

# Psychotherapy for Depression Across Different Age Groups: A Systematic Review and Meta-analysis

Pim Cuijpers, PhD; Eirini Karyotaki, PhD; Dikla Eckshtain, PhD; Mei Yi Ng, PhD; Katherine A. Corteselli, MA; Hisashi Noma, PhD; Soledad Quero, PhD; John R. Weisz, PhD

**IMPORTANCE** It is not clear whether psychotherapies for depression have comparable effects across the life span. Finding out is important from a clinical and scientific perspective.

**OBJECTIVE** To compare the effects of psychotherapies for depression between different age groups.

**DATA SOURCES** Four major bibliographic databases (PubMed, PsychINFO, Embase, and Cochrane) were searched for trials comparing psychotherapy with control conditions up to January 2019.

**STUDY SELECTION** Randomized trials comparing psychotherapies for depression with control conditions in all age groups were included.

**DATA EXTRACTION AND SYNTHESIS** Effect sizes (Hedges  $g$ ) were calculated for all comparisons and pooled with random-effects models. Differences in effects between age groups were examined with mixed-effects subgroup analyses and in meta-regression analyses.

**MAIN OUTCOMES AND MEASURES** Depressive symptoms were the primary outcome.

**RESULTS** After removing duplicates, 16 756 records were screened and 2608 full-text articles were screened. Of these, 366 trials (36 702 patients) with 453 comparisons between a therapy and a control condition were included in the qualitative analysis, including 13 (3.6%) in children (13 years and younger), 24 (6.6%) in adolescents ( $\geq 13$  to 18 years), 19 (5.2%) in young adults ( $\geq 18$  to 24 years), 242 (66.1%) in middle-aged adults ( $\geq 24$  to 55 years), 58 (15.8%) in older adults ( $\geq 55$  to 75 years), and 10 (2.7%) in older old adults (75 years and older). The overall effect size of all comparisons across all age groups was  $g = 0.75$  (95% CI, 0.67-0.82), with very high heterogeneity ( $I^2 = 80\%$ ; 95% CI: 78-82). Mean effect sizes for depressive symptoms in children ( $g = 0.35$ ; 95% CI, 0.15-0.55) and adolescents ( $g = 0.55$ ; 95% CI, 0.34-0.75) were significantly lower than those in middle-aged adults ( $g = 0.77$ ; 95% CI, 0.67-0.87). The effect sizes in young adults ( $g = 0.98$ ; 95% CI, 0.79-1.16) were significantly larger than those in middle-aged adults. No significant difference was found between older adults ( $g = 0.66$ ; 95% CI, 0.51-0.82) and those in older old adults ( $g = 0.97$ ; 95% CI, 0.42-1.52). The outcomes should be considered with caution because of the suboptimal quality of most of the studies and the high levels of heterogeneity. However, most primary findings proved robust across sensitivity analyses, addressing risk of bias, target populations included, type of therapy, diagnosis of mood disorder, and method of data analysis.

**CONCLUSIONS AND RELEVANCE** Trials included in this meta-analysis reported effect sizes of psychotherapies that were smaller in children than in adults, probably also smaller in adolescents, that the effects may be somewhat larger in young adults, and without meaningful differences between middle-aged adults, older adults, and older old adults.

JAMA Psychiatry. 2020;77(7):694-702. doi:10.1001/jamapsychiatry.2020.0164  
Published online March 18, 2020.

← Editorial page 667

+ Supplemental content

+ CME Quiz at  
jamacmelookup.com  
and CME Questions page 772

**Author Affiliations:** Author affiliations are listed at the end of this article.

**Corresponding Author:** Pim Cuijpers, PhD, Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Van der Boechorststraat 7-9, Amsterdam, Noord-Holland 1081 BT, the Netherlands (p.cuijpers@vu.nl).

Hundreds of randomized clinical trials (RCTs) have examined the effects of psychological treatments for depression. However, most of this research has been conducted separately in children and adolescents,<sup>1-3</sup> in younger<sup>4</sup> and middle-aged adults,<sup>5-8</sup> and in older adults.<sup>9-12</sup> Therefore, it is not known whether therapies have comparable effects across different age groups. Although some research has focused on differential effects of therapies in younger and older adults,<sup>13</sup> to our knowledge, there is no meta-analysis focusing on psychotherapies across the age range from children and adolescents to younger and older adults.

It is important to study effects of therapies across age groups for several reasons. First, differences identified between age groups can inform clinicians about the potential of treatments across age groups and may help inform treatment selection. Second, differential effects may also indicate differences between age groups in what procedures are required for symptom reduction, in the working mechanisms of the therapies, and even in the psychological processes involved in depression.

We conducted a systematic review and meta-analysis of psychotherapies to examine whether the reported outcomes differ by age group. In the past 15 years, we have conducted a series of meta-analyses on psychotherapies for children and adolescents on the one hand,<sup>1,2,14,15</sup> and adults and older adults on the other hand.<sup>7-9,16</sup> The child and adolescent meta-analyses have shown relatively modest mean effect sizes, whereas the adult and older adult meta-analyses on depression treatment have shown more substantial effect sizes. However, such differences can only be fairly tested when the same methods of analysis are applied consistently across age groups. By combining our databases and methods, we are able to examine the reported outcomes of psychotherapies for depression across the life span and explore potential associations.

## Methods

### Identification and Selection of Studies

We used 2 existing databases of studies on the psychotherapies for depression, 1 in adults and 1 in youths. The database in adults was used as the starting point of this meta-analysis because it is the larger one, and the studies identified through the second database were added to it. It has been described in detail elsewhere.<sup>17</sup> Four bibliographic databases were searched (PubMed, PsycINFO, Embase, and Cochrane) by combining terms indicative of depression and psychotherapies, with filters for RCTs. The full search string for PubMed is given in eAppendix 1 in the [Supplement](#). For this meta-analysis, we also checked the references of studies that were excluded from the adult database because of their focus on youths. The database is continuously updated (last update: January 1, 2019). All records were screened independently by 2 researchers, and all articles that met initial inclusion criteria were retrieved as full text. The decision to include or exclude a study was also made independently by the 2 researchers, and disagreements were resolved through discussion.

### Key Points

**Question** Do psychotherapies for depression have comparable outcomes in age groups across the life span?

**Findings** In a meta-analysis of 366 randomized clinical trials including 36 072 patients comparing psychotherapy with control conditions, psychotherapies had lower effect sizes in children and adolescents compared with adults, and no significant differences were found between middle-aged and older adults. However, conclusions are not definitive, given the low quality of many studies, the risk of publication bias, and the high heterogeneity among the studies.

**Meaning** There is a need to improve psychotherapies in children and adolescents.

The second database focused on studies of psychotherapies for depression in children and adolescents. The searches were conducted in PubMed and PsycINFO.<sup>1</sup> The searches were conducted up to December 2017 and updated to January 1, 2019, through the searches for the adult database.

We included studies that were: (1) RCTs; (2) in which a psychological treatment; (3) for depression in any age group; and (4) was compared with a control group (waiting list, usual care, pill placebo, or other control condition such as a brochure or a general discussion group). Depression could be established with a diagnostic interview (any of the versions of the *DSM* or *International Statistical Classification of Diseases and Related Health Problems* or the Research Diagnostic Criteria) or with a score greater than a cutoff on a standardized measure. We allowed different treatment formats, including individual, group, telephone, and guided self-help (through the internet or not).<sup>16</sup> We included studies on treatment-resistant and chronic depression because the patients in these studies do meet criteria for depression.<sup>18</sup> We excluded studies on bipolar and psychotic depression, on self-guided therapies without any professional support, on inpatients, and maintenance studies.

### Quality Assessment and Data Extraction

In line with previous meta-analyses,<sup>4,5,8,11</sup> we assessed the risk of bias of included studies using 4 criteria of the risk of bias assessment tool, developed by the Cochrane Collaboration<sup>19</sup>: adequate generation of allocation sequence; concealment of allocation to conditions; masking of assessors; and dealing with incomplete outcome data (this was assessed as positive when intention-to-treat analyses were conducted). All 4 items were rated as positive (the criterion was met) or negative (the criterion was not met or it was unclear). The total risk of bias score for each study was calculated as the sum of all positive scores (range 0 to 4, with 4 indicating no risk of bias). Because it is not clear whether the use of self-report measures should be rated as positive or negative,<sup>20</sup> we used 2 ways of assessing the total risk of bias score, one in which the use of self-report was rated as positive and one in which it was rated negative. Assessment of risk of bias was conducted independently by 2 researchers, and disagreements were resolved through discussion.

Table 1. Comparison of Effect Sizes of Studies of Psychotherapies Compared With Control Groups and Across Age Groups: Hedges  $g^a$ 

Comparisons	No. of Studies	$g$ (95% CI)	$I^2$ (95% CI)	$P$ Value <sup>b</sup>
All comparisons	453	0.75 (0.67-0.82)	80 (78-82)	NA
1 ES per study <sup>c</sup>				
Highest	366	0.76 (0.68-0.85)	83 (81-84)	NA
Lowest	366	0.69 (0.61-0.78)	82 (80-83)	NA
Outliers excluded ( $g \geq 1.5$ )	399	0.55 (0.51-0.59)	64 (60-67)	NA
Adjusted for publication bias <sup>d</sup>	580	0.43 (0.33-0.52)	86 (85-87)	NA
Only low RoB <sup>e</sup>	120	0.51 (0.42-0.60)	76 (72-80)	NA
Alternative <sup>e</sup>	25	0.47 (0.28-0.67)	69 (53-80)	NA
One outcome per study	453	0.77 (0.69-0.85)	81 (79-83)	NA
Specific age categories <sup>f</sup>				
Children	15	0.35 (0.15-0.55)	29 (0-62)	<.001
Adolescents	28	0.55 (0.34-0.75)	79 (71-85)	
Adults				
Young	27	0.98 (0.79-1.16)	43 (9-64)	
Middle-aged	304	0.77 (0.67-0.87)	81 (79-83)	
Older	69	0.66 (0.51-0.82)	80 (75-84)	.002
Older old	10	0.97 (0.42-1.52)	89 (81-93)	
Main age categories <sup>g</sup>				
Children and adolescents	43	0.48 (0.34-0.63)	72 (62-79)	.002
Early-middle adults	331	0.79 (0.70-0.89)	80 (78-82)	
Late adults	79	0.71 (0.56-0.85)	82 (78-85)	

Abbreviations: ES, effect size; NA, not applicable; RoB, risk of bias.

<sup>a</sup> According to the random-effects model.

<sup>b</sup> The  $P$  value indicates the significance of the difference between subgroups.

<sup>c</sup> In 72 studies, more than 1 psychotherapy arm was included (61 had 2 psychotherapy arms, 8 had 3 arms, and 3 had 4 arms). Because these effect sizes were not independent of each other, they may have artificially reduced heterogeneity and influenced the effect sizes. We conducted 2 separate analyses to examine this possibility. In the first analysis, we only included the largest effect size from each study, and in the second only the smallest effect size.

<sup>d</sup> Egger test was significant ( $P < .001$ ), and the number of imputed studies using Duvall and Tweedie trim-and-fill procedure was 126.

<sup>e</sup> We calculated low RoB in 2 ways, 1 in which the use of self-reported outcomes was rated as positive and an alternative one in which the use of self-reported outcomes was rated as negative.

<sup>f</sup> In pairwise comparisons, the difference between children and adolescents was not significant ( $Q_1 = 2.10$ ;  $P = .14$ ), but the difference between children and all other age categories was significant (young adults:  $Q_1 = 23.80$ ;  $P < .001$ ; adults:  $Q_1 = 15.59$ ;  $P < .001$ ; older adults:  $Q_1 = 6.70$ ;  $P = .01$ ; older old adults:  $Q_1 = 5.70$ ;  $P = .02$ ). Adolescents differed significantly from young adults ( $Q_1 = 10.35$ ;  $P = .001$ ) and adults ( $Q_1 = 3.96$ ;  $P = .05$ ), but there was no significant difference with older adults ( $Q_1 = 0.85$ ;  $P = .36$ ) and older old adults ( $Q_1 = 2.61$ ;  $P = .11$ ). Young adults differed from adults ( $Q_1 = 3.93$ ;  $P = .05$ ) and older adults ( $Q_1 = 6.99$ ;  $P = .01$ ), but not from older old adults ( $Q_1 = 0.00$ ;  $P = .99$ ). Adults did not differ significantly from older adults ( $Q_1 = 1.31$ ;  $P = .25$ ) and older old adults ( $Q_1 = 0.65$ ;  $P = .42$ ) and older adults did not differ significantly from older old adults ( $Q_1 = 1.46$ ;  $P = .23$ ).

<sup>g</sup> Pairwise comparisons indicated a significant difference between children and adolescents and adults ( $Q_1 = 12.21$ ;  $P < .001$ ) and between children and adolescents and older adults ( $Q_1 = 4.43$ ;  $P = .04$ ), but not between adults and older adults ( $Q_1 = 0.96$ ;  $P = .33$ ).

We examined age of the participants in different ways. First, we extracted the mean age of the sample. Second, we categorized the studies into 6 specific age categories: (1) children (mean age  $\leq 13$  years); (2) adolescents (13-18 years); (3) young adults (18-24 years); (4) middle-aged adults (24-55 years); (5) older adults (55-75 years); and (6) older old adults (75 years). Third, we clustered these 6 categories into 3 main age categories: youths (children and adolescents); early-middle adults (young adults and middle-aged adults); and late adults (older adults and older old adults). The other characteristics of the participants, interventions, and the study are presented in Table 1 (definitions and categories can be found elsewhere).<sup>17</sup>

## Outcome Measures

For each comparison between a psychotherapy and a control condition, the effect size indicating the difference in

depression between the 2 groups at posttest was calculated (Hedges  $g$ ).<sup>21</sup> Depressive symptoms were the primary outcome, and we used all measures examining depressive symptoms. Effect sizes were calculated by subtracting (at posttest) the mean score of the psychotherapy group from the mean score of the control group, divided by the pooled standard deviation at posttest, and corrected for small sample bias.<sup>21</sup> If means and standard deviations were not reported, we used the procedures of the Comprehensive Meta-Analysis (CMA) software, version 3.3070, to calculate the effect size using dichotomous outcomes, or other statistics (such as  $t$  value or  $P$  value).

If 1 study included more than 1 depression measure, we pooled the effect sizes within the study before pooling the effect sizes across studies (assuming a correlation of 1.0 between measures).<sup>22</sup> We conducted sensitivity analyses in which we included only 1 outcome measure for each study

(the algorithm for selecting measures is given in eAppendix 2 in the [Supplement](#)). All effect sizes were calculated in CMA.

### Meta-analyses

We pooled the effect sizes using the “meta” and “metafor” packages in R, and ran all analyses in R studio, version 1.1.463, for Mac (the R Foundation). We used a random-effects pooling model in all analyses. We pooled the effect sizes using the inverse variance method, with the Hartung-Knapp adjustment for the random-effects model. We calculated the  $I^2$  statistic and its 95% confidence interval to estimate heterogeneity.<sup>23</sup> We examined the risk of publication bias through the Egger test<sup>24</sup> and through the Duval and Tweedie trim-and-fill procedure,<sup>25</sup> which yields an estimate of the effect size after the publication bias has been taken into account.

We conducted subgroup analyses according to the mixed-effects model, in which effect sizes within subgroups are pooled according to the random-effects model and differences between subgroups are tested with a fixed-effects model. We conducted bivariate meta-regression analyses (with mean age as the variable associated with the effect size) and multivariate meta-regression analyses to examine the association between the mean age of the study participants and the effect size, while adjusting for other characteristics of the studies. We conducted a series of sensitivity analyses in which we included only studies with low risk of bias, only studies examining cognitive behavioral therapy, only studies with patients meeting diagnostic criteria for a mood disorder, and one in which specific target groups were excluded. In addition, we conducted a sensitivity analysis that used the multilevel meta-analysis model, which accounts for the dependency within studies.<sup>26</sup> We also calculated the relative risk (RR) of dropping out from the study (for any reason) for all comparisons. All tests were 2-sided, and *P* values were considered significant when they were smaller than .05.

## Results

### Selection and Inclusion of Studies

We examined 16 756 records, retrieved 2608 full-text articles, and excluded 2242. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart describing the inclusion process, with reasons for exclusion, is presented in eAppendix 3 in the [Supplement](#). Three hundred sixty-six RCTs (with 453 comparisons between a psychotherapy and a control group) met inclusion criteria. In these trials, 36 702 patients participated (19 544 in the treatment and 17 158 in the control conditions). Key characteristics of the included studies are given in eAppendix 4 and references in eAppendix 5 in the [Supplement](#).

### Characteristics of Included Studies

Of the 366 studies, 13 (3.6%) focused on children, 24 (6.6%) on adolescents, 19 (5.2%) on young adults (mostly college students), 242 (66.1%) on middle-aged adults, 58 (15.8%) on older adults, and 10 (2.7%) on older old adults. Main characteristics

of the studies are presented in eAppendix 6 in the [Supplement](#), separately for youths, early-middle adults, and late adults. A more detailed description can be found in eAppendix 7 in the [Supplement](#).

We examined whether the studies differed for key characteristics across age groups and found that the main age groups did indeed differ significantly on recruitment method, whether participants met criteria for a depressive disorder, and country where the study was conducted (eAppendix 6 in the [Supplement](#)). Regarding the intervention characteristics, we found that treatment format and number of treatment sessions differed significantly across age groups.

### Risk of Bias

The risk of bias in most of the included studies was considerable. One hundred ninety-nine of 366 studies (54.4%) reported an adequate sequence generation. One hundred sixty-four studies (44.8%) reported allocation to conditions by an independent (third) party. Most studies used self-report measures (including parent-report measures of youth symptoms) as outcome (*n* = 240 [65.6%]), some conducted masked outcome assessments (*n* = 97 [26.5%]), and the remainder did not mask outcome assessors or did not clarify whether they were masked (*n* = 29 [7.9%]). In 220 studies (60.1%), intention-to-treat analyses were conducted. Only 106 studies (29%) met all quality criteria. Another 161 studies (44.0%) met 2 or 3 criteria, and the 99 remaining studies (27.0%) met no or only 1 criterion. Only 21 studies (5.7%) met all quality criteria when self-report measures were rated as negative for blinded outcomes assessments.

### Overall Effects of Psychotherapy

The overall effect size of all comparisons across all age groups was *g* = 0.75 (95% CI, 0.67-0.82), with very high heterogeneity ( $I^2$  = 80%; 95% CI: 78-82). The results are presented in Table 1.

We further examined these main results in several ways (Table 1). We examined the associations of multiple comparisons in 1 trial, we excluded outliers, we limited the studies to those with low risk of bias, we conducted the analyses using an alternative approach to select outcome measures (eAppendix 2 in the [Supplement](#)), and we examined publication bias. The studies with low risk of bias resulted in a considerably smaller effect size (*g* = 0.51; 95% CI, 0.42-0.60), and heterogeneity was still high ( $I^2$  = 76%; 95% CI, 72-80). Adjusting for publication bias resulted in an effect size of *g* = 0.43 (95% CI, 0.33-0.52;  $I^2$  = 80%; 95% CI, 85-87), with 126 imputed studies.

Type of control group was associated with effect size ( $Q_2$  = 16.66; *P* for difference < .001). Therefore, we also examined the differences between the main age categories according to type of control group (Table 2). We did not find significant differences for the effect sizes across the main age categories within the studies using a waiting-list control group, but we did find highly significant differences when usual care or other control groups were used. In studies with usual care and other control groups, smaller effects were found for children and adolescents in most analyses.

Table 2. Effects of Psychotherapies Across Different Types of Control Groups and Age Categories: Hedges  $g^a$ 

Comparisons	Waiting List				Usual Care				Other Control			
	No. of Studies	<i>g</i> (95% CI)	<i>I</i> <sup>2</sup> (95% CI)	<i>P</i> Value <sup>b</sup>	No. of Studies	<i>g</i> (95% CI)	<i>I</i> <sup>2</sup> (95% CI)	<i>P</i> Value <sup>b</sup>	No. of Studies	<i>g</i> (95% CI)	<i>I</i> <sup>2</sup> (95% CI)	<i>P</i> Value <sup>b</sup>
All comparisons	183	0.91 (0.81 to 1.02)	76 (72 to 79)	NA	195	0.61 (0.52 to 0.71)	81 (78 to 83)	NA	75	0.71 (0.42 to 0.99)	83 (79 to 86)	NA
1 ES/study												
Highest	133	0.95 (0.82 to 1.08)	80 (77 to 83)	NA	171	0.64 (0.53 to 0.75)	82 (79 to 84)	NA	62	0.72 (0.37 to 1.07)	85 (81 to 88)	NA
Lowest	133	0.83 (0.71 to 0.95)	78 (75 to 82)	NA	171	0.61 (0.50 to 0.72)	82 (80 to 84)	NA	62	0.64 (0.30 to 0.97)	82 (78 to 86)	NA
Outliers excluded ( <i>g</i> ≥1.5)	151	0.67 (0.61 to 0.73)	50 (39 to 58)	NA	181	0.49 (0.43 to 0.56)	69 (64 to 73)	NA	67	0.44 (0.35 to 0.53)	61 (49 to 70)	NA
Adjusted for publication bias	243	0.57 (0.43 to 0.70)	83 (81 to 85)	NA	238	0.36 (0.24 to 0.49)	88 (86 to 89)	NA	87	0.41 (0.03 to 0.79)	88 (86 to 90)	NA
Only low risk of bias	42	0.63 (0.49 to 0.77)	79 (72 to 84)	NA	59	0.49 (0.34 to 0.63)	74 (67 to 80)	NA	19	0.33 (0.18 to 0.47)	64 (41 to 78)	NA
Specific age categories												
Children	7	0.68 (0.28 to 1.08)	11 (0 to 74)	.52	5	0.13 (−0.07 to 0.33)	0 (0 to 68)	<.001	3	0.28 (0.25 to 0.31)	0 (0 to 0)	.006
Adolescents	12	0.78 (0.35 to 1.20)	80 (67 to 88)		10	0.29 (0.12 to 0.47)	44 (0 to 73)		6	0.54 (0.02 to 1.07)	81 (58 to 91)	
Young adults	13	0.92 (0.62 to 1.22)	53 (12 to 75)		9	0.99 (0.74 to 1.23)	0 (0 to 64)		5	1.16 (0.18 to 2.15)	47 (0 to 80)	
Middle-aged adults	128	0.96 (0.82 to 1.10)	78 (74 to 82)		134	0.63 (0.50 to 0.76)	81 (78 to 84)		42	0.67 (0.17 to 1.18)	83 (77 to 87)	
Older adults	20	0.77 (0.52 to 1.01)	64 (41 to 77)		32	0.58 (0.35 to 0.80)	79 (71 to 85)		17	0.71 (0.29 to 1.13)	87 (80 to 91)	
Older old adults	3	0.95 (−0.26 to 2.17)	80 (37 to 94)		5	0.81 (−0.06 to 1.69)	87 (73 to 94)		2	1.49 (−12.68 to 15.65)	96 (90 to 99)	
Main age categories												
Children and adolescents	19	0.74 (0.46 to 1.02)	72 (55 to 82)	.19	15	0.24 (0.11 to 0.37)	28 (0 to 61)	<.001	9	0.44 (0.13 to 0.76)	69 (38 to 85)	.29
Early-middle adults	141	0.96 (0.83 to 1.10)	77 (73 to 80)		143	0.66 (0.53 to 0.78)	80 (77 to 83)		47	0.73 (0.27 to 1.19)	82 (76 to 86)	
Late adults	23	0.80 (0.58 to 1.01)	69 (53 to 80)		37	0.61 (0.40 to 0.82)	80 (73 to 85)		19	0.79 (0.36 to 1.22)	88 (83 to 92)	

Abbreviations: ES, effect size; NA, not applicable.

<sup>a</sup> According to the random-effects model.<sup>b</sup> The  $P$  value indicates the significance of the difference between the age categories but within the subgroup of studies using the type of control group.

### Differences of Effects Across the 6 Specific Age Categories

We conducted subgroup analyses to examine whether there were differences between the effect sizes found across the 6 specific age groups (Table 2). The effect size difference between the age groups was highly significant ( $Q_5 = 29.57$ ;  $P < .001$ ). The effect size for children was the lowest ( $g = 0.35$ ), and the effect size for adolescents ( $g = 0.55$ ) was also relatively low. We found no indication that the effect sizes found for older adults ( $g = 0.66$ ) and older old adults ( $g = 0.97$ ) differed considerably from those for middle-aged adults ( $g = 0.77$ ). The effect sizes across the different age groups are graphically presented in Figure 1.

In pairwise comparisons, the difference between children and adolescents was not significant ( $Q_1 = 2.10$ ;  $P = .14$ ), but the differences between children and all other age categories were significant (Table 2). Effect sizes in adolescents differed significantly from young adults ( $Q_1 = 10.35$ ;  $P = .001$ ) and middle-aged adults ( $Q_1 = 3.96$ ;  $P = .05$ ), but there was no significant difference with older adults ( $Q_1 = 0.85$ ;  $P = .36$ ) or older

old adults ( $Q_1 = 2.61$ ;  $P = .11$ ). Young adults differed from middle-aged adults ( $Q_1 = 3.93$ ;  $P = .05$ ) and older adults ( $Q_1 = 6.99$ ;  $P = .01$ ), but not from the older old adults ( $Q_1 = 0.00$ ;  $P = .99$ ). Middle-aged adults did not differ significantly from older adults ( $Q_1 = 1.31$ ;  $P = .25$ ) and older old adults ( $Q_1 = 0.65$ ;  $P = .42$ ), and older adults did not differ significantly from older old adults ( $Q_1 = 1.46$ ;  $P = .23$ ).

### Differences of Effects Across the 3 Main Age Categories

When we collapsed the age groups in 3 main age categories (youths, early-middle adults, and late adults), the difference between the 3 groups was still significant, with the smallest effect sizes found for youths ( $g = 0.48$ ) and no clear difference between early-middle adults and late adults. Pairwise comparisons for the 3 main age categories indicated a significant difference between youths and early-middle adults ( $Q_1 = 12.21$ ;  $P < .001$ ) and between youths and older adults ( $Q_1 = 4.43$ ;  $P = .04$ ) but not between early-middle adults and late adults ( $Q_1 = 0.96$ ;  $P = .33$ ).



Figure 1. Effect Sizes of Psychotherapies in Different Age Groups

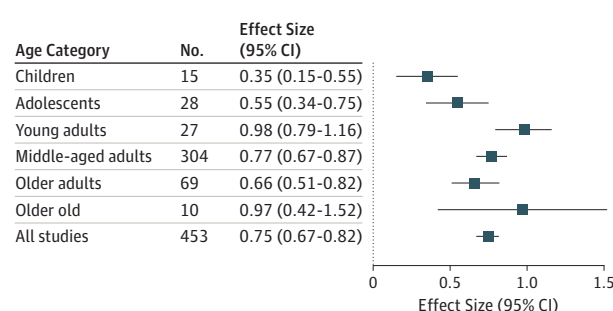
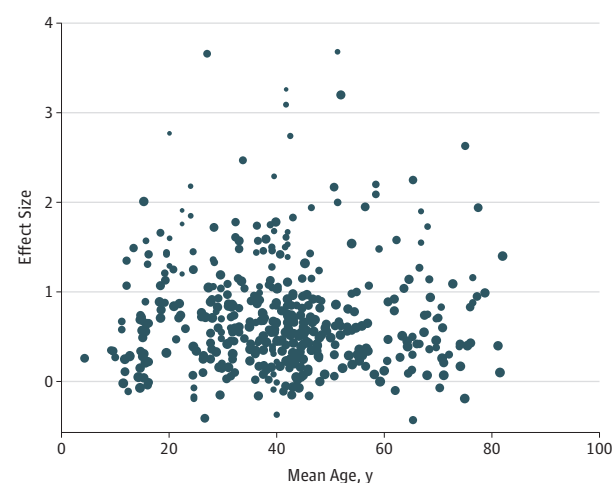


Figure 2. Bubble Plot of the Effect Size as Dependent Variable and Mean Age of the Population as Predictor



The size of the dots reflect the study sample size.

### Age as Predictor in Meta-regression Analyses

We examined the association between the effect sizes and the mean age of the study population in a series of meta-regression analyses. First, we conducted bivariate meta-regression analyses, with the effect size as the dependent variable and the mean age as the only predictor. The mean age of the population was not found to be a significant predictor (coefficient [SE], 0.0001 [0.002];  $P = .97$ ). The bubble plot is given in Figure 2.

Then we conducted multivariate meta-regression analyses with the mean age as predictor as well as the characteristics of the studies as predictors (Table 3). Mean age was still not significantly associated with the effect size. To avoid overfitting the meta-regression models, we repeated these meta-regression analyses with a (manual) stepwise backward elimination of the least significant predictor until only significant predictors remained in the model. As can be seen in Table 3 (parsimonious models), mean age was still not significantly associated with the effect size.

We also conducted multivariate meta-regression analyses with the 6 specific age categories (adults were used as the reference category; Table 3). Studies with children were significantly less effective than studies with middle-aged adults, and that was

true in the full model (coefficient [SE],  $-0.58$  [0.23];  $P = .01$ ) and the parsimonious model (coefficient [SE],  $-0.49$  [0.21];  $P = .02$ ). When we examined the 3 main age categories, studies in youths also resulted in significantly lower effect sizes in both in the full (coefficient [SE],  $-0.26$  [0.13];  $P = .01$ ) and the parsimonious model (coefficient [SE],  $-0.38$  [0.15];  $P = .04$ ).

### Sensitivity Analyses, Study Dropout, and Long-term Effects

In the sensitivity analyses (eAppendices 8-12 in the Supplement), the subgroup analyses indicating significant differences between the age groups (3 and 6 categories) were all significant or there was a trend for significance. In most analyses, the results were very similar to the main analyses. The sensitivity analyses using the multilevel meta-analysis model (eAppendix 12 in the Supplement) resulted in comparable outcomes and supported the main findings of this study.

Study dropout was significantly higher in the psychotherapy conditions compared with the control conditions (RR, 1.15; 95% CI, 1.06-1.24;  $I^2 = 40\%$ ; 95% CI, 32-48), but there were no significant differences between the 6 specific age categories nor for the 3 main age categories.

We conducted separate analyses for longer-term outcomes between 6 to 8 months after baseline, 9 to 12 months after baseline, 13 to 24 months after baseline, and more than 24 months after baseline (eAppendix 13 in the Supplement). In none of these analyses were the effects of therapies for children and adolescents significantly different from zero. Between 9 and 12 months, the differences between the main age categories were significant (with smaller effect sizes for children and adolescents;  $P = .03$ ) but not for other follow-up periods.

## Discussion

To our knowledge, this is the first meta-analysis of psychotherapy trials for depression in all age groups and the largest meta-analysis of depression psychotherapies ever conducted. We found that the effect sizes for therapies in children and adolescents were significantly smaller than those found in adults. Effect sizes were especially small in children 13 years and younger, but effect sizes for adolescents were also significantly smaller than those for adults. These differences were not supported in all meta-regression analyses in which we adjusted for other characteristics of the studies, which may indicate that the significant associations may be partly explained by other characteristics of the studies.

It is not clear how to explain the differences between children/adolescents and adults. Possibly, therapies are less effective among youths. However, these differences may also be explained by differences in study or intervention characteristics. For example, we found no significant difference between age groups when waiting lists were used as a control condition; this is important because waiting list is arguably the only control condition that is identical across all age groups. Several differences between therapies should also be noted, such as the involvement of parents in some therapies for children and adolescents and specific therapies for specific age groups, such as life review therapy for older adults. Another possibility

Table 3. Standardized Regression Coefficients of Age and Effect Sizes: Multivariate Meta-regression Analyses

Predictors	Age		6 Categories		3 Categories	
	Mean	P Value	Coefficient (SE)	P Value	Coefficient (SE)	P Value
Full models						
Mean age	0.00 (0.00)	.67	NA	NA	NA	NA
Specific age category						
Children	NA	NA	−0.58 (0.23)	.01	NA	NA
Adolescents	NA	NA	−0.22 (0.17)	.19	NA	NA
Young adults	NA	NA	0.33 (0.18)	.07	NA	NA
Middle adults	NA	NA	1 [Reference]	NA	NA	NA
Older adults	NA	NA	0.01 (0.11)	.90	NA	NA
Older old	NA	NA	0.26 (0.25)	.29	NA	NA
Main age category	NA	NA	NA	NA	NA	NA
Children and adolescents	NA	NA	NA	NA	−0.38 (0.15)	.01
Early-middle adults	NA	NA	NA	NA	1 [Reference]	NA
Late adults	NA	NA	NA	NA	0.01 (0.10)	.90
Recruitment						
Community	1 [Reference]	NA	1 [Reference]	NA	NA	NA
Clinical	−0.12 (0.12)	.32	−0.07 (0.12)	.57	−0.09 (0.12)	.43
Other	0.07 (0.11)	.53	0.19 (0.11)	.08	0.18 (0.11)	.10
Diagnosis (yes/no)	0.15 (0.08)	.08	0.17 (0.08)	.04	0.14 (0.08)	.09
Specific target group (yes/no)	−0.02 (0.10)	.85	−0.05 (0.10)	.56	−0.09 (0.10)	.36
Proportion women (continuous)	0.06 (0.19)	.77	−0.13 (0.19)	.52	−0.09 (0.19)	.65
CBT (yes/no)	0.04 (0.08)	.60	0.03 (0.08)	.63	0.04 (0.08)	.60
Format						
Individual	1 [Reference]	NA	1 [Reference]	NA	1 [Reference]	NA
Group	0.05 (0.10)	.59	0.06 (0.09)	.53	0.06 (0.09)	.53
GSH	−0.11 (0.13)	.39	−0.06 (0.13)	.66	−0.09 (0.13)	.46
Mixed/other	0.07 (0.13)	.59	0.05 (0.13)	.70	0.05 (0.13)	.71
No. of sessions (continuous)	−0.01 (0.01)	.38	−0.00 (0.01)	.68	−0.00 (0.01)	.58
Year (continuous)	0.00 (0.01)	.43	0.00 (0.00)	.80	0.00 (0.00)	.77
North America (yes/no)	−0.23 (0.09)	.01	−0.19 (0.09)	.05	−0.17 (0.09)	.08
Control group						
Waiting list	1 [Reference]		1 [Reference]		1 [Reference]	
Usual care	−0.31 (0.10)	.002	−0.32 (0.10)	<.01	−0.32 (0.10)	<.01
Other	−0.20 (0.12)	.08	−0.18 (0.11)	.12	−0.18 (0.11)	.12
Risk of bias (continuous)	−0.14 (0.04)	<.001	−0.14 (0.04)	<.001	−0.13 (0.04)	<.001
Intercept	−6.74 (9.98)	.50	−1.19 (9.78)	.90	−1.51 (9.82)	.88
Parsimonious models						
Mean age	0.00 (0.00)	.69	NA	NA	NA	NA
Specific age category						
Children	NA	NA	−0.49 (0.21)	.02	NA	NA
Adolescents	NA	NA	−0.16 (0.15)	.30	NA	NA
Young adults	NA	NA	0.39 (0.17)	.02	NA	NA
Adults	NA	NA	1 [Reference]	NA	NA	NA
Older adults	NA	NA	0.04 (0.10)	.67	NA	NA
Older old adults	NA	NA	0.28 (0.24)	.25	NA	NA
Specific age category						
Children and adolescents	NA	NA	NA	NA	−0.26 (0.13)	.04
Adults	NA	NA	NA	NA	1 [Reference]	NA
Elderly	NA	NA	NA	NA	0.03 (0.10)	.76

(continued)

Table 3. Standardized Regression Coefficients of Age and Effect Sizes: Multivariate Meta-regression Analyses (continued)

Predictors	Age		6 Categories		3 Categories	
	Mean					
	Coefficient (SE)	P Value	Coefficient (SE)	P Value	Coefficient (SE)	P Value
Recruitment: other (yes/no)	NA	NA	0.18 (0.09)	.04	NA	NA
Diagnosis (yes/no)	NA	NA	0.15 (0.08)	.05	NA	NA
Control group						
Waiting list	1 [Reference]	NA	1 [Reference]	NA	NA	NA
Usual care	−0.27 (0.08)	<.01	−0.34 (0.09)	<.001	−0.22 (0.08)	<.01
Other	−0.23 (0.11)	.04	−0.23 (0.11)	.03	NA	NA
North America (yes/no)	−0.23 (0.08)	<.01	−0.20 (0.08)	.01	−0.21 (0.08)	.01
Risk of bias (continuous)	−0.13 (0.03)	<.001	−0.14 (0.03)	<.001	−0.15 (0.03)	<.001
Intercept	1.28 (0.14)	<.001	1.22 (0.12)	<.001	1.32 (0.11)	<.001

Abbreviations: CBT, cognitive behavioral therapy; GSH, guided self-help; NA, not applicable.

is that the therapies most often used with children and adolescents are primarily age-adapted versions of therapies originally designed for adults. Those therapies may be a better fit to the needs of adults than those of young people. Another possibility is that young people's potential for recovery from depression is constrained by parental and family characteristics that youths, unlike adults, have little opportunity to escape or alter.<sup>10,11</sup> Whatever the reason, this meta-analysis brings to light potential limitations of currently available psychological treatments for depression in children and adolescents, the need for improvement of treatments for these age groups, and the need for more outcome research.

One remarkable finding was that almost all research on children and adolescents was conducted in the United States, while research in early-middle and older adults was very comparable between North America and Europe. It is difficult to understand what the causes of this difference are.

In this meta-analysis, we also found that the effect sizes of therapies were somewhat larger in young adults ( $\geq 18$  to 24 years) compared with middle-aged adults ( $\geq 24$  to 55 years). One possible reason is that many studies of young adults were conducted with college students, who may have more learning capacity and may be easier to recruit. A second reason relates to differences in the characteristics of the studies. Another option is that age may not be associated with the effects of treatment in a linear way, as is assumed by the meta-regression analyses, but that there is a curvilinear association. The data suggest that the effect sizes are small in childhood, become larger in adolescents, and grow further in early adulthood, before decreasing to more modest effect sizes in the rest of the adult populations. We did not test this because of the high heterogeneity and the fact that most studies in young adults were with college students, which may not be representative for the whole age group. However, this is certainly a possibility that should be further explored in future studies.

### Limitations

The results of this study have to be considered with caution because of several important limitations. First, although there was a large number of trials, the quality of many trials was low, and when we limited the main analyses to studies with the highest quality, the differences between the age groups were no longer significant. Second, we did not rate the number of disagreements between raters of characteristics of the studies or calculate agreement between raters. This was related to the large number of included studies and the fact that the ratings were performed over several years while building this database. Third, heterogeneity was very high in most analyses, and subgroup analyses could not explain these high levels. This suggests that the effects differed considerably across studies, and it is not clear what caused these differences. Fourth, the depression measures differed considerably across studies, with no measures that were used across all age groups. This not only contributed to the heterogeneity of the studies but also made it impossible to examine the influence of baseline severity on the effect sizes. Fifth, most studies were conducted with cognitive behavioral therapy, and the number of studies with other therapies was very limited. Sixth, because of the small number of studies, the long-term outcomes cannot be considered definitive.

### Conclusions

Despite the limitations of this meta-analysis, we can cautiously conclude that the effect sizes of studies of psychotherapy for depression are smaller in children than in adults, probably also smaller in adolescents, and may be somewhat larger in young adults, and that there are probably no meaningful differences between middle-aged adults, older adults, and older old adults.

#### ARTICLE INFORMATION

**Accepted for Publication:** January 16, 2020.

**Published Online:** March 18, 2020.

doi:10.1001/jamapsychiatry.2020.0164

**Author Affiliations:** Department of Clinical, Neuro, and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands (Cuijpers, Karyotaki); Massachusetts General

Hospital, Department of Psychiatry, Harvard Medical School, Boston (Eckstain); Department of Psychology and Center for Children and Families, Florida International University, Miami (Ng); Department of Psychology, Harvard University,



Cambridge, Massachusetts (Corteselli, Weisz); Department of Data Science, The Institute of Statistical Mathematics, Tokyo, Japan (Noma); Department of Basic, Clinical Psychology and Psychobiology, Universitat Jaume I, Castellón, Spain (Quero).

**Author Contributions:** Dr Cuijpers had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Cuijpers, Eckshtain, Quero Castellano, Weisz.

Acquisition, analysis, or interpretation of data: Cuijpers, Karyotaki, Eckshtain, Ng, Corteselli, Noma, Weisz.

Drafting of the manuscript: Cuijpers, Eckshtain, Weisz.

Critical revision of the manuscript for important intellectual content: Karyotaki, Eckshtain, Ng, Corteselli, Noma, Quero Castellano, Weisz.

Statistical analysis: Cuijpers, Karyotaki, Noma. Administrative, technical, or material support: Eckshtain, Corteselli, Weisz.

**Conflict of Interest Disclosures:** Dr Noma reported personal fees from Boehringer Ingelheim, Kyowa Hakko Kirin, and ASKA Pharmaceutical outside of the submitted work. Dr Cuijpers received expense allowances for his membership of the Board of Directors of "Mind.nl," for being Chair of the Research committee of the Dutch Council for military care and research, and for being Chair of the Mental Health Priority Area of the Wellcome Trust in London, England, in 2018. In addition, he received royalties for books he has authored or coauthored and for occasional workshops and invited addresses. Dr Weisz received payments for consulting with the Child Health and Development Institute and the National Institute of Mental Health, royalties for books he has authored and coauthored, and honoraria for workshops and invited presentations at professional meetings and conferences. No other disclosures were reported.

## REFERENCES

- Eckshtain D, Kuppens S, Ugueto A, et al. Meta-analysis: 13-year follow-up of psychotherapy effects on youth depression. *J Am Acad Child Adolesc Psychiatry*. 2020;59(1):45-63. doi:10.1016/j.jaac.2019.04.002
- Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adolescents: a meta-analysis. *Psychol Bull*. 2006;132(1):132-149. doi:10.1037/0033-2909.132.1.132
- Zhou X, Hetrick SE, Cuijpers P, et al. Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: a systematic review and network meta-analysis. *World Psychiatry*. 2015;14(2):207-222. doi:10.1002/wps.20217
- Cuijpers P, Cristea IA, Ebert DD, et al. Psychological treatment of depression in college students: a meta-analysis. *Depress Anxiety*. 2016;33(5):400-414. doi:10.1002/da.22461
- Cuijpers P, Karyotaki E, Reijnders M, Huibers MJH. Who benefits from psychotherapies for adult depression? a meta-analytic update of the evidence. *Cogn Behav Ther*. 2018;47(2):91-106. doi:10.1080/16506073.2017.1420098
- Zhang A, Franklin C, Jing S, et al. The effectiveness of four empirically supported psychotherapies for primary care depression and anxiety: a systematic review and meta-analysis. *J Affect Disord*. 2019;245:1168-1186. doi:10.1016/j.jad.2018.12.008
- van Bronswijk S, Moopen N, Beijers L, Ruhe HG, Peeters F. Effectiveness of psychotherapy for treatment-resistant depression: a meta-analysis and meta-regression. *Psychol Med*. 2019;49(3):366-379. doi:10.1017/S003329171800199X
- 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. doi:10.1080/10503307.2019.1649732
- Kirkham JG, Choi N, Seitz DP. Meta-analysis of problem solving therapy for the treatment of major depressive disorder in older adults. *Int J Geriatr Psychiatry*. 2016;31(5):526-535. doi:10.1002/gps.4358
- Jonsson U, Bertilsson G, Allard P, et al. Psychological treatment of depression in people aged 65 years and over: a systematic review of efficacy, safety, and cost-effectiveness. *PLoS One*. 2016;11(8):e0160859. doi:10.1371/journal.pone.0160859
- Cuijpers P, Karyotaki E, Pot AM, Park M, Reynolds CF III. Managing depression in older age: psychological interventions. *Maturitas*. 2014;79(2):160-169. doi:10.1016/j.maturitas.2014.05.027
- Orgeta V, Brede J, Livingston G. Behavioural activation for depression in older people: systematic review and meta-analysis. *Br J Psychiatry*. 2017;211(5):274-279. doi:10.1192/bjp.bp.117.205021
- Cuijpers P, van Straten A, Smit F, Andersson G. Is psychotherapy for depression equally effective in younger and older adults? a meta-regression analysis. *Int Psychogeriatr*. 2009;21(1):16-24. doi:10.1017/S1041610208008089
- Weisz JR, Kuppens S, Ng MY, et al. What five decades of research tells us about the effects of youth psychological therapy: a multilevel meta-analysis and implications for science and practice. *Am Psychol*. 2017;72(2):79-117. doi:10.1037/a0040360
- Weisz JR, Kuppens S, Ng MY, et al. Are psychotherapies for young people growing stronger? tracking trends over time for youth anxiety, depression, attention-deficit/hyperactivity disorder, and conduct problems. *Perspect Psychol Sci*. 2019;14:216-237. doi:10.1177/1745691618805436
- Cuijpers P, Noma H, Karyotaki E, Cipriani A, Furukawa T. Individual, group, telephone, self-help and internet-based cognitive behavior therapy for adult depression: a network meta-analysis. *JAMA Psychiatry*. 2019;76:700-707. doi:10.1001/jamapsychiatry.2019.0268
- Cuijpers P, Karyotaki E, Ciharova M. A meta-analytic database of randomized trials on psychotherapies for depression 2019. [osf.io/825c6](https://osf.io/825c6). Accessed February 5, 2020.
- Jobst A, Brakemeier EL, Buchheim A, et al. European Psychiatric Association Guidance on psychotherapy in chronic depression across Europe. *Eur Psychiatry*. 2016;33:18-36. doi:10.1016/j.eurpsy.2015.12.003
- Higgins JPT, Altman DG, Gøtzsche PC, et al; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928. doi:10.1136/bmj.d5928
- Cuijpers P, Li J, Hofmann SG, Andersson G. Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clin Psychol Rev*. 2010;30(6):768-778. doi:10.1016/j.cpr.2010.06.001
- Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis*. Academic Press; 1985.
- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis*. Wiley; 2009. doi:10.1002/9780470743386
- Borenstein M, Higgins JPT, Hedges LV, Rothstein HR. Basics of meta-analysis: I<sup>2</sup> is not an absolute measure of heterogeneity. *Res Synth Methods*. 2017;8(1):5-18. doi:10.1002/jrsm.1230
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629
- Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56(2):455-463. doi:10.1111/j.0006-341X.2000.00455.x
- Pastor DA, Lazowski RA. On the multilevel nature of meta-analysis: a tutorial, comparison of software programs, and discussion of analytic choices. *Multivariate Behav Res*. 2018;53(1):74-89. doi:10.1080/00273171.2017.1365684