

The Efficacy of Personalization Within Psychological Treatments Compared to No or Less Personalization: A Meta-Analysis and Systematic Review

Susan J. Harnas^{1, 2, 3}, Kim F. Francken^{1, 2, 3}, Fabiola Müller^{1, 2, 3}, Pythia T. Nieuwkerk^{1, 3, 4},
Mirjam A. G. Sprangers^{1, 2, 3}, Hans Knoop^{1, 2, 3}, and Annemarie M. J. Braamse^{1, 2, 3}

¹ Department of Medical Psychology, Amsterdam University Medical Center Location University of Amsterdam

² Cancer Treatment and Quality of Life, Cancer Center Amsterdam, Amsterdam, The Netherlands

³ Mental Health, Amsterdam Public Health, Amsterdam, The Netherlands

⁴ Infectious Diseases, Amsterdam Institute for Immunology and Infectious Diseases, Amsterdam, The Netherlands

Objective: This meta-analysis and systematic review assessed whether personalization within psychological treatments for mental health problems or persistent somatic symptoms improves treatment outcomes, compared to non- or less personalized treatments. **Method:** APA PsycInfo, Cochrane Library, MEDLINE, and EMBASE were searched up to January 14, 2024, for randomized controlled trials comparing within-treatment personalization against no or less personalization. Risk of bias was assessed with the Cochrane tool. A meta-analysis was performed using a random-effects model. Additionally, a three-level meta-analysis was conducted, and moderation analyses were performed. A narrative synthesis was included. **Results:** Sixteen studies were included. Risk of bias was low for two studies and high for three studies, and 11 studies had some concerns. Effect sizes were calculated for the subgroups: symptoms, smoking cessation, and treatment process outcomes. For symptoms, the pooled standardized mean difference was 0.07 (95% CI [−0.06, 0.20], $p = .28$); for smoking cessation, the pooled odds ratio was 1.12 (95% CI [0.84, 1.51], $p = .43$); and for treatment process outcomes, the pooled standardized mean difference was 0.29 (95% CI [−0.27, 0.85], $p = .31$). Treatment format (in person/online), personalization factor (preference/individual profile), or personalized treatment aspect (content/modules) did not moderate the effect of personalization. **Conclusions:** Evidence of moderate quality does not convincingly suggest that within-treatment personalization outperforms no or less personalization with respect to treatment outcome. Using evidence-based personalization strategies, future studies should clarify which degree of personalization yields clinically relevant effects for which populations, interventions, and outcomes.

What is the public health significance of this article?

Intuitively, personalization of psychological treatments is expected to enhance treatment outcomes. However, personalization is a broad concept, and psychological treatments can be personalized in different ways. Our results show that within-treatment personalization (e.g., adjusting dosages, strategies, or modules within a single treatment) does not consistently show clear benefits over non- or less personalized versions of the same or comparable treatment. More research is needed to better understand when personalization of psychological treatments adds value and which specific approaches are most effective.

Keywords: systematic review, meta-analysis, personalizing, tailoring, psychological treatment

Supplemental materials: <https://doi.org/10.1037/ccp0000964.supp>

Michael J. Constantino served as action editor.

Annemarie M. J. Braamse  <https://orcid.org/0000-0003-3216-6753>

The protocol for this study was registered with the International Prospective Register of Systematic Reviews (Registration No. CRD42022337405). Annemarie M. J. Braamse received funding from the Koningin Wilhelmina Fonds Dutch Cancer Society (Grant 11351).

Susan J. Harnas played a lead role in formal analysis, methodology, project administration, software, and writing—original draft and a supporting role in conceptualization and writing—review and editing. Kim F. Francken played a supporting role in formal analysis, project administration, and writing—review and editing. Fabiola Müller played a supporting role in conceptualization, methodology, supervision, and writing—review and editing. Pythia T. Nieuwkerk played a supporting role in formal analysis,

methodology, software, and writing—review and editing. Mirjam A. G. Sprangers played a supporting role in conceptualization, funding acquisition, methodology, and writing—review and editing and an equal role in supervision. Hans Knoop played a supporting role in funding acquisition, writing—original draft, and writing—review and editing and an equal role in conceptualization, methodology, and supervision. Annemarie M. J. Braamse played a lead role in conceptualization, funding acquisition, and writing—review and editing, a supporting role in formal analysis, project administration, software, and writing—original draft, and an equal role in methodology and supervision.

Correspondence concerning this article should be addressed to Annemarie M. J. Braamse, Department of Medical Psychology, Amsterdam University Medical Center Location University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. Email: a.m.braamse@amsterdamumc.nl

Psychological interventions for common mental health problems and/or persistent somatic symptoms often show small to moderate effect sizes when compared to care as usual or a waiting list (e.g., Cuijpers et al., 2016; Leichsenring et al., 2022; Liu et al., 2019; van Dessel et al., 2014). Continuous efforts are directed at improving the efficacy of interventions. Personalization of treatments is one strategy to improve treatment outcomes. A recent systematic review on defining and operationalizing personalized psychological treatment led to the following comprehensive definition:

A personalized psychological treatment aims to optimize treatment outcome for the individual patient by tailoring (specific elements of) treatment to unique or specific needs, preferences or other characteristics of an individual patient. It includes a systematic adaptation of treatment or a differentiation between treatment strategies based on some form of assessment, which is analyzed according to a pre-defined algorithm and applied either at the start of treatment and/or during the course of treatment. (Harnas et al., 2024, p. 478)

This definition reflects the breadth of the concept of personalization, which can be operationalized in different ways. Examples include adjusting dosages, strategies, or modules within a single treatment (within-treatment personalization); selecting between different types of treatments (between-treatment personalization); or matching therapists to patients based on patients' needs, preferences, or other characteristics. In this study, we focus on within-treatment personalization, which is applied in many symptom-focused psychological treatment studies.

Another important aspect of the definition is the systematization of personalization. Personalization should be guided by a predefined decision rule, which is replicable and based on assessments of the individual patient. Personalizing procedures has long been criticized for generally not being systematic and replicable (Kazdin, 2008; Piccirillo et al., 2019). Rapid technological developments for collecting and analyzing idiographic data provided new opportunities for the systematic personalization of psychological treatments (Piccirillo & Rodebaugh, 2019), which fueled the attention for personalization (Balaskas et al., 2021; Ng & Weisz, 2016; Norcross & Wampold, 2011), including the evaluation of its efficacy.

Personalization is expected to improve treatment efficacy and patient adherence, as it is assumed that the individual tailoring of treatment optimally matches patients' needs (U.S. Department of Health and Human Services, National Institute of Occupational Health, National Institute of Mental Health, 2008, 2015, 2020; van Beugen et al., 2014). However, head-to-head comparisons between personalized psychological treatments and a standard version of the same or a comparable treatment are rarely executed.

Two previous meta-analyses included studies targeting the comparison between personalized and non- or less personalized psychological care (Nye et al., 2023; Păsărelu et al., 2017). However, other comparison groups, including waitlist or active control groups with a different treatment, were also included, as well as treatments that were not individually tailored. These comparisons make it difficult to assess the added value of personalization itself, as differences in outcomes might be attributable to other factors, such as treatment intensity or type, rather than personalization. Only three studies (Berger et al., 2014; Coates et al., 2018; Johansson et al., 2012) in the two meta-analyses compared the efficacy of personalized treatment (PT) to a nonpersonalized version of the same treatment, of which Nye et al. (2023) included all three and Păsărelu et al.

(2017) two studies. These three studies did not show a significant difference in outcomes between the two conditions (Berger et al., 2014; Coates et al., 2018; Johansson et al., 2012), which was not reflected in the overall conclusion of both meta-analyses as they both emphasized the promise that personalization of treatment improves outcomes. With growing attention for personalization and claims about its potential benefits, it is essential to obtain a comprehensive overview of the added value of personalization itself. Therefore, we conducted a systematic review of randomized controlled trials (RCTs) in which personalization within treatment was compared to a standard version of the same or comparable treatment, investigating whether former results can be replicated.

Method

This review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (Page et al., 2021).

Eligibility Criteria

RCTs were eligible that fulfilled the following inclusion criteria: (a) Psychological treatment was individually tailored; (b) the individually tailored treatment was compared to a version of the same treatment in which no or fewer aspects of treatment were personalized; (c) the individual tailoring was applied in a systematic and replicable manner and was clearly described, including what aspect of the psychological treatment was personalized, which decision rule was used for the personalization, and at what time point(s) the personalization was applied; (d) the study population consisted of adult patients (≥ 18 years) receiving psychological treatment for a mental health problem and/or persistent somatic symptom (e.g., fatigue, pain); (e) the psychological treatment itself was an intervention aimed at modifying feelings, cognitions, and/or behaviors to stimulate coping with mental health problems and/or persistent somatic symptoms or to reduce such problems and symptoms; (f) study articles had to be published in peer-reviewed scientific journals and written in English or Dutch. Study protocols, conference abstracts, secondary analyses of RCTs, and studies evaluating group therapy were excluded.

Information Sources

APA PsycInfo (Ovid), Cochrane Library (Wiley), MEDLINE (Ovid), and EMBASE (Ovid) were searched up to January 14, 2024. The final search results were exported into EndNote, and duplicates were removed with a systematic deduplication method (Bramer et al., 2016). Included articles were used as a source to identify missing studies through backward citation searching (check for cited articles) and forward citation searching (check for citing articles).

Search Strategy

The search terms included a variety of terms for psychological treatment and RCT. Other keywords for the search included personaliz* or personalis* or tailor* or individualiz* or individualis* or individually adapt*. The final search strategy for APA PsycInfo (Ovid) can be found in Supplemental Material S1.

Selection Process

Titles and abstracts were independently screened by two authors (SH and KF), using the software tool Rayyan (Ouzzani et al., 2016). After every 5%–10% of screened titles and abstracts, the interrater reliability was calculated, and discrepancies were discussed until consensus was reached. After a kappa of at least 0.6 was reached, the remaining titles and abstracts were screened by one author (KF). A second author was consulted in case of doubt (SH). All full-text articles were independently screened for their eligibility by two authors (SH and KF), and again, discrepancies were discussed until consensus was reached. A third author (AB) was consulted when needed.

Data Collection Process

Only data of the specific study groups of interest were extracted and reported (i.e., personalized treatment group vs. non- or less personalized treatment group). Data from additional study arms (e.g., waitlist group, care as usual group) were excluded. One author (KF) extracted the data of the included studies, using a standardized data extraction sheet, which was checked by a second author (SH). Discrepancies were discussed, and a third author (AB) was consulted when necessary.

Data Items

The following data were extracted from the included studies in so far they were relevant for this review: (a) name of first author; (b) year of publication; (c) study population; (d) sample size (total and per treatment arm); (e) type of psychological treatment (e.g., cognitive behavior therapy); (f) treatment format (e.g., face-to-face, online); (g) objective(s) of personalization (e.g., increasing effectiveness); (h) aspect(s) of treatment that was/were personalized (e.g., treatment content, treatment duration); (i) the factor(s) on which the personalization was based (e.g., unique needs or preferences); (j) how the aforementioned factor(s) was/were measured (e.g., questionnaires); (k) the decision rule(s) on which the personalization was based (e.g., cutoff score on a questionnaire); (l) the time point(s) of the applied personalization (at the start and/or during treatment); (m) characteristics of the comparator intervention (control condition); (n) primary time point and other time point(s); (o) primary outcome(s); (p) result(s) of primary outcome(s), including information on the specific statistical tests employed (e.g., *F/T* tests) and corresponding *p* values and effect sizes, if reported; (q) secondary outcome(s); (r) result(s) of secondary outcomes, including information on the specific statistical tests employed (e.g., *F/T* tests) and corresponding *p* values and effect sizes, if reported.

Of note, we differentiated between primary and secondary outcomes because studies are generally powered to detect effects on the primary outcome, while analysis of the effects on secondary outcomes is more often exploratory. Similarly, we distinguished between the primary time point, which is the main measurement moment as specified in the original study design, and additional time points, which refer to other measurements completed at different time points. This approach allowed us to align our analyses as closely as possible with the study designs and analyses on which the studies were powered.

Study Risk of Bias Assessment

The Cochrane risk of bias tool for RCTs was used to assess the risk of bias (Sterne et al., 2019). For each study, an overall risk of bias was determined, categorized as high, low, or with some concerns. The studies were independently rated for risk of bias by two authors (SH, KF). Results were compared, and discrepancies were resolved. If needed, a third author (AB) was consulted.

Data Synthesis and Reporting

The data extraction for the meta-analysis was conducted by two reviewers (SH, AB) independently, and any discrepancies were resolved through discussion. We first used a random-effects model with inverse variance weights to aggregate individual effect sizes into a pooled effect estimate with a 95% confidence interval using Review Manager 5.4. For each study, we selected the primary outcome and primary time point. In studies with multiple primary outcomes, we used the outcome on which the sample size calculation was based. If this was not reported, we used the outcome with the largest effect, favoring personalized treatment, as this approach provides the best chance to detect a potential effect of personalization. In case the direction of a (nonsignificant) result was not reported, we favored the personalized condition in order to provide the best case scenario for personalization. For studies reporting multiple primary time points, we selected the first postintervention measurement for consistency. In cases where studies included multiple groups of personalized treatments compared to standard treatment, each comparison was treated as an independent data point and included separately in the meta-analysis.

Based on the nature of the primary outcomes reported in the included studies (i.e., post hoc), we organized our analysis into three main outcome categories: symptoms, smoking abstinence, and outcomes related to the treatment process. Studies in the symptom category were further divided into the following subgroups based on the primary symptom measured: (a) depression, anxiety, or distress and (b) other symptoms. For studies reporting symptoms and outcomes related to the treatment process, we extracted the standardized mean difference (SMD) and standard error as primary effect measures. When these values were not directly provided, we calculated them based on available data, such as means, standard deviations, and sample sizes. SMD values of 0.2, 0.5, and 0.8 were interpreted as small, medium, and large effects, respectively (Cohen, 1988).

For studies assessing smoking abstinence, we extracted odds ratios (*ORs*) as effect measures. If *ORs* were not provided, we calculated them from numbers of abstinent versus nonabstinent participants, proportions, and sample sizes or effect sizes.

To evaluate heterogeneity within each outcome category, we calculated the I^2 statistic. I^2 values were interpreted as follows: 25% indicating low heterogeneity, 50% moderate heterogeneity, and 75% high heterogeneity.

In addition, we conducted a three-level meta-analysis, allowing us to include multiple outcomes per study and model the dependency between studies. This type of analysis takes into account three sources of variance: the sampling variance of the observed effect sizes (Level 1), variance between effect sizes extracted from the same study (Level 2), and variance between studies (Level 3; Van den Noortgate et al., 2013). For the abstinence studies, *ORs* were

log-transformed prior to analysis, and estimates were back-transformed after analysis. If possible, we performed moderator analyses to evaluate whether the treatment format (in person vs. online), factor on which the personalization was based (preference vs. individual profile), and personalized aspect of treatment (treatment content vs. modules) were associated with effectiveness. For these analyses, we used R Studio (Version 4.3.3), the metafor package (Viechtbauer, 2015), and the R syntax as described by Assink and Wibbelink (2016), applying a 5% significance level.

Further, we performed a narrative synthesis in which we included all outcomes as defined in the included studies. We first described the results of the comparison between the personalized treatment group and the non- or less personalized treatment group on the primary outcome(s) at the predefined primary time point. If the primary time point was not specified, we chose the first posttreatment assessment. Results of the primary outcomes at other time points and of secondary outcome(s) were described separately. Finally, we examined whether studies could be grouped based on the study population, personalization method, and outcome(s).

Reporting Bias Assessment

If studies only reported an overall effect that was also based on a comparison group not relevant for the current review, we contacted the authors via email to inquire about any significant differences in the primary or secondary outcomes between the study groups of interest. If we did not receive a response within 3 weeks, we sent a reminder. If we did not receive a response within another 3 weeks, we excluded the study from the analysis.

Certainty Assessment

The strength of the scientific evidence for personalized treatment compared to non- or less personalized treatment was established according to the Grading of Recommendations Assessment, Development, and Evaluation system (Schünemann et al., 2023). This system includes four levels of certainty: high, moderate, low, and very low. The level of certainty can be downgraded based on the presence of risk of bias or limitations in the design and implementation, inconsistency or heterogeneity of results, indirectness of evidence, imprecision of results, and risk of publication bias. The level of certainty can be increased based on the presence of large effects, a dose–response gradient, and plausible control for confounding.

Transparency and Openness

Transparency and openness guidelines were followed (Nosek et al., 2015). The protocol for this meta-analysis and systematic review, including plans for the search strategy, data extraction, and analysis, was registered with the International Prospective Register of Systematic Reviews (Registration No. CRD42022337405) prior to conducting the literature search. Of note, the moderation analyses were added following peer reviewing advice after article submission and are not mentioned in the protocol. The final search strategy for APA PsycInfo (Ovid) is included in Supplemental Material S1.

Results

Study Selection

A flow diagram of the search, the selection procedure, and reasons for exclusion is provided in Figure 1. The required kappa of ≥ 0.6 was reached after screening 1,292 records.

Eighteen studies met the inclusion criteria. Three of the 18 included articles did not report on the efficacy of personalized treatment compared with a non- or less personalized treatment. Instead, an overall effect across multiple treatments was described (Batterham et al., 2018; Segal et al., 2022; Strecher et al., 2005). After contacting the corresponding authors, one study was retained in the meta-analysis and narrative synthesis after receiving relevant information from the authors (Batterham et al., 2018). The other two studies were excluded as the authors responded that the data of interest were not available for this specific question (Segal et al., 2022) or we did not receive a response (Strecher et al., 2005). Altogether, this resulted in the inclusion of 16 studies for the meta-analysis and narrative synthesis.

Study Characteristics

Details about the study characteristics per study can be found in Supplemental Material S2. The 16 included studies were published between 2004 and 2022. Six studies targeted people with anxiety or depressive disorders, and four studies focused on smokers.¹ Two studies included people with pain, and the four remaining studies included people with an alcohol use disorder, obsessive compulsive disorder, and insomnia or experiencing partner violence. The total sample sizes for the relevant groups ranged from 40 to 2,872. Most studies ($n = 12$) investigated the efficacy of (cognitive) behavioral therapy. In the remaining four studies, the treatment consisted of a safety and health intervention based on principles of trauma- and violence-informed care,² health education, a combination of different treatment strategies, or an intervention based on the trans-theoretical model of change. The treatments in nine studies were internet-based, three were delivered face-to-face, two involved printed self-help materials, and in one study, an app was used. In another study, the treatment format was unclear.

Personalized Treatment Condition

Details of the personalized condition per study can be found in Supplemental Material S3.

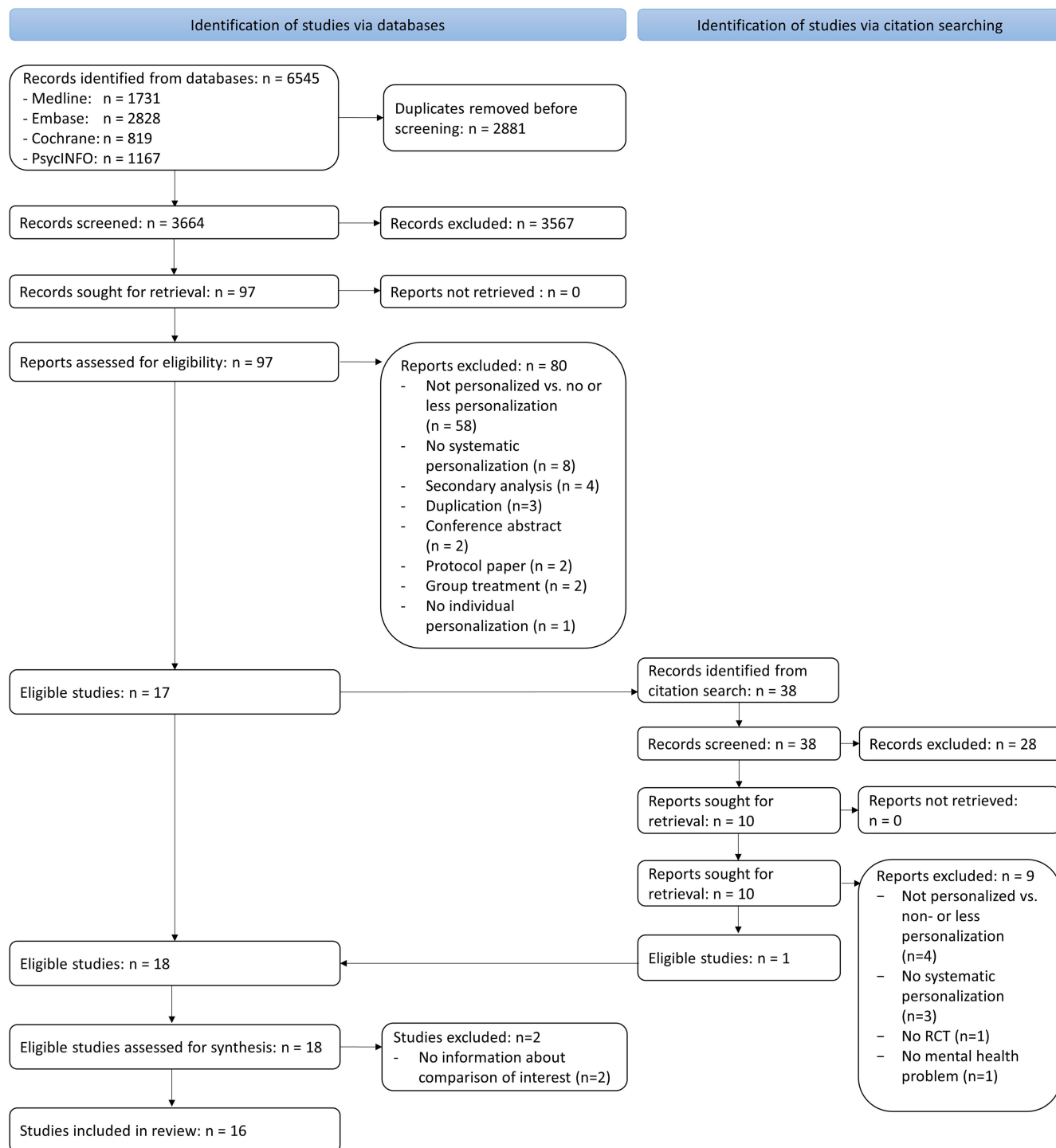
Objective of Personalization. Most studies ($n = 14$) primarily aimed to improve treatment efficacy by reducing or alleviating symptoms. Other aims of personalization were improving treatment efficiency and process outcomes, such as treatment adherence, treatment engagement, treatment satisfaction, or perceived control over treatment.

Aspects of Treatment That Were Personalized. In eight studies, treatment was divided into treatment modules, and the

¹ In alignment with our inclusion criteria, smoking cessation was included as a mental health problem because addiction is classified as a *Diagnostic and Statistical Manual of Mental Disorders* diagnosis.

² Safety and health interventions, as well as health education, were included because in these studies, they aimed to modify feelings, cognitions, and behaviors to help manage mental health problems or persistent somatic symptoms, which aligns with our criteria for psychological treatments.

Figure 1
Flow Diagram of Study Selection



Note. RCT = randomized controlled trial. See the online article for the color version of this figure.

selection and/or order of treatment modules was personalized. In the other eight studies, personalization concerned the treatment content (e.g., information, exercises, or feedback). Additionally, treatment intensity (e.g., amount of information or feedback,

number of modules, exercises, or frequency of contact with therapist) and type of treatment (e.g., a symptom-specific treatment for anxiety, worry, stress, sleep disorders, or depression) were personalized.

Factors on Which the Personalization Was Based and How These Were Measured. The personalization was mostly based on one or more characteristic(s) of the participant forming an individual profile ($n = 12$). Characteristics concerned demographic attributes and/or the presence of one or more symptoms, behaviors, or disorders. The preference of the participant (e.g., concerning the allocation of treatment modules) was the factor on which the personalization was (partly) based in two former studies and in the four remaining studies. Individual profiles and preferences were predominantly assessed with questionnaires or interviews.

Decision Rule for the Personalization. In most studies ($n = 10$), the decision rule for the personalization of treatment consisted of a (computerized) algorithm based on responses to one or more questionnaires or an assessment. For example, when a questionnaire score exceeded a predetermined threshold, a specific treatment module was allocated. In three other studies, the decision rule for the personalization was based on the rank order of the questionnaire scores. For example, a specific treatment module was selected based on the highest standardized pretreatment score on one of three questionnaires, each being associated with a different treatment module. In three studies, the decision rule for personalization followed from the participant's preference. For example, module selection and timing of support were determined by the participant's preference. The participant's preference was also taken into account in two former studies, which suggested specific treatment content based on the participant's baseline scores, after which the participant could choose which suggestion to follow.

Time Point(s) of Personalization. In all studies, the personalization took place before the start of treatment.

Control Condition

In 12 studies, the personalized treatment arm was compared to a nonpersonalized version of the same treatment. In three studies, the

personalized treatment was compared to a treatment in which fewer aspects of treatment were personalized; for example, the content of treatment was personalized based on fewer patient characteristics. One study compared the personalized treatment to a control condition with the same number of sessions and treatment modules. However, the allocation of treatment modules in the control condition was based on the preferences of a matched counterpart (based on perceived interest in and importance of a treatment module) assigned to the personalized condition, instead of the control participant's own preference.

Risk of Bias in Studies

A summary of the risk of bias assessment for each included study is provided in Figure 2. Two studies had an overall low risk of bias, 11 studies had some concerns, and three studies had an overall high risk of bias. The most common source of bias was related to the selection of the reported results (D5; some concerns, $n = 12$; high risk of bias, $n = 1$), which mainly reflected the observation that analyses were not registered beforehand ($n = 8$). Of the preregistered studies ($n = 8$), four studies did not include an analysis plan. The other four studies performed the analyses as described. The next common source of bias concerned the measurement of the outcome (D4; high risk of bias, $n = 3$; some concerns, $n = 3$), where the participant's awareness of the intervention received may have influenced his/her assessment of the outcome.

Meta-Analysis

For the first meta-analysis, a total of 16 studies were included, of which two studies included multiple groups of personalized treatments, resulting in 18 comparisons of interest.

Figure 2

The Risk of Bias Rating for Each Included Study According to the Cochrane Risk of Bias Tool for Randomized Controlled Trials

	D1	D2	D3	D4	D5	Overall		D1	D2	D3	D4	D5	Overall
Batterham, 2018 (26)	+	!	+	+	!	!	Moritz, 2016 (39)	+	+	+	+	!	!
Berger, 2014 (18)	+	+	+	!	!	!	Nissling, 2021 (30)	+	!	-	-	-	-
Coates, 2018 (20)	+	!	+	+	!	!	Okajima, 2020 (40)	+	+	+	+	+	+
Dahlin, 2022 (29)	+	!	+	!	!	!	Sauer-Zavala, 2022 (31)	!	-	+	-	!	-
Dalton, 2004 (37)	!	!	+	!	!	!	Six, 2022 (32)	+	+	!	+	+	!
Etter, 2009 (33)	+	+	!	+	!	!	Strecher, 2008 (34)	+	!	!	+	!	!
Ford-Gilboe, 2020 (41)	+	+	+	+	+	+	Van der Aalst, 2012 (35)	!	!	!	-	!	-
Kerns, 2014 (38)	!	+	+	+	!	!	Webb Hooper, 2013 (36)	!	!	!	+	!	!

Note. D1 = randomization process; D2 = deviations from the intended interventions; D3 = missing outcome data; D4 = measurement of the outcome; D5 = selection of the reported result; green = low risk of bias; yellow = some concerns; red = high risk of bias. See the online article for the color version of this figure.

Symptom Outcomes

For depression, anxiety, and distress, seven comparisons were included. Outcome measures in this category included a composite measure of multiple symptom questionnaires, Beck Depression Inventory, Penn State Worry Questionnaire, Overall Anxiety Severity and Impairment Scale, Overall Depression Severity and Impairment Scale, and Patient Health Questionnaire. As shown in Figure 3, the pooled SMD was 0.15 (95% CI [-0.06, 0.36], $p = .17$), indicating a nonsignificant difference between personalized and non- or less personalized treatment. Heterogeneity was low ($I^2 = 5\%$).

For other symptoms, five comparisons were included. Outcome measures in this category included proportion of drinking days over treatment, least pain, Posttraumatic Stress Disorder Checklist–Civilian Version, Yale-Brown Obsessive Compulsive Scale, and Insomnia Severity Index. The pooled SMD was 0.03 (95% CI [-0.14, 0.20], $p = .71$), demonstrating a nonsignificant difference between personalized and non- or less personalized treatment. Heterogeneity was low ($I^2 = 23\%$).

Finally, the overall pooled SMD for all symptom outcomes (combining depression, anxiety, distress, and other symptoms) was 0.07 (95% CI [-0.06, 0.20], $p = .28$), reflecting a nonsignificant difference between personalized and non- or less personalized treatment. Heterogeneity across the overall symptom outcomes was low ($I^2 = 14\%$).

Abstinence Outcomes

For smoking cessation, four comparisons were included. Outcome measures in this category included the proportion of

participants abstained from smoking in the past 24 hr, 7-day point-prevalence abstinence, and prolonged smoking abstinence. As shown in Figure 4, the pooled OR was 1.12 (95% CI [0.84, 1.51], $p = .43$), suggesting no significant difference in abstinence rates between personalized and non- or less personalized treatment. Heterogeneity was high ($I^2 = 75\%$).

Treatment Process Outcomes

Across two comparisons, the pooled SMD for treatment process outcomes was 0.29 (95% CI [-0.27, 0.85], $p = .31$), reflecting a nonsignificant difference between personalized and non- or less personalized treatment (see Figure 5). Outcome measures in this category included adherence ratings and perceived control over treatment. Heterogeneity was moderate to high ($I^2 = 61\%$).

Three-Level Meta-Analysis and Moderation Analyses

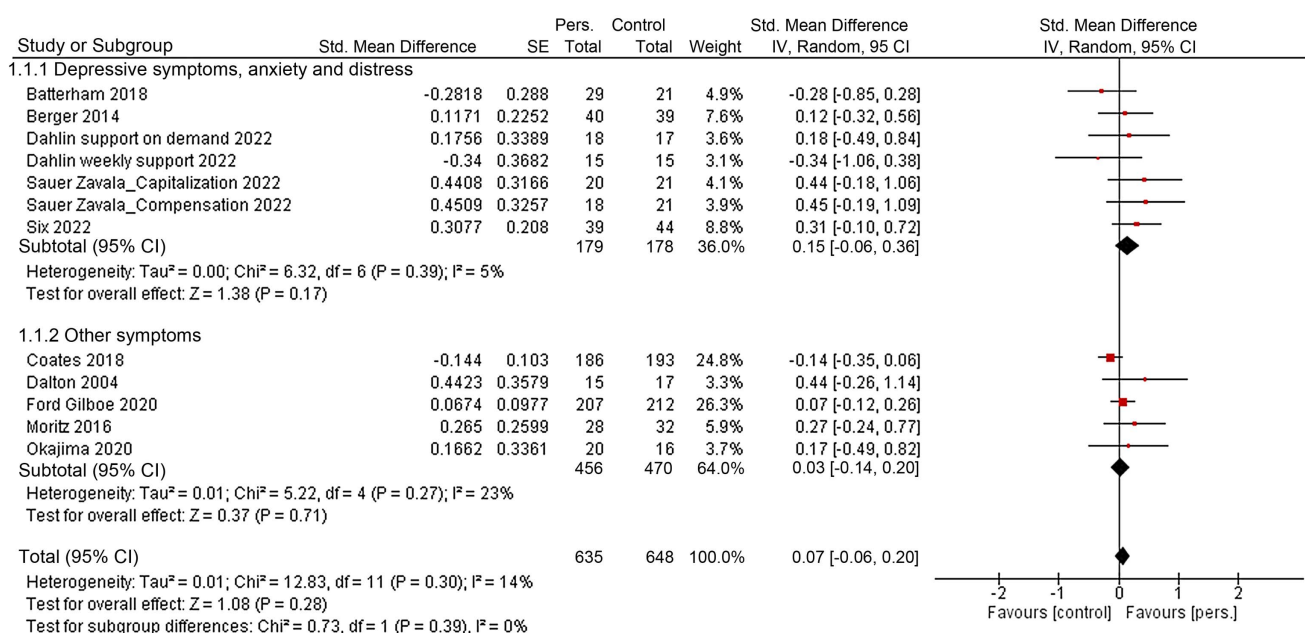
For these analyses, the same 16 studies were included, containing 35 comparisons of interest (see Table 1).

Symptom Outcomes

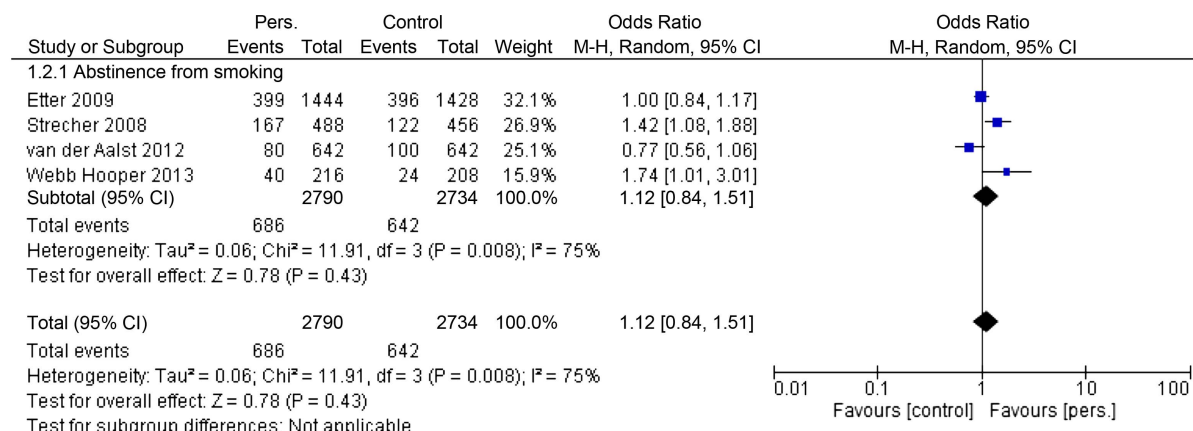
Twenty-four comparisons were included. The overall effect of personalized treatment (compared to non- or less personalized treatment) on symptom outcome was not significant, with a small effect size of $d = 0.07$ (95% CI [-0.052, 0.193], $p = .247$). None of the factors moderated the overall effect: treatment format, $\beta = 0.039$ (95% CI [-0.234, 0.312], $p = .771$); factor on which it was personalized, $\beta = -0.089$ (95% CI [-0.482, 0.303], $p = .642$); treatment aspect, $\beta = -0.059$ (95% CI [-0.320, 0.203], $p = .647$).

Figure 3

Efficacy of Personalized Versus Non- or Less Personalized Treatment on Symptom Outcomes



Note. Std. Mean Difference = standardized mean difference; SE = standard error; Pers. = perceived; CI = confidence interval; IV = inverse variance. See the online article for the color version of this figure.

Figure 4*Efficacy of Personalized Versus Non- or Less Personalized Treatment on Abstinence Outcomes*

Note. Pers. = perceived; CI = confidence interval; M-H = Mantel-Haenszel. See the online article for the color version of this figure.

Abstinence Outcomes

Seven comparisons were included. The overall effect of personalized treatment (compared to non- or less personalized treatment) was not significant: $OR\ 1.12$ (95% CI [0.73, 1.71], $p = .535$). As all putative moderators were the same across studies, no moderation analyses were performed.

Process Outcomes

Five comparisons were included. The overall effect of personalized treatment (compared to non- or less personalized treatment) on process outcome was not significant, with a small effect size of $d = 0.12$ (95% CI [-0.130, 0.379], $p = .247$). None of the factors moderated the effect: treatment format, $\beta = 0.195$ (95% CI [-1.312, 1.701], $p = .708$); treatment aspect, $\beta = 0.269$ (95% CI [-1.190, 1.729], $p = .598$).

Narrative Synthesis

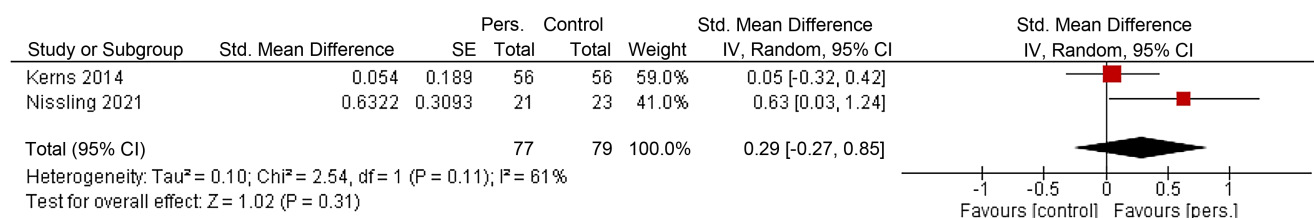
Efficacy on Primary Outcome(s) at Primary Time Point

Table 1 summarizes the efficacy results per study for the primary outcome(s) at the primary time point. The 16 studies included 32 predefined primary outcomes, of which 31 were reported in the

results. Thirteen studies, including 29 primary outcomes, did not find significant differences between personalized and non- or less personalized treatment on their primary outcome(s) at the predefined primary time point. Three studies with each one primary outcome did find significant differences at the predefined primary time point, favoring the personalized condition (9% of the total of 32 predefined primary outcomes).

Efficacy on Symptoms, Behaviors, or Diagnoses. Two studies that found superiority of personalization on their primary outcome concerned 7-day abstinence rates of smoking. Higher abstinence rates were observed among smokers at 3-month or 6-month follow-up in the personalized condition. In these two studies, treatment content was personalized based on individual profiles (e.g., demographic information and characteristics relevant to smoking [cessation]). In five other studies, including 13 assessments of symptoms, behaviors, or diagnoses as primary outcomes and in which treatment content was also personalized based on an individual profile, no significant differences between treatment arms were found. Two of these studies were also conducted within a sample of smokers and included abstinence rate as their primary outcome.

Efficacy on Process Outcomes. The third study for which a significant difference was found on the primary outcome in favor of the personalized condition concerned participants' sense of control

Figure 5*Efficacy of Personalized Versus Non- or Less Personalized Treatment on Treatment Process Outcomes*

Note. Std. Mean Difference = standardized mean difference; SE = standard error; Pers. = perceived; CI = confidence interval; IV = inverse variance. See the online article for the color version of this figure.

Table 1
Efficacy of Personalization on Primary Outcome(s) per Study

First author, year of publication (reference)	Relevant sample size (PT vs. non- or less PT)	Time point of outcome measurement	Primary outcome	Result of primary outcome
Batterham et al. (2018)	<i>n</i> total = 128 PT = 66 NPT = 62	Baseline • Symptom measures Posttreatment (2 weeks) • Symptom measures • Treatment adherence • Treatment satisfaction 3-month follow-up • Symptom measures Baseline	A composite of multiple symptom measures (see secondary outcomes), as an indicator of overall severity of mental health symptoms	No significant Time × Group interaction was found, $t(63.273) = -0.529, p = .599$.
Berger et al. (2014)	<i>n</i> total = 88	Baseline • Disorder-unspecific measures • Disorder-specific measures • SCID-I-interview After Treatment Session 4 Working alliance Posttreatment (8 weeks) • Disorder-unspecific measures • Disorder-specific measures • SCID-I-interview • Treatment satisfaction 6-month follow-up • Disorder-unspecific measures • Disorder-specific measures Baseline • Drinking behavior Posttreatment (12 weeks)	Disorder-unspecific measures: • Beck Anxiety Inventory (BAI) • Beck Depression Inventory (BDI) • Brief Symptom Inventory (BSI)	Posttreatment, no significant difference was found (all $ps > .26$; Cohen's $d = -0.05$ [BAI], 0.12 [BDI-II], and -0.02 [BSI]). <i>At 6-month follow-up, no significant Time × Group interaction was found from pretreatment to follow-up, $F(1, 58.6-80.9) = 0.03-2.06$, all $ps > .20$.</i>
Coates et al. (2018)	<i>n</i> total = 379 PT: <i>n</i> = 186 NPT: <i>n</i> = 193	• Drinking behavior Baseline and posttreatment (8 weeks) • Penn State Worry Questionnaire (PSWQ) • Symptoms of generalized anxiety disorder and anxiety • Depressive symptoms • Quality of life • Self-efficacy • Psychological flexibility • A knowledge test Baseline, posttreatment (5 weeks), 1-month follow-up, and 6-month follow-up • Cognitive-affective, behavioral, and physiologic outcomes	Drinking behavior over treatment: • Proportion of drinking days over treatment • Quantity of alcohol consumed between treatment sessions PSWQ	No significant effect of treatment condition was identified on the intercept or trajectory of the number of drinking days during treatment ($p = .162$) and on overall consumption over treatment or progression of consumption over treatment ($p = .364$). No significant effect of treatment condition, $F(1, 81) = 0.12, \eta_p^2 = 0.002$; support type, $F(1, 81) = 0.12, \eta_p^2 = 0.002$; or any interaction between treatment and support, $F(1, 81) = 0.88, \eta_p^2 = 0.010$, was found.
Dahlin et al. (2022)	<i>n</i> total = 85 PT + weekly support: <i>n</i> = 21 PT + support on demand: <i>n</i> = 22 NPT + weekly support: <i>n</i> = 21 NPT + support on demand: <i>n</i> = 21	• Drinking behavior Baseline and posttreatment (8 weeks) • Penn State Worry Questionnaire (PSWQ) • Symptoms of generalized anxiety disorder and anxiety • Depressive symptoms • Quality of life • Self-efficacy • Psychological flexibility • A knowledge test Baseline, posttreatment (5 weeks), 1-month follow-up, and 6-month follow-up • Cognitive-affective, behavioral, and physiologic outcomes	Drinking behavior over treatment: • Proportion of drinking days over treatment • Quantity of alcohol consumed between treatment sessions PSWQ	No significant effect of treatment condition was identified on the intercept or trajectory of the number of drinking days during treatment ($p = .162$) and on overall consumption over treatment or progression of consumption over treatment ($p = .364$). No significant effect of treatment condition, $F(1, 81) = 0.12, \eta_p^2 = 0.002$; support type, $F(1, 81) = 0.12, \eta_p^2 = 0.002$; or any interaction between treatment and support, $F(1, 81) = 0.88, \eta_p^2 = 0.010$, was found.
Dalton et al. (2004)	<i>n</i> total = 88 PT: <i>n</i> = 45 NPT: <i>n</i> = 43	• Brief Pain Inventory severity composite score	• Brief Pain Inventory severity composite score	Posttreatment, no significant difference was found. <i>At 1-month follow-up, no significant difference was found.</i>

(table continues)

Table 1 (continued)

First author, year of publication (reference)	Relevant sample size (PT vs. non- or less PT)	Time point of outcome measurement	Primary outcome	Result of primary outcome
Eiter (2009)	<i>n</i> total = 2,872	Posttreatment (48 hr)	Proportion of participants abstained from smoking in the past 24 hr:	At 6-month follow-up, no significant difference was found. No significant difference was found for proportion of participants abstained from smoking in the past 24 hr (personalized: 27.6% vs. nonpersonalized: 27.7%, $p = .99$), rates of quit attempts (12.1% in both groups, $p = 1.0$), and lapse or relapse rate (personalized: 25.1% vs. nonpersonalized: 23.5%, $p = .64$).
	PT: $n = 1,444$		• Did you smoke any tobacco in the past 24 hr?	
	NPT: $n = 1,428$		Quit attempts in current smokers (this outcome was based on the aforementioned question) Lapse and relapse in former smokers (this outcome was based on the aforementioned question)	
Ford-Gilboe et al. (2020)	<i>n</i> total = 531	Baseline, 3-month follow-up, 6-month follow-up, and 12-month follow-up	Depressive symptoms:	Depressive symptoms:
	PT: $n = 267$	• Depressive symptoms	• Center for Epidemiologic Studies Depression Scale-Revised	No significant Time × Group interaction was found ($p = .598$, $d = -0.18$).
	NPT: $n = 264$	• PTSD symptomology • Decisional conflict • Helpfulness of safety actions • Mastery Self-efficacy for safety planning • Social support • Experiences of coercive control Posttreatment • Decisional conflict	PTSD symptomology: • PTSD Checklist-Civilian Version	PTSD symptomatology: No significant Time × Group interaction was found ($p = .269$, $d = -0.17$).
Kerns et al. (2014)	<i>n</i> total = 128	Baseline and posttreatment (15 weeks)	Indices of engagement in treatment and adherence	Adherence ratings:
	PT: $n = 68$	• Pain severity	• Adherence/goal accomplishment (numeric rating scale)	No significant difference was found, $F(1, 110) > 1$, $\eta^2 = 0.001$.
	NPT: $n = 60$	• Physical functioning • Pain behaviors • Depressive symptom severity • Behavioral goal accomplishment During treatment • Adherence/goal accomplishment • Treatment credibility Posttreatment (15 weeks) • Adherence/goal accomplishment • Treatment “dose” • Treatment engagement • Adverse events • Patient satisfaction Baseline and posttreatment (6 weeks) • Yale-Brown Obsessive Compulsive Scale (Y-BOCS)	• Treatment “dose” (number of sessions attended) • Treatment engagement (number of sessions missed and rescheduled) • Adverse events	Average number of sessions attended: No significant difference was found, $F(1, 112) < 1$. Average number of canceled or rescheduled sessions: No significant difference was found ($F = 2.3$, $p > .10$). Adverse events: Comparison was not reported.
Moritz et al. (2016)	<i>n</i> total = 60 PT: $n = 28$		Y-BOCS OCI-R	Y-BOCS: No significant difference was found ($p > .3$).

(table continues)

Table 1 (continued)

First author, year of publication (reference)	Relevant sample size (PT vs. non- or less PT)	Time point of outcome measurement	Primary outcome	Result of primary outcome
Nissling et al. (2021)	NPT: $n = 32$	• Obsessive-Compulsive Inventory-Revised (OCI-R) • BDI		OCI-R: No significant difference was found ($p > .3$). A significant difference was found, $t(43) = 2.13$, $p = .04$, $n = 45$, favoring the personalized group.
	n total = 55	Baseline, during treatment (4 weeks), and 3-month follow-up	Perceived control over treatment	
	PT: $n = 27$ NPT: $n = 28$	• Making Decisions Scale • Generalized Anxiety Disorder seven-item scale • Montgomery Asberg Depression Rating scale • WHO Disability Assessment Schedule 12-item version Posttreatment (8 weeks) • Perceived control over treatment • Adherence to treatment • Making Decisions Scale • Generalized Anxiety Disorder seven-item scale • Montgomery Asberg Depression Rating scale • WHO Disability Assessment Schedule 12-item version		
Okajima et al. (2020)	n total = 47	Baseline, posttreatment (2 weeks), 1-month follow-up, and 3-month follow-up	Insomnia Severity Index	Insomnia severity:
	PT: $n = 24$	• Insomnia severity	Sheehan Disability Scale:	No significant difference between the two groups was reported.
	NPT: $n = 23$	• Social disabilities • Dysfunctional beliefs	• Work performance • Social life • Family life	Social disabilities: No significant difference between the two groups was reported.
Sauer-Zavala et al. (2022)	n total = 70	• Sleep reactivity • Work productivity Baseline, prior to second-stage randomization (i.e., between Sessions 5 and 6) and posttreatment (12 weeks)	The Diagnostic Interview for Anxiety, Mood, and Obsessive-Compulsive and Related Neuropsychiatric Disorders (DIAMOND) Overall Anxiety Severity and Impairment Scale (OASIS) Overall Depression Severity and Impairment Scale (ODSIS)	DIAMOND: No significant Time $\times$ Group interaction was found after first- and second-stage randomization, $F(2, 113) = 0.16$, $p = .85$. OASIS: No significant Time $\times$ Group interaction was found after first- and second-stage randomization, $F(2, 484) = 1.11$, $p = .3$. ODSIS: No significant Time $\times$ Group interaction was found after first- and second-stage randomization, $F(2, 481) = 0.15$, $p = .87$.
	PT compensation: $n = 23$	• DIAMOND		
	PT capitalization: $n = 21$ NPT: $n = 26$	• OASIS • ODSIS		
		During treatment (every week) • OASIS		
		• ODSIS Posttreatment (12 weeks) • Treatment acceptability • Treatment satisfaction		

(table continues)

Table 1 (*continued*)

First author, year of publication (reference)	Relevant sample size (PT vs. non- or less PT)	Time point of outcome measurement	Primary outcome	Result of primary outcome
Six et al. (2022)	<i>n</i> total = 94 PT: <i>n</i> = 45 NPT: <i>n</i> = 49	Baseline • PHQ-8 • Generalized Anxiety Disorder • Perceived Stress Scale <u>Posttreatment (2 weeks)</u> • PHQ-8 • Generalized Anxiety Disorder • Perceived Stress Scale • Adherence data (number of log-ins) <u>6-month follow-up</u>	Patient Health Questionnaire-8 (PHQ-8) Adherence data (number of log-ins)	PHQ-8: No significant Time × Condition effect was found, $F(1, 79) = 1.965, p = .17, \eta_p^2 = 0.024$. Adherence data: No significant difference was found ($p = .95$).
Streicher et al. (2008)	<i>n</i> total = 1,866	• 7-day point-prevalence abstinence	7-day point-prevalence abstinence. Abstinence was assessed by self-report during a telephone interview	A significant difference was found for the success stories ($p = .018$) and message source ($p = .039$), favoring the personalized condition. No other significant differences were found.
van der Aalst et al. (2012)	Success stories • PT: <i>n</i> = 488 • LPT: <i>n</i> = 456 Message source • PT: <i>n</i> = 481 • LPT: <i>n</i> = 463 Efficacy expectations • PT: <i>n</i> = 466 • LPT: <i>n</i> = 478 Outcome expectations PT: <i>n</i> = 494 LPT: <i>n</i> = 450 <i>n</i> total = 1,284 PT: <i>n</i> = 642 NPT: <i>n</i> = 642	2-year follow-up • Prolonged smoking abstinence • Quit attempts • Point-prevalent smoking abstinence • Continued smoking abstinence Baseline, <u>3-month follow-up</u> , and 6-month follow-up • 7-day point-prevalence abstinence	Prolonged smoking abstinence	No significant difference was found ($p = .11$).
Webb Hooper et al. (2013)	<i>n</i> total = 424 PT: <i>n</i> = 216 NPT: <i>n</i> = 208	• 28-day continuous abstinence • Readiness to quit • Quit attempts 3-month follow-up: • Intervention rating	7-day point-prevalence abstinence	A significant difference was found at 3-month follow-up ($p = .046$), favoring the personalized condition. <i>No significant difference was found at 6-month follow-up</i> ($p = .103$).

Note. The underlined text pertains to the primary time point of assessment. The italicized text indicates a result at a different time point. PT = personalized treatment; NPT = nonpersonalized treatment; SCID-I = Structured Clinical Interview for DSM-IV Axis I Disorders; PTSD = posttraumatic stress disorder; WHO = World Health Organization; LPT = less-personalized treatment; DSM-IV = *Diagnostic and Statistical Manual of Mental Disorders, fourth edition*.

over treatment. Participants reported a greater sense of control when different aspects of the treatment were personalized based on their preferences. In two other studies, including five assessments of the treatment process as the primary outcome and where personalized aspects of treatment were based on the participant's preference, no significant differences were found between the relevant treatment arms.

Efficacy on Primary Outcome(s) at Other Time Points

In Table 1, efficacy results per study for primary outcomes assessed at additional time points are also included. Three studies additionally assessed their primary outcome(s) at other time points than their primary time point. All three studies, including five primary outcomes, found no significant differences between treatment arms at these additional time points.

Efficacy on Secondary Outcome(s)

Three studies did not assess secondary outcomes, and 13 studies included 75 secondary outcomes. Six studies, including 30 secondary outcomes, found no difference between study arms on their secondary outcomes. Four studies found superiority of personalized psychological treatment for nine of the 27 outcomes. Two studies showed mixed results. One of these two studies included two secondary outcomes and three treatment arms and showed superiority of one of two personalized treatment arms and the non-personalized treatment compared with the other personalized treatment for one outcome. The other study included seven secondary outcomes and showed superiority of personalized treatment on one outcome at 1-month follow-up and superiority of non-personalized treatment on two outcomes at 6-month follow-up. Thus, for 11 of the total of 75 secondary outcomes, the personalized condition outperformed the non- or less personalized condition (15%). One study found superiority of nonpersonalized treatment for one of nine secondary outcomes. Further details on the efficacy of personalization on secondary outcomes can be found in Supplemental Material S4.

Risk of Bias

Of the three studies that found superiority of personalized treatment for their primary outcome at their primary time point (see meta-analysis), one study had an overall high risk of bias, and two had some concerns. Of the four studies that only found superiority of personalization on secondary outcomes, one study had a high risk of bias, two had some concerns, and one study had a low risk of bias.

Certainty of Evidence

Given that all studies were RCTs, the initial level of certainty was high. Subsequently, the level of evidence was downgraded due to the overall risk of bias (one level). Whereas there were some concerns about inconsistency of results and imprecision based on the narrative synthesis, we did not further rate down the quality of evidence because the results of the meta-analysis did not demonstrate a significant effect, even in studies with a large number of participants. Therefore, the overall body of evidence received a moderate quality rating, with no clear evidence for the superiority of

personalized psychological treatment compared to non- or less personalized treatment.

Discussion

The results of the meta-analysis showed small, nonsignificant differences between personalized and non- or less personalized treatment on outcomes measuring symptoms, smoking abstinence, and treatment outcomes. The effect of personalization was not significantly moderated by treatment format (in person vs. online), factor on which personalization was based (preference vs. individual profile), or treatment aspect that was personalized (treatment content vs. modules). Narrative synthesis of data showed that more than three quarters of the included studies (13 out of 16) did not find a significant difference between personalized treatment and non- or less personalized treatment on their primary outcome(s) at the predefined primary time point. Only three studies found superiority of personalization on three out of 32 primary outcomes at their predefined primary time point (9%). No studies found a significant difference on primary outcomes assessed at other time points, and most studies also did not find significant differences on their secondary outcomes. When a significant difference was found, this was mostly in favor of the personalized condition but not exclusively. In conclusion, this meta-analysis and systematic review does not show convincing evidence supporting the idea that personalization within treatment improves the outcome of psychological treatments compared to the same treatment in which no or less personalization is applied.

The discrepancies between our findings and those of the two previously conducted meta-analyses might be predominantly explained by differences in the assessment of the efficacy of personalization, the search strategy, and the targeted population, as only three of our included studies overlapped with the other reviews (Nye et al., 2023; Păsărelu et al., 2017). First, we only included studies that evaluated individual tailoring. In contrast, this only applied to five (out of 14) studies in the review of Păsărelu et al. (2017) and 10 (out of 17) studies from Nye et al. (2023). Second, we applied stricter criteria with regard to the personalization approach, requiring it to be clearly described, systematic, and replicable. For this review, this resulted in the exclusion of two studies (Johansson et al., 2012; Vernmark et al., 2010) included in at least one of the other reviews (Nye et al., 2023; Păsărelu et al., 2017). Of note, both studies reported superiority of the personalized condition. Third, we only included studies in which personalized psychological treatment was compared to the same or comparable treatment with no or less personalization, while the other two meta-analyses also included different kinds of comparators, such as a waitlist or an active control group with a different treatment (Nye et al., 2023; Păsărelu et al., 2017). For this review, this resulted in the exclusion of five studies (Carlbring et al., 2011; Lutz et al., 2022; Nordgren et al., 2014; Schulte et al., 1992; Silfvernagel et al., 2012), of which three reported superiority of the personalized condition (Carlbring et al., 2011; Nordgren et al., 2014; Silfvernagel et al., 2012). We identified 13 additional studies that were not included in the two meta-analyses, for which we have three possible explanations. First, we included the term "tailor" in our search strategy, enhancing our ability to find more relevant studies. Second, our inclusion criterion with regard to study population was broader. Third, our search covered a longer time span.

The lack of convincing evidence for the superiority of within-treatment personalization appears counterintuitive. Perhaps other forms of personalization, such as treatment matching or matching therapists to patients, might yield different insights. Additionally, it is important to consider the evidence base for personalization strategies. It is assumed that for personalization to be effective, the selected factor on which the personalization is based should have a demonstrated impact on treatment outcomes (Huibers et al., 2015). Also, the decision rules guiding personalization should be grounded in empirical evidence (Deisenhofer et al., 2024). In the 10 included studies that used a decision rule consisting of a (computerized) algorithm, the evidence base varied. Four studies used personalized information in their counseling letters, informed by theoretical models and based on findings from previous trials. Other studies used validated questionnaires with ($n = 2$) or without ($n = 4$) evidence-based cutoff values, and in three of the latter studies without evidence for linking these items to treatment modules. Personalization strategies based on research or validated algorithms are more likely to lead to meaningful improvements compared to those relying solely on clinical judgment or theoretical criteria.

There are other explanations for the lack of convincing evidence for the superiority of within-treatment personalization, for example, a limited contrast between personalized and non- or less personalized treatment arms. However, it is still unclear what degree of personalization is needed to achieve clinically relevant effects. Moreover, studies may suffer from insufficient power. When comparing two active treatments, one might anticipate smaller effect sizes, necessitating sufficiently large sample sizes to enable the detection of such small effects. In case only small effects are achieved, these gains could still be valuable at the population level. It is also possible that personalization offers less added value than one may expect. Nonspecific factors of therapy can serve as powerful mediators, for example, a strong therapeutic alliance facilitated by a skilled therapist (Baier et al., 2020). Consequently, an evidence-based psychological treatment based on group research may feel personalized to patients, even when following a standardized protocol.

Limitations of the evidence provided by this review include the exclusion of two eligible studies from our meta-analysis and narrative synthesis due to lack of information, despite our efforts to obtain information from the authors; the restriction of publications in English or Dutch; and data extraction by one reviewer and a check of the extracted data for each study by a second reviewer. Additionally, the control group of one study was included in two comparisons of our meta-analysis. The confidence intervals for these comparisons should have been slightly wider to appropriately account for the shared control group.

Strengths of the review process include the preregistration of the study protocol, a systematic search conducted across multiple databases, the use of backward and forward searching to identify potentially missing studies, contacting authors for additional information, the independent screening of full-text articles, assessment of risk of bias by two authors, independent data extraction for the meta-analysis by two reviewers, and the narrative synthesis involving multiple rounds of review and discussion.

In conclusion, there is a repeated call for the personalization of psychological treatment. However, the findings of this meta-analysis and narrative review do not provide convincing evidence that personalization within treatment consistently outperforms non-

or less personalized versions of the same or comparable treatment. For clinical practice, this suggests that while personalization within treatment can be applied, its added value over standard approaches remains uncertain. Future research should focus on identifying factors that predict or moderate treatment outcomes and use these insights to develop evidence-based personalization strategies to justify the resources invested in these approaches and to disentangle which personalization approaches work for whom and under what circumstances. Furthermore, future studies should aim to clarify which degree of personalization is needed to achieve clinically relevant effects.

References

- Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, 12(3), 154–174. <https://doi.org/10.20982/tqmp.12.3.p154>
- Baier, A. L., Kline, A. C., & Feeny, N. C. (2020). Therapeutic alliance as a mediator of change: A systematic review and evaluation of research. *Clinical Psychology Review*, 82, Article 101921. <https://doi.org/10.1016/j.cpr.2020.101921>
- Balaskas, A., Schueller, S. M., Cox, A. L., & Doherty, G. (2021). Ecological momentary interventions for mental health: A scoping review. *PLOS ONE*, 16(3), Article e0248152. <https://doi.org/10.1371/journal.pone.0248152>
- Batterham, P. J., Calear, A. L., Farrer, L., McCallum, S. M., & Cheng, V. W. S. (2018). FitMindKit: Randomised controlled trial of an automatically tailored online program for mood, anxiety, substance use and suicidality. *Internet Interventions*, 12, 91–99. <https://doi.org/10.1016/j.invent.2017.08.002>
- Berger, T., Boettcher, J., & Caspar, F. (2014). Internet-based guided self-help for several anxiety disorders: A randomized controlled trial comparing a tailored with a standardized disorder-specific approach. *Psychotherapy*, 51(2), 207–219. <https://doi.org/10.1037/a0032527>
- Bramer, W. M., Giustini, D., de Jonge, G. B., Holland, L., & Bekhuis, T. (2016). De-duplication of database search results for systematic reviews in endnote. *Journal of the Medical Library Association*, 104(3), 240–243. <https://doi.org/10.3163/1536-5050.104.3.014>
- Carlbring, P., Maurin, L., Törngren, C., Linna, E., Eriksson, T., Sparthar, E., Strååt, M., Marquez von Hage, C., Bergman-Nordgren, L., & Andersson, G. (2011). Individually-tailored, internet-based treatment for anxiety disorders: A randomized controlled trial. *Behaviour Research and Therapy*, 49(1), 18–24. <https://doi.org/10.1016/j.brat.2010.10.002>
- Coates, J. M., Gullo, M. J., Feeney, G. F. X., Young, R. M., & Connor, J. P. (2018). A randomized trial of personalized cognitive-behavior therapy for alcohol use disorder in a public health clinic. *Frontiers in Psychiatry*, 9, Article 297. <https://doi.org/10.3389/fpsy.2018.00297>
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Lawrence Erlbaum.
- Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. (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>
- Dahlin, M., Johansson, A., Romare, H., Carlbring, P., & Andersson, G. (2022). Worry-specific versus self-tailored internet-based treatments for generalized anxiety disorder, with scheduled support or support on demand: A pilot factorial design trial. *Internet Interventions*, 28, Article 100531. <https://doi.org/10.1016/j.invent.2022.100531>
- Dalton, J. A., Keefe, F. J., Carlson, J., & Youngblood, R. (2004). Tailoring cognitive-behavioral treatment for cancer pain. *Pain Management Nursing*, 5(1), 3–18. [https://doi.org/10.1016/S1524-9042\(03\)00027-4](https://doi.org/10.1016/S1524-9042(03)00027-4)
- Deisenhofer, A. K., Barkham, M., Beierl, E. T., Schwartz, B., Aafjes-van Doorn, K., Beevers, C. G., Berwian, I. M., Blackwell, S. E., Bockting,

- C. L., Brakemeier, E. L., Brown, G., Buckman, J. E. J., Castonguay, L. G., Cusack, C. E., Dalgleish, T., de Jong, K., Delgadillo, J., DeRubeis, R. J., Driessen, E., ... Cohen, Z. D. (2024). Implementing precision methods in personalizing psychological therapies: Barriers and possible ways forward. *Behaviour Research and Therapy*, 172, Article 104443. <https://doi.org/10.1016/j.brat.2023.104443>
- Etter, J.-F. (2009). Comparing computer-tailored, internet-based smoking cessation counseling reports with generic, untailored reports: A randomized trial. *Journal of Health Communication*, 14(7), 646–657. <https://doi.org/10.1080/10810730903204254>
- Ford-Gilboe, M., Varcoe, C., Scott-Storey, K., Perrin, N., Wuest, J., Wathen, C. N., Case, J., & Glass, N. (2020). Longitudinal impacts of an online safety and health intervention for women experiencing intimate partner violence: Randomized controlled trial. *BMC Public Health*, 20(1), Article 260. <https://doi.org/10.1186/s12889-020-8152-8>
- Harnas, S. J., Knoop, H., Sprangers, M. A. G., & Braamse, A. M. J. (2024). Defining and operationalizing personalized psychological treatment—A systematic literature review. *Cognitive Behaviour Therapy*, 53(5), 467–489. <https://doi.org/10.1080/16506073.2024.2333345>
- Huibers, M. J., Cohen, Z. D., Lemmens, L. H., Arntz, A., Peeters, F. P., Cuijpers, P., & DeRubeis, R. J. (2015). Predicting optimal outcomes in cognitive therapy or interpersonal psychotherapy for depressed individuals using the personalized advantage index approach. *PLOS ONE*, 10(11), Article e0140771. <https://doi.org/10.1371/journal.pone.0140771>
- Johansson, R., Sjöberg, E., Sjögren, M., Johnsson, E., Carlbring, P., Andersson, T., Rousseau, A., & Andersson, G. (2012). Tailored vs. standardized internet-based cognitive behavior therapy for depression and comorbid symptoms: A randomized controlled trial. *PLOS ONE*, 7(5), Article e36905. <https://doi.org/10.1371/journal.pone.0036905>
- Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, 63(3), 146–159. <https://doi.org/10.1037/0003-066X.63.3.146>
- Kerns, R. D., Burns, J. W., Shulman, M., Jensen, M. P., Nielson, W. R., Czapinski, R., Dallas, M. I., Chatkoff, D., Sellinger, J., Heapy, A., & Rosenberger, P. (2014). Can we improve cognitive-behavioral therapy for chronic back pain treatment engagement and adherence? A controlled trial of tailored versus standard therapy. *Health Psychology*, 33(9), 938–947. <https://doi.org/10.1037/a0034406>
- Leichsenring, F., Steinert, C., Rabung, S., & Ioannidis, J. P. A. (2022). The efficacy of psychotherapies and pharmacotherapies for mental disorders in adults: An umbrella review and meta-analytic evaluation of recent meta-analyses. *World Psychiatry*, 21(1), 133–145. <https://doi.org/10.1002/wps.20941>
- Liu, J., Gill, N. S., Teodorczuk, A., Li, Z. J., & Sun, J. (2019). The efficacy of cognitive behavioural therapy in somatoform disorders and medically unexplained physical symptoms: A meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, 245, 98–112. <https://doi.org/10.1016/j.jad.2018.10.114>
- Lutz, W., Deisenhofer, A. K., Rubel, J., Bennemann, B., Giesemann, J., Poster, K., & Schwartz, B. (2022). Prospective evaluation of a clinical decision support system in psychological therapy. *Journal of Consulting and Clinical Psychology*, 90(1), 90–106. <https://doi.org/10.1037/ccp0000642>
- Moritz, S., Stepulovs, O., Schröder, J., Hottenrott, B., Meyer, B., & Hauschildt, M. (2016). Is the whole less than the sum of its parts? Full versus individually adapted metacognitive self-help for obsessive-compulsive disorder: A randomized controlled trial. *Journal of Obsessive-Compulsive and Related Disorders*, 9, 107–115. <https://doi.org/10.1016/j.jocrd.2016.04.001>
- Ng, M. Y., & Weisz, J. R. (2016). Annual research review: Building a science of personalized intervention for youth mental health. *Journal of Child Psychology and Psychiatry*, 57(3), 216–236. <https://doi.org/10.1111/jcpp.12470>
- Nissling, L., Kraepelien, M., Kaldo, V., Hange, D., Larsson, A., Persson, M., & Weineland, S. (2021). Effects of patient-driven iCBT for anxiety in routine primary care and the relation between increased experience of empowerment and outcome: A randomized controlled trial. *Internet Interventions*, 26, Article 100456. <https://doi.org/10.1016/j.invent.2021.100456>
- Norcross, J. C., & Wampold, B. E. (2011). What works for whom: Tailoring psychotherapy to the person. *Journal of Clinical Psychology*, 67(2), 127–132. <https://doi.org/10.1002/jclp.20764>
- Nordgren, L. B., Hedman, E., Etienne, J., Bodin, J., Kadowaki, A., Eriksson, S., Lindkvist, E., Andersson, G., & Carlbring, P. (2014). Effectiveness and cost-effectiveness of individually tailored internet-delivered cognitive behavior therapy for anxiety disorders in a primary care population: A randomized controlled trial. *Behaviour Research and Therapy*, 59, 1–11. <https://doi.org/10.1016/j.brat.2014.05.007>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab2374>
- Nye, A., Delgadillo, J., & Barkham, M. (2023). Efficacy of personalized psychological interventions: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, 91(7), 389–397. <https://doi.org/10.1037/ccp0000820>
- Okajima, I., Akitomi, J., Kajiyama, I., Ishii, M., Murakami, H., & Yamaguchi, M. (2020). Effects of a tailored brief behavioral therapy application on insomnia severity and social disabilities among workers with insomnia in Japan: A randomized clinical trial. *JAMA Network Open*, 3(4), Article e202775. <https://doi.org/10.1001/jamanetworkopen.2020.2775>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), Article 210. <https://doi.org/10.1186/s13643-016-0384-4>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ*, 372(71), Article n71. <https://doi.org/10.1136/bmj.n71>
- Păsărelu, C. R., Andersson, G., Bergman Nordgren, L., & Dobrea, A. (2017). Internet-delivered transdiagnostic and tailored cognitive behavioral therapy for anxiety and depression: A systematic review and meta-analysis of randomized controlled trials. *Cognitive Behaviour Therapy*, 46(1), 1–28. <https://doi.org/10.1080/16506073.2016.1231219>
- Piccirillo, M. L., Beck, E. D., & Rodebaugh, T. L. (2019). A clinician's primer for idiographic research: Considerations and recommendations. *Behavior Therapy*, 50(5), 938–951. <https://doi.org/10.1016/j.beth.2019.02.002>
- Piccirillo, M. L., & Rodebaugh, T. L. (2019). Foundations of idiographic methods in psychology and applications for psychotherapy. *Clinical Psychology Review*, 71, 90–100. <https://doi.org/10.1016/j.cpr.2019.01.002>
- Sauer-Zavala, S., Southward, M. W., Stumpp, N. E., Semcho, S. A., Hood, C. O., Garlock, A., & Urs, A. (2022). A smart approach to personalized care: Preliminary data on how to select and sequence skills in trans-diagnostic CBT. *Cognitive Behaviour Therapy*, 51(6), 435–455. <https://doi.org/10.1080/16506073.2022.2053571>
- Schulte, D., Kunzel, R., Pepping, G., & Schulte-Bahrenberg, T. (1992). Tailor-made versus standardized therapy of phobic patients. *Advances in Behaviour Research and Therapy*, 14(2), 67–92. [https://doi.org/10.1016/0146-6402\(92\)90001-5](https://doi.org/10.1016/0146-6402(92)90001-5)
- Schünemann, H. J., Higgins, J. P. T., Vist, G. E., Glasziou, P., Akl, E. A., Skoetz, N., & Guyatt, G. H. (2023). Completing 'summary of findings'

- tables and grading the certainty of the evidence. In J. P. T. Higgins, T. J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 6.4, pp. 375–402). Cochrane.
- Segal, A., Pine, D. S., & Bar-Haim, Y. (2022). Personalized attention control therapy for PTSD: Effectiveness and moderators of outcome in a randomized controlled trial. *Psychological Medicine*, 52(12), 2365–2375. <https://doi.org/10.1017/s0033291720004304>
- Silfvermangel, K., Carlbring, P., Kabo, J., Edström, S., Eriksson, J., Månson, L., & Andersson, G. (2012). Individually tailored internet-based treatment for young adults and adults with panic attacks: Randomized controlled trial. *Journal of Medical Internet Research*, 14(3), Article e65. <https://doi.org/10.2196/jmir.1853>
- Six, S. G., Byrne, K. A., Aly, H., & Harris, M. W. (2022). The effect of mental health app customization on depressive symptoms in college students: Randomized controlled trial. *JMIR Mental Health*, 9(8), Article e39516. <https://doi.org/10.2196/39516>
- Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H. Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., ... Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *The BMJ*, 366, Article 14898. <https://doi.org/10.1136/bmj.14898>
- Stecher, V. J., Marcus, A., Bishop, K., Fleisher, L., Stengle, W., Levinson, A., Fairclough, D. L., Wolfe, P., Morra, M., Davis, S., Warnecke, R., Heimendinger, J., & Nowak, M. (2005). A randomized controlled trial of multiple tailored messages for smoking cessation among callers to the cancer information service. *Journal of Health Communication*, 10(1), 105–118. <https://doi.org/10.1080/10810730500263810>
- Stecher, V. J., McClure, J. B., Alexander, G. L., Chakraborty, B., Nair, V. N., Konkel, J. M., Greene, S. M., Collins, L. M., Carlier, C. C., Wiese, C. J., Little, R. J., Pomerleau, C. S., & Pomerleau, O. F. (2008). Web-based smoking-cessation programs: Results of a randomized trial. *American Journal of Preventive Medicine*, 34(5), 373–381. <https://doi.org/10.1016/j.amepre.2007.12.024>
- U.S. Department of Health and Human Services, National Institute of Occupational Health, National Institute of Mental Health. (2008). *National Institute of Mental Health strategic plan* (DHHS Publication No. 08-6368). <https://www.nimh.nih.gov/about/strategic-planning-reports/index.shtml>
- U.S. Department of Health and Human Services, National Institute of Occupational Health, National Institute of Mental Health. (2015). *National Institute of Mental Health strategic plan for research* (DHHS Publication No. 15-6368). <https://www.nimh.nih.gov/about/strategic-planning-reports/index.shtml>
- U.S. Department of Health and Human Services, National Institute of Occupational Health, National Institute of Mental Health. (2020). *NIMH strategic plan for research* (NIH Publication No. 20-MH-8096). <https://www.nimh.nih.gov/about/strategic-planning-reports/index.shtml>
- van Beugen, S., Ferwerda, M., Hoeve, D., Rovers, M. M., Spillekom-van Koulil, S., van Middendorp, H., & Evers, A. W. (2014). Internet-based cognitive behavioral therapy for patients with chronic somatic conditions: A meta-analytic review. *Journal of Medical Internet Research*, 16(3), Article e88. <https://doi.org/10.2196/jmir.2777>
- Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, 45(2), 576–594. <https://doi.org/10.3758/s13428-012-0261-6>
- van der Aalst, C. M., de Koning, H. J., van den Bergh, K. A. M., Willemsen, M. C., & van Klaveren, R. J. (2012). The effectiveness of a computer-tailored smoking cessation intervention for participants in lung cancer screening: A randomised controlled trial. *Lung Cancer*, 76(2), 204–210. <https://doi.org/10.1016/j.lungcan.2011.10.006>
- van Dessel, N., den Boeft, M., van der Wouden, J. C., Kleinstäuber, M., Leone, S. S., Terluin, B., Numans, M. E., van der Horst, H. E., & van Marwijk, H. (2014). Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults. *Cochrane Database of Systematic Reviews*, 2014(11), Article CD011142. <https://doi.org/10.1002/14651858.CD011142.pub2>
- Vernmark, K., Lenndin, J., Bjärehed, J., Carlsson, M., Karlsson, J., Oberg, J., Carlbring, P., Eriksson, T., & Andersson, G. (2010). Internet administered guided self-help versus individualized e-mail therapy: A randomized trial of two versions of CBT for major depression. *Behaviour Research and Therapy*, 48(5), 368–376. <https://doi.org/10.1016/j.brat.2010.01.005>
- Viechtbauer, W. (2015). *Meta-analysis package for R* [Computer software]. <https://cran.r-project.org/web/packages/metafor/metafor.pdf>
- Webb Hooper, M., Rodríguez de Ybarra, D., & Baker, E. A. (2013). The effect of placebo tailoring on smoking cessation: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 81(5), 800–809. <https://doi.org/10.1037/a0032469>

Received June 14, 2024

Revision received May 9, 2025

Accepted May 13, 2025 ■