditions (using post-treatment data) was carried out to investigate differences between the age groups. Adult papers showed a medium overall effect size ($g = 0.49, 95 \% CI = 0.22-0.77$), significantly favouring CBT over other treatments ($Z = 3.53, p < .001$). However, heterogeneity among adult papers was considerable ($Q (24) = 140.11, P < .001, I^2 = 90 \%$). OA papers showed a small effect ($g = 0.45, 95 \% CI = 0.24-0.67$), significantly favouring CBT over other conditions ($Z = 4.14, p < .001$). Heterogeneity among OA papers was considered not significant ($Q (11) = 12.60, P < .32, I^2 = 22 \%$). Fig. 3 shows effect sizes for studies included in analysis for adults and OA individually, displays the random effects model for each age group and overall random effects model, and displays the test for difference between the two age groups. Funnel plots can be viewed for adults and OA studies individually in Figs. 4 and 5. When investigating heterogeneity ($Q$) between age groups, there is no statistically significant difference ($Q_{between} (df = 1) = 0.06, p = .89$) between adults and OA for CBT outcome compared to other treatment conditions.

### 2.4.4. Sensitivity analysis of CBT efficacy compared to active control condition

Using post-treatment data, a sensitivity analysis was completed with a random effects model and compared CBT efficacy to active control conditions (psychotherapy and pharmacotherapy) groups only. This analysis involved 20 papers (16 adults and 4 OA) which included a CBT group and an active treatment comparison group. A non-significant effect ($z = 1.96, p = .05$) was found for adults, Hedges’ $g = 0.29, 95 \% CI 0.00-0.57$. Heterogeneity of adult papers was considered significant and very high at ($Q (15) = 84.18, P < .001, I^2 = 87 \%$). OA articles also showed a non-significant effect ($z = 1.85, p = .06$), with Hedges’ $g = 0.30, 95 \% CI -0.02$ to $0.62$. Heterogeneity among OA papers was considered not statistically significant and was low ($Q (3) = 1.66, p = .06, I^2 = 0 \%$). The overall effect size showed $g = 0.29 (95 \% CI 0.05$ to $0.53$), a statistically significant, small effect favouring CBT ($Z = 2.39, p < .02$). Duval and Tweedie trim-and-fill method suggested two studies are potentially missing. Including these two studies in calculations increased the overall effect size, $g = 0.63 (95 \% CI 0.39$ to $0.86$), indicating there was no evidence that publication bias produced an over-estimate of the true effect.

A moderator analysis considered whether there were differences between adults and OA studies. A test of heterogeneity between adults and OA papers found not statistically significant difference ($Q_{between} (1) = 0.03, p = .86$) between age groups when comparing CBT to active treatments only.

### 2.4.5. Sensitivity analysis of CBT efficacy compared to passive control conditions

Using post-treatment data, a random effects model compared CBT to passive control conditions only (i.e. treatment as usual, wait list, delayed treatment, or placebo). Analysis included 17 papers (nine adults and eight OA), each of which compared a CBT group to a control condition. Statistically significant results were found in both age groups, with a large effect for adults ($g = 1.00, 95 \% CI 0.44$ to $1.56, Z = 3.56, p < .001$), and medium effect for OA ($g = 0.55, 95 \% CI 0.25$ to $0.85, Z = 3.61, p < .001$). Overall, a medium effect size ($g = 0.75, 95 \% CI 0.47$ to $1.04$) was found significantly favouring CBT ($Z = 5.17, p < .001$). Duval and Tweedie trim-and-fill method did not produce a result, indicating the calculation could not usefully improve on the overall effect.

Between age group heterogeneity analysis was not significant.
Results demonstrate, based on the basis of post-treatment data, that CBT is an efficacious treatment for depression, a finding consistent with previous reviews and meta-analyses (Cuijpers et al., 2013a; Cuijpers et al., 2019; Mulder et al., 2017). However, the main aim of this meta-analysis was to compare efficacy of CBT treatment for depression between adults and OA adopting the same methodology across age ranges in a single study which is considered important (Cuijpers et al., 2020a, 2020b). When examining CBT treatment efficacy between adults and OA, there are no statistically significant differences in outcome. Our moderator analyses suggested that this was true regardless of the type of control condition that trials utilised, i.e., the lack of effect of a moderation effect for age group cannot be attributed to there being more of one type of control condition for adults or OA. This finding for CBT is consistent with broader findings of no significant differences in depression treatment efficacy for psychotherapy more broadly, when adult and OA participant data are compared (Cuijpers et al., 2009; Cuijpers et al., 2020b).

Subgroup analyses comparing age groups were conducted with a very small number of OA studies, revealing more about the quality difference that may exist in the research literatures for this age group.
Fig. 4. Funnel plot for effect sizes (Hedge’s g) from post-treatment analysis of CBT compared to other conditions for adults.

Fig. 5. Funnel plot for effect sizes (Hedge’s g) from post-treatment analysis of CBT compared to other conditions for older adults.
compared to research completed with adults. The OA studies tended to be published earlier and often did not compare to active controls, suggesting there is a need for more contemporary research trials examining the efficacy of CBT with OA. Some suggestions for this research include greater use of ITT design, recruitment of much older participants (i.e. mean age of studies should be 70+ years of age as a minimum) and greater direct comparison of outcome with other psychotherapies.

This lack of support for differential effects of CBT for depression is clinically important. Notably, whilst literature search criteria created an age group overlap, the ages of the population sample from papers in the two age groups included in this meta-analysis do not overlap. As such, practitioners working with depressed OA should expect the same treatment outcome as when working with other clinical populations. The result reported here confirms that the misconception that OA are less likely to benefit from ‘talking cures’ (Laidlaw and Wilkinson, 2020) is exactly that, an erroneous misconception, and one that harms the access of OA to evidence-based psychological treatment for depression. The lack of any difference in outcome for CBT between the age groups (adults vs OA) argues for equality of access to psychological treatment. As such, there is no logical or other reason to deny OA with depression access to evidence-based psychological therapies. The data reported here has implications in terms of training for therapists who should be expected to treat OA as well as adults. An addition to CBT training for common mental health condition should include some reference to challenging age-stereotypes and a focus on understanding normal ageing so that OA receive comparable access to psychological treatment for depression.

The finding reported here is consistent with ‘naturalistic’ outcome data reported from the improving access to psychological therapies (IAPT) national roll of evidence-based protocol delivered CBT. From national naturalistic data, reported in the IAPT minimum dataset outcomes, it is reported that OA benefit from CBT as much as adults do and with much lower attrition rates reported for OA (Chaplin et al., 2015; Pettit et al., 2017; Plotkin, 2014).

When considering the current meta-analysis reported here, lower heterogeneity (I²) is observed among OA papers compared to adult papers suggesting that there is greater consistency of outcome for OA and much greater variability in outcome for adults. This may be partially explained by the difference in indices of outcome. In particular, many of the OA CBT studies reported completer analyses rather than intention to treat analyses. Moreover, many of the OA studies are smaller and were published earlier than the adult studies. This is of concern as recent years have seen the world experience a significant demographic transition (UN, 2019), with OA are living longer and in the main living more of life disability free (Guzman-Castillo et al., 2017).

An important limitation to consider is the challenge of comparability of the adults and OA studies in that not all OA studies adopted an ITT methodology. Adapting the ITT criterion for OA studies, taking into account the context of different time periods the OA research was conducted in, and standards of research at the time the research was completed, should prevent unnecessarily excluding relevant and valid studies. Though, as many of the OA studies did not adopt ITT this may suggest there are qualitative differences in the scientific quality of these studies. This may be so, however when conducting sub-group analyses comparing the effect size of all studies to the effect size of ITT only papers (i.e. all adult studies, and only the OA studies adopting ITT) there was no statistical significant difference. Nevertheless, this adaptation demonstrates a gap in contemporary research evidence for OA, which if available could provide a more methodologically equal comparison.

Another consideration, and limitation of this meta-analysis, was the criterion for research to be carried out in Western society countries. Whilst set in order to maintain homogeneity in results, this limits generalisability of findings from analyses. Nevertheless, results offer important implications for Western society countries, and this meta-analysis prompts interest for further research to be carried out in non-Western countries so as to investigate these effects on a wider scale.

Many of the earlier OA studies recruited much younger participants (i.e. in their 60s and 70s) than one is likely see in clinical settings. This is an important consideration as it suggests there is a need to have more RCTs that consider the efficacy of CBT for people who are aged more than 75–85 years. As this is the fastest growing population (UN, 2019) grouping in most of the developed and developing world, there is an urgent need for CBT to be evaluated as a treatment of late life with the oldest-old (those aged 85 years plus). This is not just a consideration of chronological age but a recognition that the oldest-old are likely to have higher levels of physical comorbidity and to have faced great challenges associated with ageing. More research is therefore urgently needed to consider how CBT can meet the needs of the new cohort of OA. Research with the oldest-old can act as an important guide for therapists in clinical practices who are undoubtedly likely to be asked to treat the oldest-old. This important omission of research knowledge highlighted by the MA reported here should not be allowed to persist.

Some caution should be exercised when interpreting the results due to the heterogeneity of papers in the meta-analysis which suggests variability of findings. Moderator analyses showed relative low heterogeneity for OA studies compared to high heterogeneity for adult studies, suggesting overall heterogeneity is largely influenced by adult studies and offering a more robust finding for OA studies. Nevertheless, our publication bias enquiries suggested effects found in the analyses were robust and this meta-analysis does, therefore, not appear to overestimate outcomes, especially within the OA studies included in this review (Cuijpers et al., 2013a).

4. Conclusion

This meta-analysis detected no statistically significant difference between adults and OA for CBT efficacy in depression treatment, challenging clinical misconceptions related to working with OA. Practitioners applying CBT with OA can expect equivalent outcomes for this age group to that achievable with adults. As such there can be no justification to deny access to psychological therapy to OA.

Research included in this meta-analysis resulted in methodological differences as many studies reporting outcome with OA did not report ITT. Some caution with interpreting results reflect considerations such as overlapping age criteria between adults and OA, and participants in the OA studies are relatively young given contemporary demographic changes. In addition to a demand for more contemporary research involving OA, this current meta-analysis concludes an urgent need for the evaluation of CBT in late life, particularly with oldest-old (aged 85 years plus).

Role of funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

None.

Acknowledgements

None.

References


Alessa Werson is a Graduate Research Assistant at the University of Exeter where she works on enhancing resilience and widening diversity in the IAPT workforce. She is also an Honorary Assistant Psychologist at Livewell Southwest at Insight, the Early Intervention Psychosis service. As a qualified PWP, she also trains and supervises practitioners with strength-based CBT as Honorary Clinical Supervisor at the University of Exeter.

Richard Meiser-Stedman is a Professor of Clinical Psychology and based at the University of East Anglia. He conducts research into how children and young people (and their families) cope after traumas (e.g. car accidents, assaults, abuse), and what might help them to recover more quickly. He is a co-investigator on the PYCES trial (the treatment of PTSD in preschool children) and the PROTECT study (looking at how families shape children’s responses to trauma).

Ken Laidlaw is a Professor of Clinical Psychology and Programme Director for the Doctoral Programme in Clinical Psychology as well as Director of postgraduate research in CEDAR at the College of Life and Environmental Sciences, University of Exeter. Ken does research with older people and is interested in Health Psychology, Abnormal Psychology and Clinical Psychology. He is especially interested in quality of life and wisdom research in later life in terms of how this research data can be integrated into CBT to augment treatment outcome.