Journal of Affective Disorders 375 (2025) 456–464

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

How do young-old and old-old adults benefit from cognitive-behavioral therapy compared to working-age adults? A large multicenter naturalistic study

Nicolas Wrede [a,*,1], Mareike C. Hillebrand [b,1], Anne Katrin Risch [a], Georg W. Alpers [c,d], Stephan Bartholdy [e], Eva-Lotta Brakemeier [e], Anne-Kathrin Bräscher [f], Timo Brockmeyer [g,h], Hanna Christiansen [i,j], Monika Equit [k], Lydia Fehm [l], Thomas Forkmann [m], Julia Glombiewski [n], Jens Heider [o], Sylvia Helbig-Lang [p], Andrea Hermann [q], Christiane Hermann [r], Jürgen Hoyer [s], Tina In-Albon [t], Tim Klucken [u], Tania M. Lincoln [v], Lea Ludwig [v], Ulrike Lueken [w,x], Wolfgang Lutz [y], Jürgen Margraf [z,j], Tanja Michael [aa], Patrizia Odyniec [ab], Anya Pedersen [ac], Babette Renneberg [ad], Julian Rubel [ae], Almut Rudolph [af], Henning Schöttke [ag], Brian Schwartz [y], Rudolf Stark [q], Tobias Teismann [z], Julia Velten [z], Katja Werheid [ah], Ulrike Willutzki [ab], Michael Witthöft [f], Gabriele Wilz [a]

[a] Department of Counseling and Clinical Intervention, Institute of Psychology, Friedrich Schiller University Jena, Germany
[b] Department of Molecular Psychology, Ulm University, Germany
[c] Department of Psychology, School of Social Sciences, University of Mannheim, Germany
[d] Otto-Selz-Institute, University of Mannheim, Germany
[e] Clinical Psychology and Psychotherapy, University of Greifswald, Germany
[f] Department of Clinical Psychology, Psychotherapy, and Experimental Psychopathology, Johannes Gutenberg University of Mainz, Germany
[g] Clinical Psychology and Psychotherapy, University of Goettingen, Germany
[h] Clinical Psychology and Translational Psychotherapy, University of Münster, Germany
[i] Clinical Child and Adolescent Psychology, Philipps University Marburg, Germany
[j] German Center for Mental Health (DZPG), partner site Bochum-Marburg, Germany
[k] Clinical Psychology and Psychotherapy, Saarland University, Germany
[l] Department of Psychology, Center for Psychotherapy (ZPHU), Humboldt-Universität zu Berlin, Germany
[m] Department of Clinical Psychology and Psychotherapy, University of Duisburg-Essen, Germany
[n] Pain and Psycholtherapy Research Lab, Department for Psychology, University Kaiserlautern-Landau, Germany
[o] Clinical Psychology and Psychotherapy, University Kaiserslautern-Landau, Germany
[p] Psychotherapie-Ausbildung Hamburg, Universität Hamburg, Germany
[q] Psychotherapy and Systems Neuroscience, Justus Liebig University Giessen, Germany
[r] Clinical Psychology and Psychotherapy, Justus Liebig University Giessen, Germany
[s] Clinical Psychology and Psychotherapy, Technische Universität Dresden, Germany
[t] Clinical Psychology and Psychotherapy of Children and Adolescents, University of Kaiserslautern-Landau, Germany
[u] Clinical Psychology and Psychotherapy, Siegen University, Germany
[v] Clinicial Psychology and Psychotherapy, Universität Hamburg, Germany
[w] Department of Psychology, Humboldt-Universität zu Berlin, Germany
[x] German Center for Mental Health (DZPG), partner site Berlin-Potsdam, Germany
[y] Clinical Psychology and Psychotherapy, Department of Psychology, University of Trier, Germany
[z] Mental Health Research and Treatment Center, Faculty of Psychology, Ruhr University Bochum, Germany
[aa] WIPS Saarbrücken, Saarland University, Germany
[ab] Clinical Psychology and Psychotherapy, Department of Psychology and Psychotherapy, University Witten/Herdecke, Germany
[ac] Clinical Psychology and Psychotherapy, University of Kiel, Germany
[ad] Clinical Psychology and Psychotherapy, Freie Universität Berlin, Germany
[ae] Clinical Psychology and Psychotherapy of Adulthood, School of Human Sciences, Osnabrueck University, Germany
[af] Clinical Psychology and Psychotherapy, Department of Psychology, University of Leipzig, Germany

* Corresponding author at: Friedrich-Schiller-University Jena, Institute of Psychology, Department of Counseling and Clinical Intervention, Humboldtstraße 11, 07743 Jena, Germany.
E-mail address: nicolas.wrede@uni-jena.de (N. Wrede).

https://doi.org/10.1016/j.jad.2025.01.145
Received 7 August 2024; Received in revised form 24 January 2025; Accepted 27 January 2025
Available online 29 January 2025

[ag] Clinical Psychology and Psychotherapy, Department of Psychology, University of Osnabrück, Germany
[ah] Clinical Neuropsychology and Psychotherapy, Faculty of Psychology and Sports Science, Bielefeld University, Germany

ARTICLE INFO

Keywords:
Depression
Anxiety
Older adults
Oldest-old
Effectiveness
Cognitive-behavioral therapy

ABSTRACT

Background: Although meta-analyses suggest comparable efficacy of cognitive-behavioral therapy (CBT) in older adults compared to working-age adults, little is known about its effectiveness in naturalistic settings across different age groups. Hence, this study compared symptom change, attrition rates, and treatment duration in outpatient CBT between working-age adults (18–64 years), young-old adults (65–74 years), and old-old adults (≥ 75 years).
Methods: We analyzed a large naturalistic dataset comprising 9081 patients between 18 and 96 years receiving outpatient CBT in Germany. Using propensity score matching, we examined differences in treatment response, remission, attrition, and duration between comparable groups of working-age, young-old, and old-old adults.
Results: Response and remission rates did not differ between the three age groups in terms of the Brief Symptom Inventory (BSI-53) and patient- and clinician-rated subjective improvement (CGI-I). Young-old and old-old adults showed lower rates of response and remission on the Beck Depression Inventory (BDI-II). These differences were limited to items assessing somatization, which may be related to normal aging. Treatment duration was shorter in young-old and old-old adults compared to working-age adults. Attrition rates did not differ.
Limitations: The samples of older adults were relatively small and probably selective. Especially, home-bound, vulnerable older adults may be underrepresented. Further, the observational study design limits interpretability of findings.
Conclusions: Young-old and old-old adults seem to benefit from outpatient CBT to a similar extent as working-age adults. Potential bias in outcome measures due to age-related somatic complaints should be acknowledged in practice and future research.

1. Background

The global demographic trend of population aging continues to progress. Across member countries of the Organization for Economic Co-operation and Development (OECD), the proportion of individuals aged 65 and over is projected to increase from 18.0 % in 2021 to 26.7 % by 2050. Simultaneously, the share of those aged 80 and above is expected to approximately double, rising from 4.8 % to 9.8 % during the same period (OECD, 2023). Given the high prevalence of common mental disorders among older adults (Andreas et al., 2017; Hu et al., 2022; Volkert et al., 2013), the demand for effective treatments targeting this population is steadily increasing.

Meta-analyses of randomized controlled trials (RCTs) provide robust evidence supporting the efficacy of cognitive-behavioral therapy (CBT) across a broad spectrum of mental disorders (Cuijpers et al., 2016; Hofmann et al., 2012). However, the majority of these RCTs have been conducted with working-age adult samples, leaving older adults vastly underrepresented in the literature (Fordham et al., 2021). As a result, the strong evidence for the general efficacy of CBT cannot be confidently generalized to older adult populations. Additionally, there is a widespread stereotypical belief that CBT is less effective in older adults compared to younger cohorts (Frost et al., 2019; Kessler and Blachetta, 2020). Specifically, some practitioners stereotype older adults as less likely to benefit from psychotherapy due to assumptions of diminished learning abilities, cognitive rigidity, and limited potential for change due to a shorter remaining life expectancy (Bodner et al., 2018). In line with this belief, older adults are currently found to be underrepresented in psychotherapeutic treatments, while psychotropic medication is prescribed more frequently in this population (Gellert et al., 2021; Sanglier et al., 2011).

To address the concerns about the generalizability of CBT efficacy due to the underrepresentation of older adults in existing research, an increasing number of meta-analyses have been conducted to specifically synthesize the available evidence on the efficacy of CBT in older populations. These meta-analyses indicate that CBT is efficacious in treating common mental disorders among older adults, such as depression (Cuijpers et al., 2006; Davison et al., 2024; Gould et al., 2012a; Holvast et al., 2017) and anxiety disorders (Gonçalves and Byrne, 2012; Gould et al., 2012b; Hall et al., 2016). Additionally, some meta-analyses have directly compared the efficacy of CBT between older and working-age adults. For instance, Kishita and Laidlaw (2017) found no significant differences in the efficacy of CBT for generalized anxiety disorder between older and working-age adults, although treatment effects were descriptively smaller in the older cohort. Similarly, regarding CBT for depression, several meta-analyses reported no significant differences in efficacy between older adults and working-age adults (Cuijpers et al., 2020; Cuijpers et al., 2009; Werson et al., 2022).

While RCTs are critical for establishing causal treatment effects, several factors limit their generalizability to routine clinical care settings. RCTs typically test standardized CBT protocols in homogeneous patient groups (e.g., excluding patients with comorbidities), whereas CBT is delivered more flexibly in routine practice to diverse populations (e.g., often including comorbid patients) (Nathan et al., 2000; Waller and Turner, 2016). Additionally, challenges of routine clinical practice, such as treatment selection, age-related stereotypes, and barriers to access, are controlled for in RCTs but may still influence the uptake, delivery, and effectiveness of CBT in older adults. Therefore, complementing RCTs with naturalistic observational studies using routine care data is essential to evaluate the generalizability of findings (Franklin et al., 2019).

Recent studies have analyzed large-scale routine care datasets on psychological treatments in the UK, revealing that older adults have significantly less access to psychological therapies, with CBT being particularly underutilized in this population (Chaplin et al., 2015; Saunders et al., 2021). Despite receiving fewer treatment sessions, older adults achieved symptom reductions comparable to, or even exceeding, those observed in working-age adults across a broad range of psychological treatments. Further, rates of treatment attrition were lower in older adults compared to working-age adults (Chaplin et al., 2015; Pettit et al., 2017; Saunders et al., 2021). Moreover, studies analyzing routine care datasets specifically on CBT regarding differences between older and working-age adults consistently found that older adults benefit from CBT to a similar extent as their younger counterparts (Karlin et al., 2013; Karlin et al., 2015; Pomerleau et al., 2023). Further, consistent with the findings on various psychological treatments mentioned above, these studies indicated that, in CBT, treatment attrition may be lower in older

[1] Co-first authorship: Nicolas Wrede and Mareike C. Hillebrand contributed equally to this work.

adults (Karlin et al., 2015) and that older adults tend to receive fewer sessions (Pomerleau et al., 2023).

Although previous naturalistic studies provide valuable insights into differences in symptom change, attrition, and treatment duration in CBT between older and working-age adults, they have generally treated older adults as a homogeneous group, typically including individuals aged 65 years and older (Chaplin et al., 2015; Karlin et al., 2013; Karlin et al., 2015; Saunders et al., 2021). However, due to the observed delay in age-related physical and cognitive decline in industrialized countries over recent decades, it has been suggested that individuals aged 65 to 74 years be classified as young-old, as many in this age group maintain high levels of physical and cognitive functioning (Brailean et al., 2016; Ouchi et al., 2017). Therefore, age-related factors that may necessitate adaptions to CBT, such as somatic comorbidities, cognitive decline, or social isolation (Laidlaw and Kishita, 2015; Laidlaw and Pachana, 2009), may be more prevalent in old-old adults (≥ 75 years) compared to young-old adults (65–74 years). Hence, it has been recommended to distinguish between these groups when evaluating the effectiveness of CBT in older age (Kishita and Laidlaw, 2017; Laidlaw, 2021).

To date, the only naturalistic study examining symptom change during psychological interventions that differentiated between subgroups of older adults included only individuals up to the age of 74 (Pettit et al., 2017). Therefore, it remains unclear whether the findings of comparable symptom change during CBT between older and working-age adults extend to those aged 75 and older. While Chaplin et al. (2015) reported particularly limited access to psychological treatments for individuals aged 75 and above, no study has yet compared the symptom change, attrition, and treatment duration in CBT under naturalistic conditions in this age group with those of working-age adults.

1.1. Research objective

Given the demographic shift and high prevalence of mental disorders in older adults, there is a critical need for effective treatments. While RCTs indicate comparable efficacy of CBT in older adults compared to working-age adults, naturalistic studies can provide insights into its comparative effectiveness in routine clinical settings between different age groups. This study contributes to the limited body of research comparing symptom change, attrition, and treatment duration in outpatient CBT between older and working-age adults under naturalistic conditions. Importantly, the study extends previous research by examining young-old adults (65–74 years) and old-old adults (≥ 75 years) separately, thereby recognizing the potential differences in the risk of physical and cognitive decline between these groups.

2. Methods

2.1. Participants

The study's data were collected from outpatient university clinics of the KODAP research network (acronym for "Coordination of Data Acquisition at Research Clinics for Psychotherapy"). The KODAP Network was founded in 2013. It aims to merge and jointly evaluate routine assessments from multiple university outpatient clinics to gain enhanced insights into research and clinical care questions (Deisenhofer et al., 2024; In-Albon et al., 2019; Margraf et al., 2021; Teismann et al., 2024a; Teismann et al., 2024b; Velten et al., 2017, 2025). The study procedures of the KODAP network have been registered in the German Clinical Trials Register (DRKS00015883) and approved by the Ethics Committee of the Faculty of Psychology at Ruhr University Bochum. All patients provided informed consent prior to participation.

We analyzed a retrospective dataset comprising 15,296 adult patients who initiated CBT at 30 KODAP clinics between 2014 and 2022. The clinics classified treatments as either regularly completed (n = 7886), prematurely terminated (n = 1963), ongoing (n = 5360), currently suspended (n = 41), or not reimbursed (n = 46). Regularly completed treatments were defined as those in which the patient and therapist mutually agreed to terminate therapy, or when the reimbursable session limit set by health insurance was reached. In contrast, prematurely terminated treatments were defined as those where therapy ended without mutual agreement, even though additional sessions were available under the session quotas. For this study, we included all regularly completed and prematurely terminated treatments.

To initiate treatment, all patients were required to have at least one ICD-10 diagnosis for which CBT is considered an appropriate treatment. Patients with missing data regarding their primary diagnosis (n = 768) were excluded resulting in a final dataset of 9081 patients aged between 18 and 96 years. Sample characteristics are presented in Table 1. More than one-third of the patients (40.0 %, n = 3630) had a depressive disorder (including major depressive disorder and dysthymia) as their primary diagnosis, while approximately one-quarter (28.3 %, n = 2569) had an anxiety disorder (including agoraphobia, panic disorder, social phobia, specific phobias, generalized anxiety disorder, obsessive-compulsive disorder, and post-traumatic stress disorder) as their primary diagnosis. More than one-third of patients (40.7 %, n = 3696) was assigned more than one diagnosis. Most diagnoses were based on structured clinical interviews (92.6 %, n = 8406) such as the Structured Clinical Interview for DSM-IV (SCID; Wittchen et al., 1997) or DSM-5 (Beesdo-Baum et al., 2019) and the Diagnostic Interview for Mental Disorders (DIPS; Margraf et al., 2017). Further, a small percentage of diagnoses were based on screening questionnaires (5.6 %, n = 513) or on the therapist's clinical judgment (0.4 %, n = 34). 126 patients (1.4 %) had missing data on the diagnostic method used.

2.2. Therapists and treatment

Patients received CBT from a total of 1238 therapists, with an average caseload of 7.34 patients per therapist (SD = 8.65). Of these therapists, 396 (32.0 %) were fully trained in CBT, while 746 (60.3 %) were in the advanced stages of a 3-year (full-time) or 5-year (part-time) postgraduate CBT training program. Additionally, 79 therapists (6.4 %) treated patients both before and after completing their training during the observation period, and 17 therapists (1.4 %) had missing data regarding their CBT training status. The therapists had a mean age of 31.09 years (SD = 5.13 years) and 1037 (83.8 %) were female. During treatment, 781 patients (8.6 %) experienced at least one therapist change.

All therapies were conducted in university-based outpatient CBT clinics and were reimbursed as CBT treatments by health insurance providers in Germany. Therapies were mostly individual therapies (93.7 %, n = 8512), but also included group treatments (0.6 %, n = 55), and combined treatments (5.7 %, n = 514). For long-term CBT treatments (> 24 sessions), reimbursement requires the submission of a CBT-based individualized case conceptualization and treatment plan aligned with the patient's primary diagnosis. Further, the participating clinics implement rigorous quality monitoring in terms of diagnostics, case conceptualization, treatment delivery, and documentation. In line with German training regulations, CBT therapists in training received supervision at least after every fourth session. Therefore, it can be reasonably assumed that the treatments adhered to guideline-based CBT for the patient's primary diagnosis, and that therapists were subjected to rigorous quality monitoring and supervision.

2.3. Measures

The Beck Depression Inventory-II (BDI-II; Hautzinger et al., 2009), a widely used instrument in both research and clinical practice, was employed to measure depressive symptoms before treatment initiation and after treatment completion. This instrument comprises 21 items, each rated on a 4-point scale from 0 to 3, yielding a total score ranging from 0 to 63, with higher scores indicating greater severity of depression. The BDI-II demonstrated good internal consistency in previous

Table 1
Sample characteristics in the subsamples of working-age adults (18–64 years), young-old adults (65–74 years), and old-old adults (≥ 75 years).

	Working-age adults (*N* = 8793)	Young-old adults (*N* = 222)	Old-old adults (*N* = 66)	F / χ^2
Age, *M (SD)*	35.72 (12.66)	68.29 (2.62)	78.92 (3.82)	2120.1***
Female, *N* (%)	5604 (63.8)	153 (68.9)	53 (80.3)	10.02**
A-level degree, *N* (%)	4324 (58.8)	64 (37.6)	17 (35.4)	40.77***
In a relationship, *N* (%)	3723 (50.7)	109 (61.6)	23 (46.9)	8.57*
Treatment setting, *N* (%)				
Individual	8262 (94.0)	199 (89.6)	51 (77.3)	
Group	44 (0.5)	6 (2.7)	5 (7.6)	85.32***
Both	487 (5.5)	17 (7.7)	10 (15.2)	
Clinic type, *N* (%)				
Research	1635 (20.9)	67 (33.2)	34 (52.3)	54.42***
Training	6179 (79.1)	135 (66.8)	31 (47.7)	
Session number, *M (SD)*	36.01 (23.09)	29.35 (20.02)	20.17 (15.06)	47.32***
Primary diagnosis, *N* (%)				
Substance use disorder	101 (1.1)	4 (1.8)	0 (0.0)	
Schizophrenia	130 (1.5)	0 (0.0)	0 (0.0)	
MDD (single)	1109 (12.6)	23 (10.4)	7 (10.6)	
MDD (recurrent)	1967 (22.4)	43 (19.4)	10 (15.2)	
MDD (remitted)	141 (1.6)	3 (1.4)	0 (0.0)	
Dysthymia	317 (3.6)	10 (4.5)	0 (0.0)	
Agoraphobia/Panic Disorder	594 (6.8)	25 (11.3)	7 (10.6)	
Social Phobia	579 (6.6)	2 (0.9)	1 (1.5)	193.61***
Specific Phobia	236 (2.7)	13 (5.9)	2 (3.0)	
GAD	209 (2.4)	12 (5.4)	5 (7.6)	
Other phobic or anxiety disorders	173 (2.0)	8 (3.6)	6 (9.1)	
OCD	299 (3.4)	6 (2.7)	2 (3.0)	
Adjustment disorder	858 (9.8)	27 (12.2)	10 (15.2)	
PTSD	385 (4.4)	4 (1.8)	1 (1.5)	
Somatoform disorder	477 (5.4)	22 (9.9)	11 (16.7)	
Eating disorder	300 (3.4)	2 (0.9)	0 (0.0)	
Psychological/ behavioral factors associated with disorders/diseases classified elsewhere	145 (1.6)	5 (2.3)	2 (3.0)	
Borderline Personality Disorder	231 (2.6)	0 (0.0)	0 (0.0)	
Other	542 (6.2)	13 (5.9)	2 (3.0)	
Number of diagnoses, *M (SD)*	1.58 (0.82)	1.37 (0.64)	1.26 (0.54)	24.24***
BDI-II before treatment, *M (SD)*	21.62 (11.50)	18.21 (10.70)	15.41 (8.92)	27.95***
BSI-53 before treatment, *M (SD)*	1.07 (0.64)	0.84 (0.57)	0.76 (0.49)	32.93***
Previously received psychotherapy, *N* (%)	3738 (42.5)	109 (49.1)	38 (57.6)	9.78**

Note. Sample characteristics were derived from non-imputed data. MDD = Major Depressive Disorder, GAD = Generalized Anxiety Disorder, OCD = Obsessive Compulsive Disorder, PTSD = Post Traumatic Stress Disorder. F = F-statistics from one-way ANOVA to test differences between age groups on continuous outcomes, χ^2 = test statistics from χ^2-test to test differences between age groups on categorial outcomes, *: $p < .05$; **: $p < .01$, ***: $p < .001$.

studies (α = 0.89; Erford et al., 2016) and in the current sample (α = 0.92).

The Brief Symptom Inventory-53 (BSI-53; Franke, 2000) was utilized to assess psychopathological symptoms before treatment initiation and after treatment completion. This instrument contains 53 items, each rated on a 5-point scale from 0 ("not at all") to 4 ("extremely"), with higher scores reflecting greater symptom severity. The BSI-53 captures a wide range of psychopathological symptoms on nine subscales (somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, psychoticism). A mean score across all items was used as a global indicator of symptom distress. Internal consistency for this measure was high in previous studies (α = 0.96; Prinz et al., 2013) and in the current sample (α = 0.96).

The Global Improvement subscale of the Clinical Global Impression Scale (CGI-I; Busner and Targum, 2007) was used to assess subjective symptom improvement. After treatment completion, both patients and therapists rated the patient's improvement during therapy on a 7-point scale, ranging from 1 ("very much improved since the initiation of treatment") to 7 ("very much worse since the initiation of treatment"). The CGI-I provides a valuable complement to symptom questionnaire change scores by offering a holistic evaluation of improvement from both perspectives. This broader assessment is particularly suitable given the heterogeneity of the sample in this study.

All assessments after treatment completion were intended to be conducted regardless of whether the treatment was regularly completed or prematurely terminated.

2.4. Data analysis

We performed data analyses using R version 4.4.1 (R Core Team, 2024). We provide statistical input and output code for all analyses of this study in the Supplementary Material.

2.4.1. Multiple imputation

Missing data were handled using multiple imputation by chained equations, implemented with the R package *mice* version 3.16.0 (van Buuren and Groothuis-Oudshoorn, 2011). Imputation was conducted at the item level before computing total scores of measurement scales (Gottschall et al., 2012). To ensure that all informative variables of the dataset could be used to impute missing values, the predictor matrix was created based on correlations and usable cases using the "quickpred" function within *mice*. A minimum correlation of $r = 0.2$ and at least 50 % usable cases were used as thresholds. We generated 50 imputed datasets. Data analyses were performed separately on each of the 50 imputed datasets and then pooled (Rubin, 1987).

2.4.2. Propensity score matching

In all analyses, age was used as a categorical variable encompassing three age groups: working-age adults (18–64 years), young-old adults (65–74 years), and old-old adults (≥ 75 years). An important source of bias in naturalistic studies is confounding, i.e., that risk factors for treatment outcome may be imbalanced between compared age groups in our study. We addressed this source of bias by propensity score matching to compare subsamples of working-age adults that were balanced with young-old or old-old adults regarding relevant covariates. Throughout data analysis, we applied a stepped approach using three models which stepwise included relevant covariates.

Model 1 was based on the full non-matched dataset, and we conducted analyses while controlling for initial symptom scores on BDI-II and BSI-53. Model 2 was performed using matched data based on patient-related and therapist-related covariates, ensuring that older adults were compared with younger patients who had similar characteristics and were treated by comparable therapists. Model 3 was conducted using matched data based on patient-related, therapist-related, and treatment-related covariates, ensuring that older adults were compared with working-age adults who not only had similar characteristics and therapists but were also treated in a comparable manner. Notably, treatment duration was not included as a covariate in model 3 for analyses regarding treatment attrition and duration. A detailed description of all variables used for matching is provided in the Supplementary Material.

In all analyses conducted within model 1, age group was included as a categorical predictor of the respective outcome, and pairwise contrasts were performed to examine group differences. For all analyses within

model 2 and model 3, two separate matching procedures were conducted, one to match a subsample of working-age adults that has optimal covariate balance with young-old adults (65–74 years), and another one to match a subsample of working-age adults that has optimal covariate balance with old-old adults (≥ 75 years). These matched group pairs were then analyzed separately to compare the respective outcome between young-old adults and working-age adults and between old-old adults and working-age adults.

Given the underrepresentation of young-old and old-old adults in the sample compared to the much larger group of working-age adults, traditional 1:1 matching poses the risk of reduced precision by excluding a substantial proportion of the working-age group (Austin and Mamdani, 2006). Specifically, 1:1 matching would pair each older adult with only the single most similar working-age counterpart, thereby disregarding numerous other potentially comparable individuals. For example, only 66 (0.8 %) working-age adults would be selected as the comparison group for old-old adults (≥ 75 years), leaving 8727 (99.2 %) working-age adults excluded from analysis. To address this limitation, full optimal matching was employed, which permits one observation from one group to be matched with one or more similar observations from the other group, as long as their similarity falls within a predefined threshold (Austin and Stuart, 2017). This method maximizes the use of available data and thereby enhances the precision of group comparisons. We used a caliper of 0.2 standard deviations of the logit of the propensity score as a threshold for matching (Austin, 2011). Covariate balance after matching was assessed by standardized mean differences and variance ratios (Austin, 2009).

As our analyses were based on multiply imputed data, we combined full optimal matching with multiple imputation using a within-matching strategy, wherein matching was conducted separately within each imputed dataset (Leyrat et al., 2017). The effects estimated from each matched imputed dataset were then pooled to provide overall estimates. The described matching procedure was performed using the R package *MatchThem* version 1.2.1 (Pishgar et al., 2021).

2.4.3. Symptom change

To evaluate age group differences in symptom change during treatment, we analyzed several indicators of treatment response and remission:

Response on the BDI-II and BSI-53 was defined by reliable symptom improvement following the approach outlined by Jacobson and Truax (1991). Individual pre-post difference scores were divided by the standard error of the difference, which was calculated based on the standard deviation and internal consistency of the respective scale at baseline. Improvements of >1.96 standard errors were considered reliable. Based on this method, response on the BDI-II was defined by an absolute improvement of at least 9 points, whereas response on the BSI-53 was defined by an improvement of 0.35 points. As changes within the non-clinical range were not of interest for our study, only patients that met the clinical cutoff of at least 13 on the BDI-II (von Glischinski et al., 2019), or at least 0.60 on the BSI-53 (Schmitz et al., 2000) before treatment were included in the analyses.

Response on the CGI-I was defined by a rating of either 1 ("very much improved") or 2 ("much improved"). Response based on CGI-I rated by therapists and patients were analyzed separately.

Remission on the BDI-II was defined by achieving reliable symptom improvement to a score below the clinical cutoff of 13, whereas remission on the BSI-53 was defined by achieving reliable symptom improvement to a score below 0.60.

For the analysis of age group differences in treatment response and remission, logistic regression models were employed using the full sample. Sensitivity analyses were conducted by repeating the analyses regarding symptom change using data exclusively from treatment completers, i.e., patients that did not prematurely terminate treatment (n = 7416). Additionally, we performed disorder-specific sensitivity analyses only including the 3630 patients with a depressive disorder as their primary diagnosis and only including the 2569 patients with an anxiety disorder as their primary diagnosis.

2.4.4. Treatment attrition and duration

To examine differences in treatment attrition between age groups, premature treatment discontinuation was predicted by age group in logistic regression models. Differences in treatment duration were examined by predicting session number by age group in analyses of covariance (ANCOVAs) using heteroscedasticity-consistent standard errors.

3. Results

3.1. Covariate balance after propensity score matching

Covariate balance was excellent in all matching procedures regarding young-old adults (all SMDs between −0.01 and 0.02; all variance ratios between 0.86 and 1.07) and at least acceptable in all matching procedures regarding old-old adults (all SMDs between −0.07 and 0.04; all variance ratios between 0.60 and 1.19). Detailed information on covariate balance after matching is provided in the Supplementary Material.

3.2. Symptom change

Groupwise rates of treatment response and remission as well as results of logistic regression analysis are given in Table 2. After controlling for relevant covariates through propensity score matching, no significant differences between young-old and working-age adults or between old-old and working-age adults were found in response and remission rates on the BSI-53 or in response rates on the CGI-I rated by patients and therapists. Additionally, direct comparisons between young-old and old-old adults within model 1 showed no significant differences regarding the BSI-53 (response: $b = -0.41$, $SE = 0.43$, $p = .34$, $OR = 0.66$; remission: $b = -0.54$, $SE = 0.44$, $p = .22$, $OR = 0.58$) or regarding response on the CGI-I rated by patients ($b = -0.23$, $SE = 0.35$, $p = .51$, $OR = 0.79$) and therapists ($b = -0.28$, $SE = 0.32$, $p = .37$, $OR = 0.75$).

However, both young-old and old-old adults exhibited significantly lower rates of response and remission on the BDI-II compared to working-age adults, even after propensity score matching. Direct comparisons within model 1 revealed significantly lower BDI-II response rates in old-old compared to young-old adults ($b = -0.99$, $SE = 0.51$, $p = .049$, $OR = 0.37$), while there were no significant differences in remission rates ($b = -0.89$, $SE = 0.57$, $p = .12$, $OR = 0.41$).

Results from sensitivity analyses only including treatment completers mostly replicated the results of the main analyses. Further, sensitivity analyses only including patients with depressive or, respectively, anxiety disorders did not reveal any significant age group differences that were not observed in the full sample. The results of the sensitivity analyses are provided in the Supplementary Material.

3.2.1. Exploratory post-hoc analysis

Given the considerable differences in the results between the BDI-II and the other outcome measures, we conducted exploratory post-hoc analyses to investigate potential reasons for these discrepancies. Research suggests that somatic symptoms may significantly influence BDI-II scores (Thombs et al., 2010; Wedding et al., 2007) and this may distort the BDI-II scores in older adults (Georgi et al., 2019; Trentini et al., 2005). A common separation of the BDI-II into the subscales "somatic-affective" (12 items; e.g., loss of energy, change in sleeping pattern, changes in appetite, loss of interest in sex) and "cognitive" (9 items; e.g., sadness, pessimism, self-dislike, suicidal ideation) (Subica et al., 2014), suggests that 57.1 % of all BDI-II items are related to somatic-affective symptoms. In contrast, the subscale "somatization" of the BSI-53 only comprises seven items which corresponds to 13.2 % of all items. Hence, the distorting influence of somatic complaints on

Table 2
Results from logistic regression analysis comparing treatment response and remission in young-old adults (65–74 years) and old-old adults (≥ 75 years) with working-age adults (18–64 years).

		Young-old adults vs. working-age adults				Old-old adults vs. working-age adults			
		Response/remission rates, %		Logistic regression		Response/remission rates, %		Logistic regression	
Outcome	Model	Working-age	Young-old	*b* (*SE*)	Odds Ratio	Working-age	Old-old	*b* (*SE*)	Odds Ratio
Response BDI-II	Model 1	62.3	46.0	−0.60 (0.20)**	0.55	62.3	23.4	−1.59 (0.47)***	0.20
	Model 2	58.4	46.5	−0.52 (0.22)*	0.60	55.1	24.5	−1.44 (0.55)**	0.24
	Model 3	56.6	47.0	−0.42 (0.22)$^{+}$	0.65	51.7	22.9	−1.33 (0.58)*	0.26
Response BSI-53	Model 1	61.7	52.3	−0.29 (0.19)	0.75	61.7	41.3	−0.70 (0.38)$^{+}$	0.50
	Model 2	58.6	52.5	−0.28 (0.20)	0.76	56.3	43.0	−0.57 (0.43)	0.57
	Model 3	57.8	52.5	−0.24 (0.20)	0.78	54.3	42.9	−0.52 (0.48)	0.60
Response CGI-I (Pat)	Model 1	71.5	72.3	−0.05 (0.18)	0.96	71.5	68.7	−0.28 (0.31)	0.76
	Model 2	73.1	72.0	−0.05 (0.19)	0.95	73.6	68.2	−0.26 (0.33)	0.77
	Model 3	72.3	72.1	−0.01 (0.19)	0.99	69.6	69.8	0.00 (0.39)	1.00
Response CGI-I (Th)	Model 1	58.9	56.4	−0.18 (0.15)	0.83	58.9	50.8	−0.47 (0.29)	0.63
	Model 2	60.2	56.0	−0.17 (0.16)	0.84	59.6	50.6	−0.38 (0.32)	0.69
	Model 3	58.4	56.1	−0.09 (0.15)	0.91	53.5	52.5	−0.05 (0.34)	0.95
Remission BDI-II	Model 1	43.2	30.3	−0.68 (0.22)**	0.51	43.2	17.2	−1.57 (0.52)**	0.21
	Model 2	43.6	30.7	−0.56 (0.22)*	0.57	43.8	17.6	−1.34 (0.61)*	0.26
	Model 3	41.3	31.0	−0.46 (0.22)*	0.63	40.1	18.5	−1.12 (0.62)$^{+}$	0.33
Remission BSI-53	Model 1	35.3	33.2	−0.23 (0.20)	0.80	35.3	26.3	−0.77 (0.40)$^{+}$	0.46
	Model 2	37.8	33.3	−0.20 (0.21)	0.81	41.2	29.0	−0.56 (0.43)	0.57
	Model 3	36.6	33.3	−0.15 (0.21)	0.86	37.9	28.6	−0.46 (0.46)	0.63

Note. Model 1 only included initial outcome scores as covariates, model 2 was based on propensity score matching using patient- and therapist-related covariates, model 3 was based on propensity score matching using patient-, therapist-, and treatment-related covariates. For response and remission analyses regarding BDI-II and BSI-53, only patients with clinical intake symptoms on the respective scale were included (6234 working-age adults, 129 young-old adults, and 33 old-old adults for the BDI-II; 6619 working-age adults, 140 young-old adults, and 39 old-old adults for the BSI-53). *b* = logistic regression coefficient, *SE* = standard error, $^{+}$: $p < .10$, *: $p < .05$; **: $p < .01$, ***: $p < .001$.

change scores may be stronger for the BDI-II, than for other measures.

To explore the influence of somatic symptoms on our results, we compared pre-post changes on subscales of the BDI-II and BSI-53 between age groups in ANCOVAs within model 3. We analyzed change on the BDI-II subscales "somatic-affective" and "cognitive" and on the BSI-53 separated into the "somatization" subscale and the rest of all items. To ensure comparability of results, pre-post changes on the subscales were standardized in these ANCOVAs. We found that young-old adults had significantly smaller pre-post change on the BDI-II compared to working-age adults on the somatic-affective subscale ($b = 0.13$, $SE = 0.06$, $p = .03$), but not on the cognitive subscale ($b = 0.02$, $SE = 0.05$, $p = .67$). Further, differences between old-old adults and working-age adults were significant on the somatic-affective subscale ($b = 0.35$, $SE = 0.12$, $p = .003$), but not on the cognitive subscale ($b = 0.17$, $SE = 0.11$, $p = .10$). Similarly, young-old adults showed significantly smaller pre-post change on the somatization subscale of the BSI-53 ($b = 0.12$, $SE = 0.06$, $p = .04$), but not on the remaining scale ($b = -0.02$, $SE = 0.05$, $p = .77$). In the same way, old-old adults showed significantly lower pre-post change on the somatization subscale ($b = 0.33$, $SE = 0.13$, $p = .01$), but not on the remaining scale ($b = 0.02$, $SE = 0.09$, $p = .81$).

3.3. Treatment attrition and duration

3.3.1. Treatment attrition

Unadjusted numbers of premature treatment termination were 1615 (18.4 %) working-age adults, 41 (18.5 %) young-old adults, and 9 (13.6 %) old-old adults. There were no significant differences in treatment attrition between working-age and young-old adults in any of the models (model 1: $b = 0.07$, $SE = 0.18$, $p = .67$, $OR = 1.08$; model 2: $b = 0.02$, $SE = 0.19$, $p = .89$, $OR = 1.03$; model 3: $b = 0.02$, $SE = 0.18$, $p = .92$, $OR = 1.02$). Similarly, there were no significant differences in treatment attrition between old-old adults and working-age adults (model 1: $b = -0.24$, $SE = 0.36$, $p = .50$, $OR = 0.78$; model 2: $b = -0.39$, $SE = 0.41$, $p = .34$, $OR = 0.68$; model 3: $b = -0.58$, $SE = 0.46$, $p = .21$, $OR = 0.56$). Direct comparisons between young-old and old-old adults within model 1 revealed no significant differences ($b = -0.32$, $SE = 0.40$, $p = .43$, $OR = 0.73$).

3.3.2. Treatment duration

Young-old adults had significantly lower session counts compared to working-age adults in model 1 ($b = -5.17$, $SE = 1.43$, $p < .001$). However, these differences were only marginally significant when further covariates were accounted for (model 2: $b = -3.01$, $SE = 1.61$, $p = .06$; model 3: $b = -2.99$, $SE = 1.63$, $p = .07$). In the old-old adults, treatment duration was significantly shorter compared to working-age adults in all models (model 1: $b = -13.93$, $SE = 1.87$, $p < .001$; model 2: $b = -10.25$, $SE = 2.36$, $p < .001$; model 3: $b = -9.98$, $SE = 2.59$, $p < .001$). Direct comparisons between young-old and old-old adults within model 1 revealed a significantly shorter treatment duration in old-old adults compared to young-old adults ($b = -8.76$, $SE = 2.34$, $p < .001$).

4. Discussion

With the aim to evaluate how older adults benefit from outpatient CBT compared to working-age adults, this study analyzed a large clinical routine dataset from 9081 CBT outpatients aged 18 to 96 years. Our study is one of the few to compare symptom change, attrition, and treatment duration in CBT under naturalistic conditions between older and working-age adults and is notably the first to distinguish between young-old (65–74 years) and old-old adults (≥ 75 years). Therefore, the results provide nuanced insights into age-related differences in the delivery and outcome of CBT in naturalistic settings.

Overall, our results suggest that the symptom reduction during outpatient CBT under naturalistic conditions is comparable between older adults and working-age adults. This aligns with meta-analyses demonstrating similar efficacy of CBT across these age groups (Kishita and Laidlaw, 2017; Werson et al., 2022) and with previous naturalistic studies (Chaplin et al., 2015; Karlin et al., 2013; Karlin et al., 2015; Pettit

et al., 2017; Pomerleau et al., 2023; Saunders et al., 2021). Importantly, our findings contribute to the literature by showing that this also holds true for adults aged 75 years and above. Therefore, the common pessimism surrounding the use of CBT in older adults is unwarranted. Our findings, together with studies showing lower perceived competence among therapists in training when treating older patients compared to younger ones (Kessler and Blachetta, 2020), support recent calls for the increased integration of geropsychology into CBT training curricula (Becker et al., 2020; Hinrichsen et al., 2018).

In addition to these encouraging results, our findings revealed a measurement-specific pattern that warrants attention. Specifically, we observed lower rates of response and remission in older adults on the BDI-II, but not on the BSI-53 or the CGI-I. The BDI-II may be less suitable for older adults due to distortion by age-related functional limitations (Georgi et al., 2019). Fittingly, age group differences were confined to the somatic-affective subscale of the BDI-II, suggesting that the somatic focus may be problematic for older populations. Further, the finding that changes on the somatic-affective subscale were particularly small in old-old adults supports the explanation that the observed differences in symptom improvement can be attributed to age-related somatic complaints. Future research and practice should critically consider potential bias in outcome measures in older adults due to age-related somatic symptoms.

In line with previous research, our findings suggest that older age is not a risk factor for premature treatment termination in CBT under naturalistic conditions (Chaplin et al., 2015; Karlin et al., 2013; Pomerleau et al., 2023; Saunders et al., 2021). However, consistent with earlier studies, we found that older adults received fewer CBT sessions compared to working-age adults (Pomerleau et al., 2023; Saunders et al., 2021). Importantly, our results extend this evidence by showing that this difference in session count is even more pronounced in old-old adults compared to young-old adults. Although the relationship between treatment duration and CBT effectiveness is unclear, with some evidence suggesting that the number of sessions is not associated to better outcome (Cuijpers et al., 2013), the causes behind this pattern warrant further exploration. Age-related stereotypes, for example, may influence both patients' and therapists' expectations of treatment success, leading to earlier treatment termination for older adults (Gellert et al., 2021; Kessler and Blachetta, 2020). Additionally, access barriers such as limited mobility may impede older individuals' ability to engage in outpatient psychotherapy for extended periods (Lavingia et al., 2020).

4.1. Limitations and future research

The present study has several limitations to be mentioned. First, the subsample of older adults that show up in outpatient psychotherapy clinics probably represents a selective subsample of older adults with mental disorders (Gellert et al., 2021). Especially, home-bound, vulnerable older adults may be not represented (Tegeler et al., 2020). Consequently, the results of this study cannot be readily extrapolated to the general population of older adults with mental disorders, but only refer to older adults that seek outpatient psychotherapy.

Second, certain diagnostic groups were underrepresented among the older patients in our sample, which limits the generalizability of our findings to these populations. This includes, for example, older individuals with substance use disorders, eating disorders, schizophrenia, or personality disorders. Hence, our conclusions primarily apply to older adults with depressive, anxiety, and somatoform disorders rather than across diagnostic categories.

Third, the relatively small group sizes of older adults compared to working-age adults limit statistical power of group comparisons. Hence, we may not have been able to detect small differences between these groups.

Fourth, the narrow range of available symptom measures presents a notable limitation. While these measures collectively capture a broad spectrum of common psychopathological symptoms, they are not equally suited to adequately reflect meaningful symptom changes across all included disorders.

Fifth, despite our efforts to account for relevant covariates using propensity score matching, unmeasured covariates could confound the relationship between age group and outcomes. Due to the use of naturalistic data, only a limited set of variables was available, which introduces the risk that important covariates were not captured. For example, physical health status, cognitive functioning, or additional pharmacological treatment could not be addressed in our analyses.

Sixth, the uncertainty surrounding treatment fidelity presents another limitation that is common in studies using routine care data. Although we reasonably assume that treatments adhered to guideline-based CBT, this assumption cannot be directly verified due to the lack of fidelity assessments. This may be particularly problematic as evidence suggests that patient age may influence the treatment decisions made by therapists (Kessler and Schneider, 2017). Hence, the comparability of treatments received by older and younger patients remains uncertain.

Finally, the observational design of our study presents an important limitation. Pre-post symptom changes do not necessarily equate to causal treatment effects, as they fail to account for how symptoms would have developed if the patient had not received treatment (Langkaas et al., 2018). Hence, potential systematic age group differences in this counterfactual untreated symptom development could not be accounted for. Future studies should include untreated control groups to provide more accurate insights into age group differences in the treatment effects of CBT under naturalistic conditions. Thereby, recent methods to obtain control conditions for naturalistic studies may be useful (Kaiser et al., 2023).

5. Conclusion

This study is one of the few that compared symptom change, attrition, and treatment duration in outpatient CBT between older and working-age adults and the first that thereby distinguished between young-old (65–74 years) and old-old adults ($\geq$ 75 years). The results indicate that symptom reduction and attrition during CBT is comparable between working-age adults, young-old adults, and old-old adults, while young-old and old-old adults tend to receive fewer sessions. The findings thus contribute to the growing body of evidence contradicting the persistent belief that older individuals are unable to benefit from CBT.

CRediT authorship contribution statement

Nicolas Wrede: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Mareike C. Hillebrand:** Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Anne Katrin Risch:** Writing – review & editing, Investigation, Conceptualization. **Georg W. Alpers:** Writing – review & editing, Investigation. **Stephan Bartholdy:** Writing – review & editing, Investigation. **Eva-Lotta Brakemeier:** Writing – review & editing, Investigation. **Anne-Kathrin Bräscher:** Writing – review & editing, Investigation. **Timo Brockmeyer:** Writing – review & editing, Investigation. **Hanna Christiansen:** Writing – review & editing, Investigation. **Monika Equit:** Writing – review & editing, Investigation. **Lydia Fehm:** Writing – review & editing, Investigation. **Thomas Forkmann:** Writing – review & editing, Investigation. **Julia Glombiewski:** Writing – review & editing, Investigation. **Jens Heider:** Writing – review & editing, Investigation. **Sylvia Helbig-Lang:** Writing – review & editing, Investigation. **Andrea Hermann:** Writing – review & editing, Investigation. **Christiane Hermann:** Writing – review & editing, Investigation. **Jürgen Hoyer:** Writing – review & editing, Project administration, Investigation. **Tina In-Albon:** Writing – review & editing, Investigation. **Tim Klucken:** Writing – review & editing, Investigation. **Tania M. Lincoln:** Writing – review & editing, Investigation. **Lea Ludwig:** Writing – review & editing, Investigation. **Ulrike Lueken:** Writing – review & editing, Investigation. **Wolfgang Lutz:** Writing – review & editing,

Investigation. **Jürgen Margraf:** Writing – review & editing, Project administration, Investigation. **Tanja Michael:** Writing – review & editing, Investigation. **Patrizia Odyniec:** Writing – review & editing, Investigation. **Anya Pedersen:** Writing – review & editing, Investigation. **Babette Renneberg:** Writing – review & editing, Investigation. **Julian Rubel:** Writing – review & editing, Investigation. **Almut Rudolph:** Writing – review & editing, Investigation. **Henning Schöttke:** Writing – review & editing, Investigation. **Brian Schwartz:** Writing – review & editing, Investigation. **Rudolf Stark:** Writing – review & editing, Investigation. **Tobias Teismann:** Writing – review & editing, Investigation. **Julia Velten:** Writing – review & editing, Project administration, Investigation. **Katja Werheid:** Writing – review & editing, Investigation. **Ulrike Willutzki:** Writing – review & editing, Investigation. **Michael Witthöft:** Writing – review & editing, Investigation. **Gabriele Wilz:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization.

Role of the funding source

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

Declaration of competing interest

None.

Acknowledgements

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2025.01.145.

References

Andreas, S., Schulz, H., Volkert, J., Dehoust, M., Sehner, S., Suling, A., Ausín, B., Canuto, A., Crawford, M., Da Ronch, C., Grassi, L., Hershkovitz, Y., Muñoz, M., Quirk, A., Rotenstein, O., Santos-Olmo, A.B., Shalev, A., Strehle, J., Weber, K., Härter, M., 2017. Prevalence of mental disorders in elderly people: the European MentDis_ICF65+ study. Br. J. Psychiatry 210 (2), 125–131. https://doi.org/10.1192/bjp.bp.115.180463.

Austin, P.C., 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat. Med. 28 (25), 3083–3107. https://doi.org/10.1002/sim.3697.

Austin, P.C., 2011. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm. Stat. 10 (2), 150–161. https://doi.org/10.1002/pst.433.

Austin, P.C., Mamdani, M.M., 2006. A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. Stat. Med. 25 (12), 2084–2106. https://doi.org/10.1002/sim.2328.

Austin, P.C., Stuart, E.A., 2017. Estimating the effect of treatment on binary outcomes using full matching on the propensity score. Stat. Methods Med. Res. 26 (6), 2505–2525. https://doi.org/10.1177/0962280215601134.

Becker, T., Martin, F., Wilz, G., Risch, A.K., Kessler, E.-M., Forstmeier, S., 2020. Psychotherapie im höheren Lebensalter in der Psychotherapieausbildung. Z. Klin. Psychol. Psychother. 49 (3), 172–181. https://doi.org/10.1026/1616-3443/a000593.

Beesdo-Baum, K., Zaudig, M., Wittchen, H.-U., 2019. Strukturiertes Klinisches Interview für DSM-5®-Störungen – Klinische Version - Deutsche Bearbeitung des Structured Clinical Interview for DSM-5® Disorders – Clinician Version. Hogrefe, Göttingen, Germany.

Bodner, E., Palgi, Y., Wyman, M.F., 2018. Ageism in mental health assessment and treatment of older adults. In: Ayalon, L., Tesch-Römer, C. (Eds.), Contemporary Perspectives on Ageism. Springer International Publishing, pp. 241–262. https://doi.org/10.1007/978-3-319-73820-8_15.

Brailean, A., Huisman, M., Prince, M., Prina, A.M., Deeg, D.J.H., Comijs, H., 2016. Cohort differences in cognitive aging in the longitudinal aging study Amsterdam. The Journals of Gerontology: Series B 73 (7), 1214–1223. https://doi.org/10.1093/geronb/gbw129.

Busner, J., Targum, S.D., 2007. The clinical global impressions scale: applying a research tool in clinical practice. Psychiatry (Edgmont) 4 (7), 28–37.

Chaplin, R., Farquharson, L., Clapp, M., Crawford, M., 2015. Comparison of access, outcomes and experiences of older adults and working age adults in psychological therapy. Int. J. Geriatr. Psychiatry 30 (2), 178–184. https://doi.org/10.1002/gps.4122.

Cuijpers, P., van Straten, A., Smit, F., 2006. Psychological treatment of late-life depression: a meta-analysis of randomized controlled trials. Int. J. Geriatr. Psychiatry 21 (12), 1139–1149. https://doi.org/10.1002/gps.1620.

Cuijpers, P., van Straten, A., Smit, F., Andersson, G., 2009. Is psychotherapy for depression equally effective in younger and older adults? A meta-regression analysis. Int. Psychogeriatr. 21 (1), 16–24. https://doi.org/10.1017/S1041610208008089.

Cuijpers, P., Huibers, M., Daniel Ebert, D., Koole, S.L., Andersson, G., 2013. How much psychotherapy is needed to treat depression? A metaregression analysis. J. Affect. Disord. 149 (1), 1–13. https://doi.org/10.1016/j.jad.2013.02.030.

Cuijpers, P., Cristea, I.A., Karyotaki, E., Reijnders, M., Huibers, M.J.H., 2016. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World Psychiatry 15 (3), 245–258. https://doi.org/10.1002/wps.20346.

Cuijpers, P., Karyotaki, E., Eckshtain, D., Ng, M.Y., Corteselli, K.A., Noma, H., Quero, S., Weisz, J.R., 2020. Psychotherapy for depression across different age groups: a systematic review and meta-analysis. JAMA Psychiatry 77 (7), 694–702. https://doi.org/10.1001/jamapsychiatry.2020.0164.

Davison, T.E., Bhar, S., Wells, Y., Owen, P.J., You, E., Doyle, C., Bowe, S.J., Flicker, L., 2024. Psychological therapies for depression in older adults residing in long-term care settings. Cochrane Database Syst. Rev. 3. https://doi.org/10.1002/14651858.CD013059.pub2.

Deisenhofer, A.-K., Hehlmann, M.I., Rubel, J.A., Lutz, W., Schwartz, B., Bräscher, A.-K., Christiansen, H., Fehm, L., Glombiewski, J.A., Heider, J., Helbig-Lang, S., Hermann, A., Hoyer, J., In-Albon, T., Lincoln, T., Margraf, J., Risch, A.K., Schöttke, H., Schulze, L., Odyniec, P., 2024. Love yourself as a therapist, doubt yourself as an institution? Therapist and institution effects on outcome, treatment length, and dropout. Psychother. Res. 1-14. https://doi.org/10.1080/10503307.2024.2352749.

Erford, B.T., Johnson, E., Bardoshi, G., 2016. Meta-analysis of the English version of the Beck Depression Inventory–Second Edition. Meas. Eval. Couns. Dev. 49 (1), 3–33. https://doi.org/10.1177/0748175615596783.

Fordham, B., Sugavanam, T., Edwards, K., Stallard, P., Howard, R., das Nair, R., Copsey, B., Lee, H., Howick, J., Hemming, K., Lamb, S.E., 2021. The evidence for cognitive behavioural therapy in any condition, population or context: a meta-review of systematic reviews and panoramic meta-analysis. Psychol. Med. 51 (1), 21–29. https://doi.org/10.1017/S0033291720005292.

Franke, G., 2000. BSI. Brief Symptom Inventory - Deutsche Version. Manual. Beltz, Göttingen.

Franklin, J.M., Glynn, R.J., Martin, D., Schneeweiss, S., 2019. Evaluating the use of nonrandomized real-world data analyses for regulatory decision making. Clin. Pharmacol. Ther. (St. Louis, MO, U. S.) 105 (4), 867–877. https://doi.org/10.1002/cpt.1351.

Frost, R., Beattie, A., Bhanu, C., Walters, K., Ben-Shlomo, Y., 2019. Management of depression and referral of older people to psychological therapies: a systematic review of qualitative studies. Br. J. Gen. Pract. 69 (680), e171–e181. https://doi.org/10.3399/bjgp19X701297.

Gellert, P., Lech, S., Kessler, E.-M., Herrmann, W., Döpfmer, S., Balke, K., Oedekoven, M., Kuhlmey, A., Schnitzer, S., 2021. Perceived need for treatment and non-utilization of outpatient psychotherapy in old age: two cohorts of a nationwide survey. BMC Health Serv. Res. 21 (1), 442. https://doi.org/10.1186/s12913-021-06384-6.

Georgi, H., Vlckova, K., Lukavsky, J., Kopecek, M., Bares, M., 2019. Beck depression inventory-II: self-report or interview-based administrations show different results in older persons. Int. Psychogeriatr. 31 (5), 735–742. https://doi.org/10.1017/S1041610218001187.

Gonçalves, D.C., Byrne, G.J., 2012. Interventions for generalized anxiety disorder in older adults: systematic review and meta-analysis. J. Anxiety Disord. 26 (1), 1–11. https://doi.org/10.1016/j.janxdis.2011.08.010.

Gottschall, A.C., West, S.G., Enders, C.K., 2012. A comparison of item-level and scale-level multiple imputation for questionnaire batteries. Multivar. Behav. Res. 47 (1), 1–25. https://doi.org/10.1080/00273171.2012.640589.

Gould, R.L., Coulson, M.C., Howard, R.J., 2012a. Cognitive behavioral therapy for depression in older people: a meta-analysis and meta-regression of randomized controlled trials. J. Am. Geriatr. Soc. 60 (10), 1817–1830. https://doi.org/10.1111/j.1532-5415.2012.04166.x.

Gould, R.L., Coulson, M.C., Howard, R.J., 2012b. Efficacy of cognitive behavioral therapy for anxiety disorders in older people: a meta-analysis and meta-regression of randomized controlled trials. J. Am. Geriatr. Soc. 60 (2), 218–229. https://doi.org/10.1111/j.1532-5415.2011.03824.x.

Hall, J., Kellett, S., Berrios, R., Bains, M.K., Scott, S., 2016. Efficacy of cognitive behavioral therapy for generalized anxiety disorder in older adults: systematic review, meta-analysis, and meta-regression. Am. J. Geriatr. Psychiatry 24 (11), 1063–1073. https://doi.org/10.1016/j.jagp.2016.06.006.

Hautzinger, M., Keller, F., Kühner, C., 2009. BDI-II. Beck-Depressions-Inventar. Revision, 2nd ed. Pearson Assessment, Frankfurt am Main, Germany.

Hinrichsen, G.A., Emery-Tiburcio, E.E., Gooblar, J., Molinari, V.A., 2018. Building foundational knowledge competencies in professional geropsychology: Council of Professional Geropsychology Training Programs (CoPGTP) recommendations. Clin. Psychol. Sci. Pract. 25 (2). https://doi.org/10.1111/cpsp.12236.

Hofmann, S.G., Asnaani, A., Vonk, I.J.J., Sawyer, A.T., Fang, A., 2012. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cogn. Ther. Res. 36 (5), 427–440. https://doi.org/10.1007/s10608-012-9476-1.

Holvast, F., Massoudi, B., Oude Voshaar, R.C., Verhaak, P.F.M., 2017. Non-pharmacological treatment for depressed older patients in primary care: a systematic

review and meta-analysis. PLoS One 12 (9), e0184666. https://doi.org/10.1371/journal.pone.0184666.
Hu, T., Zhao, X., Wu, M., Li, Z., Luo, L., Yang, C., Yang, F., 2022. Prevalence of depression in older adults: a systematic review and meta-analysis. Psychiatry Res. 311, 114511. https://doi.org/10.1016/j.psychres.2022.114511.
In-Albon, T., Christiansen, H., Imort, S., Krause, K., Schlarb, A., Schneider, S., Schwarz, D., Weber, L., Velten, J., 2019. Forschungsnetzwerk KODAP. Z. Klin. Psychol. Psychother. 48 (1), 40–50. https://doi.org/10.1026/1616-3443/a000528.
Jacobson, N.S., Truax, P., 1991. Clinical significance : a statistical approach to defining meaningful change in psychotherapy research. J. Consult. Clin. Psychol. 59, 12–19. https://doi.org/10.1037/0022-006X.59.1.12.
Kaiser, T., Brakemeier, E.-L., Herzog, P., 2023. What if we wait? Using synthetic waiting lists to estimate treatment effects in routine outcome data. Psychother. Res. 33 (8), 1043–1057. https://doi.org/10.1080/10503307.2023.2182241.
Karlin, B.E., Trockel, M., Brown, G.K., Gordienko, M., Yesavage, J., Taylor, C.B., 2013. Comparison of the effectiveness of cognitive behavioral therapy for depression among older versus younger veterans: results of a national evaluation. The Journals of Gerontology: Series B 70 (1), 3–12. https://doi.org/10.1093/geronb/gbt096.
Karlin, B.E., Trockel, M., Spira, A.P., Taylor, C.B., Manber, R., 2015. National evaluation of the effectiveness of cognitive behavioral therapy for insomnia among older versus younger veterans. Int. J. Geriatr. Psychiatry 30 (3), 308–315. https://doi.org/10.1002/gps.4143.
Kessler, E.-M., Blachetta, C., 2020. Age cues in patients' descriptions influence treatment attitudes. Aging Ment. Health 24 (1), 193–196. https://doi.org/10.1080/13607863.2018.1515889.
Kessler, E.-M., Schneider, T., 2017. Do treatment attitudes and decisions of psychotherapists-in-training depend on a patient's age? The Journals of Gerontology: Series B 74 (4), 620–624. https://doi.org/10.1093/geronb/gbx078.
Kishita, N., Laidlaw, K., 2017. Cognitive behaviour therapy for generalized anxiety disorder: is CBT equally efficacious in adults of working age and older adults? Clin. Psychol. Rev. 52, 124–136. https://doi.org/10.1016/j.cpr.2017.01.003.
Laidlaw, K., 2021. Cognitive behavioral therapy with older people. In: Handbook of Cognitive Behavioral Therapy: Applications, vol. 2. American Psychological Association, pp. 751–771. https://doi.org/10.1037/0000219-024.
Laidlaw, K., Kishita, N., 2015. Age-appropriate augmented cognitive behavior therapy to enhance treatment outcome for late-life depression and anxiety disorders. GeroPsych 28 (2), 57–66. https://doi.org/10.1024/1662-9647/a000128.
Laidlaw, K., Pachana, N.A., 2009. Aging, mental health, and demographic change: challenges for psychotherapists. Prof. Psychol. Res. Pract. 40 (6), 601–608. https://doi.org/10.1037/a0017215.
Langkaas, T.F., Wampold, B.E., Hoffart, A., 2018. Five types of clinical difference to monitor in practice. Psychotherapy 55 (3), 241–254. https://doi.org/10.1037/pst0000194.
Lavingia, R., Jones, K., Asghar-Ali, A.A., 2020. A systematic review of barriers faced by older adults in seeking and accessing mental health care. J. Psychiatr. Pract. 26 (5). https://doi.org/10.1097/PRA.0000000000000491.
Leyrat, C., Seaman, S.R., White, I.R., Douglas, I., Smeeth, L., Kim, J., Resche-Rigon, M., Carpenter, J.R., Williamson, E.J., 2017. Propensity score analysis with partially observed covariates: how should multiple imputation be used? Stat. Methods Med. Res. 28 (1), 3–19. https://doi.org/10.1177/0962280217713032.
Margraf, J., Cwik, J.C., Pflug, V., Schneider, S., 2017. Strukturierte klinische interviews zur Erfassung psychischer Störungen über die Lebensspanne: Gütekriterien und Weiterentwicklungen der DIPS-Verfahren. [structured clinical interviews for mental disorders across the life span: psychometric quality and further developments of the DIPS open access interviews.]. Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis 46 (3), 176–186. https://doi.org/10.1026/1616-3443/a000430.
Margraf, J., Hoyer, J., Fydrich, T., In-Albon, T., Lincoln, T., Lutz, W., Schlarb, A., Schöttke, H., Willutzki, U., Velten, J., 2021. The cooperative revolution reaches clinical psychology and psychotherapy: an example from Germany. Clinical Psychology in Europe 3 (1), e4459. https://doi.org/10.32872/cpe.4459.
Nathan, P.E., Stuart, S.P., Dolan, S.L., 2000. Research on psychotherapy efficacy and effectiveness: between Scylla and Charybdis? Psychol. Bull. 126 (6), 964–981. https://doi.org/10.1037/0033-2909.126.6.964.
OECD, 2023. Demographic trends. In: Health at a Glance 2021: OECD Indicators. OECD Publishing, Paris. https://doi.org/10.1787/b52b873c-en.
Ouchi, Y., Rakugi, H., Arai, H., Akishita, M., Ito, H., Toba, K., Kai, I., 2017. Redefining the elderly as aged 75 years and older: proposal from the Joint Committee of Japan Gerontological Society and the Japan Geriatrics Society. Geriatr Gerontol Int 17 (7), 1045–1047. https://doi.org/10.1111/ggi.13118.
Pettit, S., Qureshi, A., Lee, W., Stirzaker, A., Gibson, A., Henley, W., Byng, R., 2017. Variation in referral and access to new psychological therapy services by age: an empirical quantitative study. Br. J. Gen. Pract. 67 (660), e453–e459. https://doi.org/10.3399/bjgp17X691361.
Pishgar, F., Greifer, N., Leyrat, C., Stuart, E., 2021. MatchThem:: matching and weighting after multiple imputation. The R Journal 13 (2), 292–305. https://doi.org/10.32614/RJ-2021-073.
Pomerleau, V.J., Sekhon, H., Bajsarowicz, P., Demoustier, A., Rej, S., Myhr, G., 2023. Do older adults respond to cognitive behavioral therapy as well as younger adults? An analysis of a large, multi-diagnostic, real-world sample. Int. J. Geriatr. Psychiatry 38 (6), e5953. https://doi.org/10.1002/gps.5953.
Prinz, U., Nutzinger, D.O., Schulz, H., Petermann, F., Braukhaus, C., Andreas, S., 2013. Comparative psychometric analyses of the SCL-90-R and its short versions in patients with affective disorders. BMC Psychiatry 13 (1), 104. https://doi.org/10.1186/1471-244X-13-104.
R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Rubin, D.B., 1987. Multiple Imputation for Nonresponse in Surveys. Wiley. https://doi.org/10.1002/9780470316696.ch1.
Sanglier, T., Saragoussi, D., Milea, D., Auray, J.-P., Valuck, R.J., Tournier, M., 2011. Comparing antidepressant treatment patterns in older and younger adults: a claims database analysis. J. Am. Geriatr. Soc. 59 (7), 1197–1205. https://doi.org/10.1111/j.1532-5415.2011.03457.x.
Saunders, R., Buckman, J.E.J., Stott, J., Leibowitz, J., Aguirre, E., John, A., Lewis, G., Cape, J., Pilling, S., 2021. Older adults respond better to psychological therapy than working-age adults: evidence from a large sample of mental health service attendees. J. Affect. Disord. 294, 85–93. https://doi.org/10.1016/j.jad.2021.06.084.
Schmitz, N., Hartkamp, N., Franke, G.H., 2000. Assessing clinically significant change: application to the SCL-90–R. Psychol. Rep. 86 (1), 263–274. https://doi.org/10.2466/pr0.2000.86.1.263.
Subica, A.M., Fowler, J.C., Elhai, J.D., Frueh, B.C., Sharp, C., Kelly, E.L., Allen, J.G., 2014. Factor structure and diagnostic validity of the Beck Depression Inventory–II with adult clinical inpatients: comparison to a gold-standard diagnostic interview. Psychol. Assess. 26 (4), 1106–1115. https://doi.org/10.1037/a0036998.
Tegeler, C., Beyer, A.K., Hoppmann, F., Ludwig, V., Kessler, E.M., 2020. Current state of research on psychotherapy for home-living vulnerable older adults with depression. Z. Gerontol. Geriatr. 53 (8), 721–727. https://doi.org/10.1007/s00391-020-01805-3.
Teismann, T., Brakemeier, E.L., Brockmeyer, T., Christiansen, H., Fehm, L., Forkmann, T., Glombiewski, J., Heider, J., Hermann, A., Hoyer, J., In-Albon, T., Kaiser, T., Klucken, T., Lincoln, T.M., Lutz, W., Margraf, J., Odyniec, P., Pedersen, A., Renneberg, B.,…Velten, J., 2024a. Amelioration of suicidal ideation in routine care psychotherapy: preliminary findings from a large multicenter assessment. Journal of Affective Disorders Reports 100843. https://doi.org/10.1016/j.jadr.2024.100843.
Teismann, T., Forkmann, T., Glaesmer, H., Alpers, G.W., Brakemeier, E.L., Brockmeyer, T., Christiansen, H., Fehm, L., Glombiewski, J., Heider, J., Hermann, A., Hoyer, J., Kaiser, T., Klucken, T., Lincoln, T.M., Lutz, W., Margraf, J., Pedersen, A., Renneberg, B., In-Albon, T., 2024b. Prevalence of suicidal ideation in German psychotherapy outpatients: a large multicenter assessment. J. Affect. Disord. 351, 971–976. https://doi.org/10.1016/j.jad.2024.02.019.
Thombs, B.D., Ziegelstein, R.C., Pilote, L., Dozois, D.J., Beck, A.T., Dobson, K.S., Fuss, S., de Jonge, P., Grace, S.L., Stewart, D.E., Ormel, J., Abbey, S.E., 2010. Somatic symptom overlap in Beck Depression Inventory-II scores following myocardial infarction. Br. J. Psychiatry 197 (1), 61–66. https://doi.org/10.1192/bjp.bp.109.076596.
Trentini, C.M., Xavier, F.M., Chachamovich, E., Rocha, N.S., Hirakata, V.N., Fleck, M.P., 2005. The influence of somatic symptoms on the performance of elders in the Beck Depression Inventory (BDI). Brazilian Journal of Psychiatry 27 (2), 119–123. https://doi.org/10.1590/s1516-44462005000200009.
van Buuren, S., Groothuis-Oudshoorn, K., 2011. mice: multivariate imputation by chained equations in R. J. Stat. Softw. 45 (3), 1–67. https://doi.org/10.18637/jss.v045.i03.
Velten, J., Margraf, J., Benecke, C., Berking, M., In-Albon, T., Lincoln, T., Lutz, W., Schlarb, A., Schöttke, H., Willutzki, U., Hoyer, J., 2017. Methodenpapier zur Koordination der Datenerhebung und -auswertung an Hochschul- und Ausbildungsambulanzen für Psychotherapie (KODAP). Z. Klin. Psychol. Psychother. 46 (3), 169–175. https://doi.org/10.1026/1616-3443/a000431.
Velten, J., Christiansen, H., Hoyer, J., In-Albon, T., Lincoln, T., Lutz, W., Margraf, J., Schöttke, H., Stark, R., Werheid, K., Willutzki, U., Alpers, G.W., Bartholdy, S., Berger, E.-M., Brakemeier, E.-L., Bräscher, A.-K., Brockmeyer, T., Dziobek, I., Fehm, L., Rubel, J.A., 2025. Effectiveness of cognitive behavioral therapy for adult mental disorders: A large-scale naturalistic study across 29 university outpatient clinics. Behav. Res. Ther. 104691. https://doi.org/10.1016/j.brat.2025.104691.
Volkert, J., Schulz, H., Härter, M., Wlodarczyk, O., Andreas, S., 2013. The prevalence of mental disorders in older people in Western countries – a meta-analysis. Ageing Res. Rev. 12 (1), 339–353. https://doi.org/10.1016/j.arr.2012.09.004.
von Glischinski, M., von Brachel, R., Hirschfeld, G., 2019. How depressed is "depressed"? A systematic review and diagnostic meta-analysis of optimal cut points for the Beck Depression Inventory revised (BDI-II). Qual. Life Res. 28 (5), 1111–1118. https://doi.org/10.1007/s11136-018-2050-x.
Waller, G., Turner, H., 2016. Therapist drift redux: why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. Behav. Res. Ther. 77, 129–137. https://doi.org/10.1016/j.brat.2015.12.005.
Wedding, U., Koch, A., Röhrig, B., Pientka, L., Sauer, H., Höffken, K., Maurer, I., 2007. Requestioning depression in patients with cancer: contribution of somatic and affective symptoms to Beck's Depression Inventory. Ann. Oncol. 18 (11), 1875–1881. https://doi.org/10.1093/annonc/mdm353.
Werson, A.D., Meiser-Stedman, R., Laidlaw, K., 2022. A meta-analysis of CBT efficacy for depression comparing adults and older adults. J. Affect. Disord. 319, 189–201. https://doi.org/10.1016/j.jad.2022.09.020.
Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., Zaudig, M., 1997. SKID I. Strukturiertes Klinisches Interview für DSM-IV. In: Achse I: Psychische Störungen. Interviewheft und Beurteilungsheft. Eine deutschsprachige, erweiterte Bearb. d. amerikanischen Originalversion des SKID I. Hogrefe.