Scandinavian Journal of Psychology, 2024, 65, 311–320 DOI: 10.1111/sjop.12976

Review Article

Clinically representative therapy for Nordic adult outpatients with common mental health problems: A systematic review and meta-analysis

MARTIN BRATTMYR,[1] MARTIN SCHEVIK LINDBERG,[1,2] JAKOB LUNDQVIST,[1] LARS-GÖRAN ÖST,[3] STIAN SOLEM,[1] ODIN HJEMDAL[1] and AUDUN HAVNEN[1,4]

[1]*Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway*
[2]*Mental Health Care Services, Trondheim Municipality, Trondheim, Norway*
[3]*Department of Psychology, Stockholm University, Stockholm, Sweden*
[4]*Nidaros Community Mental Health Centre, Division of Psychiatry, St. Olavs University Hospital, Trondheim, Norway*

Brattmyr, M., Lindberg, M.S., Lundqvist, J., Öst, L.-G., Solem, S., Hjemdal, O. & Havnen, A. (2024). Clinically representative therapy for Nordic adult outpatients with common mental health problems: A systematic review and meta-analysis. *Scandinavian Journal of Psychology*, *65*, 311–320.

There is a knowledge gap regarding clinically representative therapy given in routine settings, that is treatment as usual (TAU), for patients with common mental health problems (CMHP). This review and meta-analysis aimed to investigate what characterizes clinically representative therapy in Nordic routine clinics and meta-analyze the outcome of such treatment. Databases (PubMed, EMBASE, PsychINFO, and SveMed+) were searched for TAU, CMHP, and Nordic countries, together with backward and forward search in Scopus (7 November 2022). Studies were either randomized controlled trials (RCT) or open trials, using prospective study designs, examining heterogeneous outpatient groups in routine treatment. Within- and between-group effect sizes (ES), using random effects model, and moderator analyses were calculated. Eleven studies (n = 1,413), demonstrated a small to moderate within-group ES with high heterogeneity (g = 0.49, I^2 = 90%). ESs in RCTs were significantly smaller than in open trials. TAU had a marginally smaller ES (g = −0.21; adjusted for publication bias g = −0.06) compared to a broad set of clinical interventions. Clinically representative therapy in the Nordic countries demonstrated a wide variety of characteristics and also a marginally lower ES compared to other interventions. The ESs were smaller than other meta-analyses examining evidence-based treatments in routine treatment.

Key words: Clinically representative treatment, treatment as usual (TAU), care as usual, common mental health problems, meta-analysis.

Martin Brattmyr, Norwegian University of Science and Technology NTNU, Department of Psychology, NO-7491, Trondheim, Norway. Email: martin.brattmyr@ntnu.no

INTRODUCTION

Research on the treatment outcome of routine care for patients with common mental health problems (CMHP), such as depressive or anxiety disorders (National Institute for Health and Care Excellence [NICE], 2011), has been neglected (Hewlett & Moran, 2014). Routine mental health treatment may be equated to treatment as usual (TAU). Rather than viewing TAU as a generic control condition to evidence-based treatments (EBT), the effect of TAU may also be used as an indicator of what outcome to expect for patients who undergo treatment in mental health care facilities (Kazdin, 2015). Meanwhile, research on TAU is challenging, mainly due to its ambiguity (Freedland, Mohr, Davidson & Schwartz, 2011; Kazdin, 2015; Wampold, Budge, Laska *et al.*, 2011; Watts, Turnell, Kladnitski, Newby & Andrews, 2015). One of the biggest limitations of TAU research is that many researchers do not describe the contents of it (Wampold *et al.*, 2011; Watts *et al.*, 2015). Therefore, TAU is often a *laissez faire* (Freedland *et al.*, 2011) and in many randomized control trials (RCTs) not even intended to be therapeutic (Wampold *et al.*, 2011). As a result, weaker TAU leads to greater effect sizes (ES) for the intervention of interest in direct comparisons (Cuijpers, Karyotaki, Reijnders & Ebert, 2019; Cuijpers, Quero, Papola, Cristea & Karyotaki, 2021; Wampold *et al.*, 2011; Watts *et al.*, 2015).

A second challenge is the use of TAU across different research contexts. Patients randomized to TAU can expect inferior treatment or receive less attention than the intervention group (Freedland *et al.*, 2011). A third challenge comes with the variation in the availability and utilization of outpatient services across the world (World Health Organization [WHO], 2021). This in turn may affect the ES of TAU, for example between the US compared to Scandinavia (Löfholm, Brännström, Olsson & Hansson, 2013). Thus, Cuijpers *et al.* (2021) recommended meta-analyses to recruit studies within one country and one setting, although with the number of expected studies and resources in consideration.

Overall, there is substantial evidence of the effectiveness of psychological and psychopharmacological treatment for patients over a wide range of disorders (Leichsenring, Steinert, Rabung & Ioannidis, 2022), and EBT in routine care (Wakefield, Kellett, Simmonds-Buckley, Stockton, Bradbury & Delgadillo, 2021), but few studies have examined TAU restricted to so-called clinically representative therapy. In a review of meta-analyses, three increasingly stringent and cumulative criteria defined clinically representative treatment (Shadish, Matt, Navarro *et al.*, 1997). First, patients had to be referred in a conventional manner into routine clinics, with regular therapists having regular caseloads. Second, treatments had to be unaltered by the researchers. Finally, patients had to have a spread of mental health problems and background characteristics, while therapists were free to use a variety of techniques, and not trained for the specific study. With over 500 studies examined, only one fulfilled all these criteria

(Shadish *et al.*, 1997). An updated review found more studies (Shadish, Matt, Navarro & Phillips, 2000). The random effect size was $d = 0.41$ and effects increased with larger treatment dose and use of specific outcome measures. However, no statistically significant association between the degree of clinical representativeness and ES was found.

Others have reviewed TAU for depression and anxiety but restricted to be a control condition to cognitive behavior therapy (CBT; Watts *et al.*, 2015), EBT (Wampold *et al.*, 2011) or guideline-adherent interventions (Setkowski, Boogert, Hoogendoorn, Gilissen & van Balkom, 2021). In studies where TAU was a psychotherapy intervention for depression and anxiety disorders, five studies had a between-group ES for depression of $g = 0.44$ and for anxiety of $g = 0.34$ in favor of CBT (Watts *et al.*, 2015). A more stringent criterion of TAU as an active treatment revealed three studies with a between-group ES of $d = 0.33$ in favor of EBT (Wampold *et al.*, 2011). However, the heterogeneity for TAU in general was substantial in both these meta-analyses. Guideline-adherent therapies have also shown larger ES than TAU across nine different diagnosis-specific studies, with a between-group ES of $d = 0.29$ (Setkowski *et al.*, 2021). Further, an umbrella review of 102 meta-analyses compared psychotherapy with active TAU, with an ES of $d = 0.36$ in favor of psychotherapy (Leichsenring *et al.*, 2022). Within-group analysis of routine treatment for patients with depression and anxiety, treated with EBT in the Improving Access to Psychological Therapies (IAPT) program, had an ES of $d = 0.87$ and 0.88 for symptoms of depression and anxiety (Wakefield *et al.*, 2021).

To increase knowledge about the treatment that most people with CMHP receive, an updated systematic review of clinically representative therapies is necessary. In this review the Nordic countries were selected, due to their many similarities in population characteristics and health care utilization. The Nordic region is often considered a distinct region in international comparisons of health care systems, characterized by high-trust, high-taxation, open economies (Lyttkens, Christiansen, Häkkinen, Kaarboe, Sutton & Welander, 2016). Despite a policy for evidence-based treatments across the Nordic countries, there is reason to believe that there is a great deal of non-adherence to these by therapists (Bergmark, Sundberg, Markström & Rosenberg, 2022). Although there is a lot of evidence regarding the effectiveness of evidence-based treatment, there appears to be a knowledge gap regarding the effect of the treatment provided in routine mental health care. A systematic review and meta-analysis on TAU in the Nordic region may serve as a benchmark of the effect of treatment that most patients receive in routine mental health care facilities, while at the same time unveiling the characteristics of TAU in these countries.

In this review we aimed to examine: (1) the contents of clinically representative mental health outpatient treatment for adults with CMHP in the Nordic countries; and (2) provide a meta-analysis regarding its treatment effects.

METHOD

In the present review, TAU was defined in accordance with the strictest definition of clinically representative therapy (Shadish *et al.*, 1997). Treatment had to be active, unaltered by the researchers, and conducted in publicly available outpatient facilities where patients are referred through usual clinical routes. Thus, research done in private care and university clinics were excluded. Common mental health problems (CMHP) were defined according to, but not limited to, the clinical guidelines by NICE (2011), which use the term for depressive and anxiety disorders. Unlike common mental disorders (Hewlett & Moran, 2014), CMHP was conceptualized to include subclinical populations. Most patients with CMHP are treated within primary health care (NICE, 2011), but CMHP are also the most common disorders in heterogeneous outpatient treatment in secondary care. Thus, heterogeneous outpatient facilities were defined as publicly available primary or secondary care facilities aimed to treat CMHP, including depressive and anxiety disorders as well as related mental health problems.

Patients had to display heterogeneity regarding demography and mental health problems. Records examining only specific disorder groups, or for example, severe mental health disorders (defined by the authors themselves), or only suicidal patients were excluded. Therapists could also not rely on a specific technique, thus records where for example, CBT was termed TAU and was the only treatment intervention, were excluded. Both RCTs and open trials using prospective study designs were included.

Search strategy

A systematic search in the electronic databases PubMed, EMBASE, PsycINFO, and SveMed+ was conducted using no time-limit, which is further described in a preregistered protocol on PROSPERO (CRD42020213988). The search string is presented in Appendix S1. First, a search strategy for PubMed was designed. It was then adapted to the other databases regarding syntax and search field tags. The search strategy included variations and synonyms of mental health problems, "routine outpatient treatment" and outcome measurements, together with the Nordic countries: Denmark, Finland, Iceland, Norway, and Sweden.

Terms were combined with Boolean operators (OR/AND) along with truncation. Nordic countries were searched in all fields, other terms in title/abstract and corresponding index terms. Duplicates were discarded first in Endnote and then in the web tool Rayyan (Ouzzani, Hammady, Fedorowicz & Elmagarmid, 2016). Reference lists and citing reports of the records read for full text were identified using scopus.com, thus backwards and forward citation searching was conducted. Three researchers (MB, ML & JL), independently and blinded from each other, screened titles and abstracts for eligibility according to inclusion and exclusion criteria. Conflicting results were discussed to reach consensus. Records eligible for full text reading, but later excluded were documented, and the reason for the first discovered exclusion criterion was plotted. The following criteria for exclusion were used: not Nordic, no heterogeneous adult sample, not outpatient facility, not TAU, and not prospective study. Records using the same dataset as one already excluded or included were marked as secondary data if no new information of interest were presented. A data extraction table was designed, piloted, and used to extract data. For this review, PRISMA guidelines were followed (see Appendix S2).

Quality assessment

The Downs and Black (1998) checklist for assessment of the methodological quality of both randomized and non-randomized studies of health care interventions was applied. The checklist uses subscales regarding reporting, external validity, bias of the measurement/outcome, confounding selection of study subjects and power. The instrument was modified (equal to Bear, Edbrooke-Childs, Norton, Krause & Wolpert, 2020; questions regarding the intervention group were not used, that is, items 14 and 21–24), thus yielding a maximum score of 23, where higher scores indicated better quality. Correspondingly, cut-offs (as suggested by Hooper, Jutai, Strong & Russell-Minda, 2008) were adjusted, where less than 10 meant the study demonstrated poor, 10–14 fair, 15–20 good, and 21–23 excellent methodological quality. For the power analysis item, the research group determined a maximum score of 1 for sufficient power, and power for all studies was manually calculated. Two researchers

conducted the assessment independently and blinded, and intraclass correlation was calculated. Conflicting results were discussed in the research team until full consensus was reached. The scores were used to analyze the quality regarding external and internal validity, but also using an overall score in meta-regression analysis. The interrater reliability for the quality scale resulted in a moderate level of agreement (intraclass correlation coefficient = 0.74, 95% confidence interval [CI] = 0.66–0.79, $p < 0.001$).

Statistical analyses

IBM SPSS Statistics (version 27) and Comprehensive Meta-Analysis (version 3.3.070; Borenstein, Hedges, Higgins & Rothstein, 2014) were used for statistical analysis. Both between-group and within-group meta-analyses were conducted with random effects model using Hedges' *g*, presented by forest plot. Generic patient reported outcome measures or subscales measuring broader change in symptoms were used as primary outcome measures. Well established global psychotherapeutic outcome measures (e.g., as reviewed by Tarescavage & Ben-Porath, 2014) were preferred. If more than one measure of interest was reported, the most suitable measure was decided in the research group.

For studies reporting several points of assessment after the post-treatment assessment, the final point of assessment was used as follow-up. The intervention was assumed to affect the post-measure and standard deviation, thus as recommended by Lakens (2013), the following formula for ES was used: (*M*pre − *M*post)/*SD*pre. Due to unknown pre-post-treatment correlations, this was imputed at 0.5 (Follmann, Elliott, Suh & Cutler, 1992). For studies presenting results for completers only analysis, the post sample size was used. Heterogeneity was estimated with the *Q*-value and I^2 (Higgins, Thomas, Chandler *et al*., 2021). Risk of publication bias was analyzed by inspection of Egger's regression intercept (Egger, Davey Smith, Schneider & Minder, 1997), and by Duval and Tweedie's (2000) trim-and-fill method.

The following variables were extracted for qualitative synthesis and to examine sources of heterogeneity: main intervention, study design (RCT or open trial), generic outcome measure, other outcome measures, country, level of care (primary or secondary care), therapist profession, format of therapy, type of psychological intervention, percent declined, percent excluded before study started, non-starters, attrition to follow-up, quality of the studies, data-collection years, number of patients at start of treatment, percent female, mean age, working status, civil status, non-nativity, education, diagnosis, session mean, mean duration, percent pharmacotherapy, and weeks between pre and post measurement. To examine sources of heterogeneity, subgroup analysis was used for categorical variables, and meta-regression for continuous moderators. The Cochrane handbook recommends at least 10 studies as the lowest number for conducting subgroup analysis or meta-regression (Higgins *et al*., 2021). Thus, subgroup-analysis and meta-regression were applied if variables from at least 10 studies provided data on the variable in question.

RESULTS

Study selection

Four database searches (October 14, 2020; November 30, 2021; March 8, 2022, and November 7, 2022) and four backwards and forward searches were conducted (December 7, 2020; January 3, 2022; March 23, 2022, and November 7, 2022; see Fig. 1). A total of 7202 records were screened. Out of these, 119 reports were read in full text, and 12 studies were accepted (see Appendix S3 for full texts excluded with reason). Out of these, four articles provided follow-up data (Arvidsdotter, Marklund & Taft, 2014; Bratberg, Leira, Granan *et al*., 2021; Koksvik, Linaker, Gråwe, Bjørngaard & Lara-Cabrera, 2018; Rise, Eriksen, Grimstad & Steinsbekk, 2016). Since most articles had results by completers only (except Arvidsdotter *et al*., 2014, who provided intention to treat results), per protocol was used.

Eleven studies provided a generic outcome measure (see Appendix S4). For two studies, the 32-item behavior and symptom identification scale (BASIS-32) was chosen as it was more comprehensive than the four-item outcome rating scale (ORS). Three studies provided disorder-specific outcome measures, four studies provided measures of self-reported health, and three studies provided outcomes measuring social functioning. For meta-analytic data-synthesis, only generic outcome measures were chosen, due to too few studies providing other outcomes of interest.

Quality assessment

The included studies demonstrated a great variation in methodological quality measured by a modified version of Downs and Black (1998) (see Appendix S5). The mean quality score was 11.5 (*SD* = 3.3) out of 23 (range 8–17). Four studies demonstrated poor quality, six fair, and two good quality, while none demonstrated excellent quality. None of the included studies described adverse effects of the intervention, two described sufficiently the patient characteristics of attrition, and two controlled for it in the analyses.

Study characteristics

Design and attrition. Seven studies were RCTs and five were open trials. As a generic outcome measure, three studies used the 90-item Symptom Checklist (SCL-90), three BASIS-32, and five used others (see Appendix S4). One study did not present outcome using a generic symptom measure, and was excluded for quantitative synthesis (Ramirez, Ekselius & Ramklint, 2008). Three studies did not present clear inclusion or exclusion criteria except reasons why patients declined (Østergård, O'Toole, Svendsen & Hougaard, 2020;, Rise *et al*., 2016; Werbart, Levin, Andersson & Sandell, 2013). Number of declined, excluded, non-starting patients, and drop-outs (attrition) spanned considerably (see Appendix S4). However, many studies did not specify non-starters, thus attrition statistics could have been negatively affected.

Facility and patient characteristics. Six studies were conducted in Sweden, five in Norway, and one in Denmark (see Table 1). No studies conducted in Finland or Iceland were included. Four studies examined treatment in primary care, seven examined specialist mental health services or secondary care, and one study examined both primary care and specialist mental health services (correspondingly 30% and 70% of the patients; Werbart *et al*., 2013). All studies had a majority of female patients (average 72%, range 55%–90%), with mean age of 31.7 years (range 22–42 mean years).

Diagnoses were reported in eight studies, three using DSM-IV criteria (Hansson, Rundberg, Österling, Öjehagen & Berglund, 2013; Ramirez *et al*., 2008; Werbart *et al*., 2013), three ICD-10 criteria (Bratberg *et al*., 2021; Brattland, Koksvik, Burkeland *et al*., 2018; Møllersen, Sexton & Holte, 2009) and

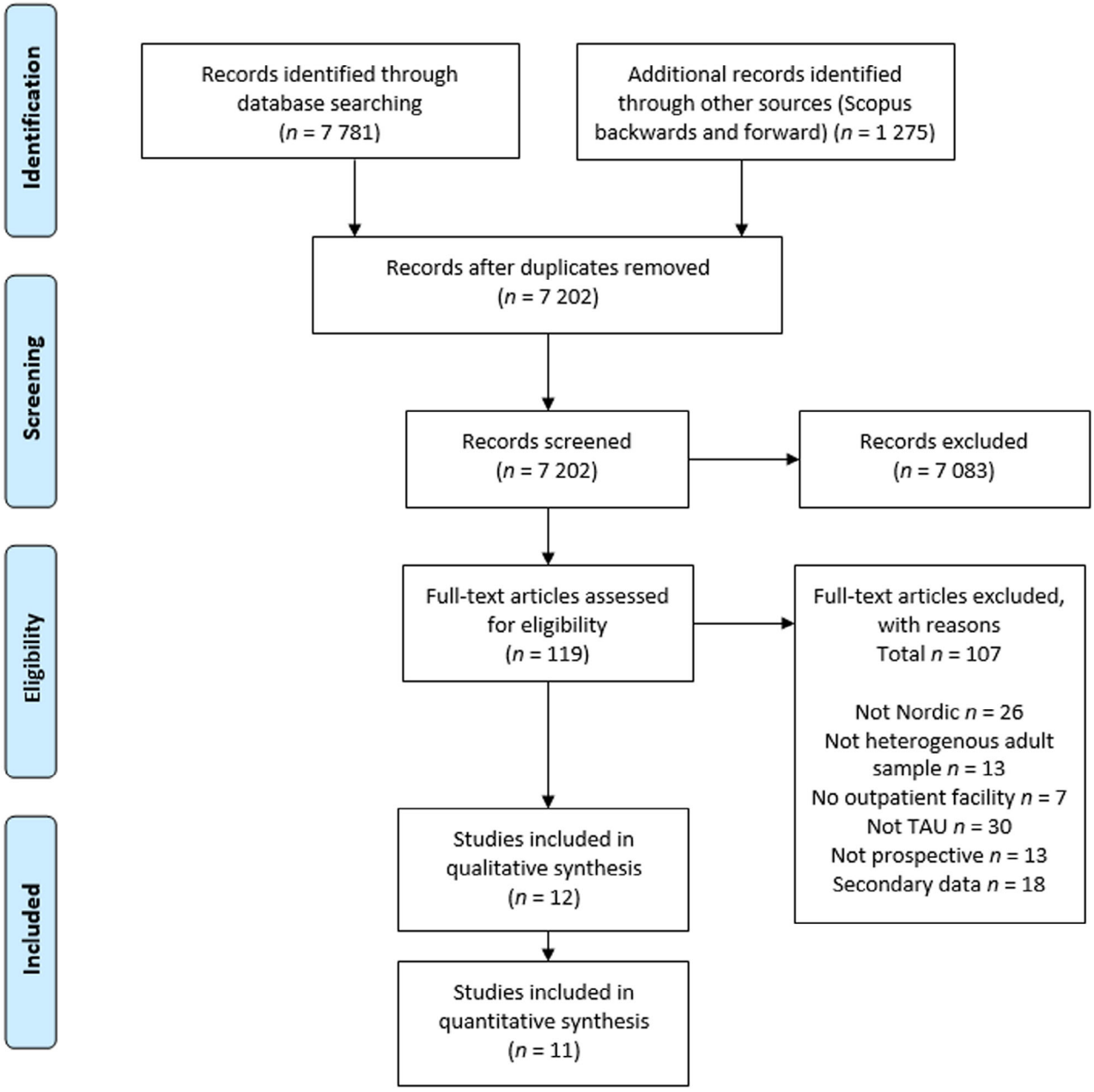


Fig. 1. PRISMA flow chart of included studies.

two without a specified diagnostic system (Arvidsdotter *et al*., 2014; Østergård *et al*., 2020). Patients with depressive, or anxiety disorders/mental health problems were reported in eight studies (depression range 28%–73%, anxiety range 20%–67%), with an estimated total proportion of diagnoses of 43% depression and 34% anxiety disorders.

Therapists and treatment characteristics. The most commonly reported professions at the primary health care level were psychologists and social workers (see Table 2). At the secondary level, psychiatrists and psychologists were most frequently reported, but also social workers, psychiatric nurses, nursing assistants, physiotherapists, occupational therapists, and milieu therapists were reported. Four studies did not report health care professionals' background.

Type of treatment also varied: two studies reported psychotherapy as the only treatment given, seven studies reported psychotherapy and psychopharmacotherapies used in combination. Three studies did not explicitly report the format of therapy (Koksvik *et al*., 2018; Ramirez *et al*., 2008; Rise *et al*., 2016). All treatments were given face-to-face individually or in group. The most common treatment interventions were CBT, but also metacognitive-, psychodynamic-, support-, psychoeducational-, systemic-, humanistic-, and existential-therapies were reported. Seven studies did not report if specific treatment interventions were given.

Six studies presented mean number of sessions per patient (M = 7.13, range 4–27 sessions), and one study presented range, where 52% had 1–10 visits (Hansson *et al*., 2013). Four articles presented mean duration (M = 100 days, range 56 days–

Table 1. *Facility and patient characteristics*

Study	Data collection	Country	Level of care	TAU *n*	Female %	Age *M* (*SD*)	Working status	Civil status	Disorders/problems
Arvidsdotter, Marklund, and Taft (2014)	2010–2011	Sweden	Primary	40	88	40 (9.1)	Unknown	Unknown	Dep 30% Anx 20%
Bratberg *et al.* (2021)	2014–2016	Norway	Secondary	35	64	31.7 (10.7)	56% employed	Unknown	Dep 28% Anx 59%
Brattland *et al.* (2018)	2012–2016	Norway	Secondary	85	60	34.6 (12)	48% not working	51% single	Dep 28% Anx 28%
Elfström *et al.* (2013)	Unknown	Sweden	Primary	133	71	42.4 (12.9)	Unknown	Unknown	Unknown
Hansson *et al.* (2013)	2007–2008	Sweden	Secondary	186	69	39 (14.1)	39% unemployed	66% single	Dep 33% Anx 24%
Koksvik *et al.* (2018)	2009–2013	Norway	Secondary	40	76	37.09 (12.8)	23.9% employed	40% married	Unknown
Møllersen *et al.* (2009)	1999–2001	Norway	Secondary	335	61	36.01 (11.9)	60% employed	Unknown	Dep 43% Anx 31%
Østergård *et al.* (2020)	2014–2016	Denmark	Primary	740	75	25.19 (4.6)	Unknown	Unknown	Dep 43% Anx 30%
Ramírez *et al.* (2008)	2002–2004	Sweden	Secondary	191	80	22.4 (1.9)	Unknown	Unknown	Dep 73%[b] Anx 67%
Rise *et al.* (2016)	2010– Unknown	Norway	Secondary	38	55	29.2 (unknown)	15.8% working	37% living alone	Unknown
Sundquist *et al.* (2017)	2012	Sweden	Primary	105	90	41 (11)	Unknown	65% married[a]	Unknown
Werbart *et al.* (2013)	2007–2010	Sweden	Primary & Secondary	180	74	36 (10.9)	Unknown	Unknown	Dep 42% Anx 24%

[a]Information provided in Sundquist *et al.*, 2019,
[b]Depressive disorders, dysthymia and bipolar disorder.

7 months), and three presented frequencies, between weekly to monthly sessions. Three studies provided mean number of sessions for both the primary intervention and TAU: one study reported more sessions provided to the intervention (5%; Bratberg *et al.*, 2021), and two studies reported more sessions provided to TAU (7% in Brattland *et al.*, 2018; 21% for group treatment, 5% for individual treatment in Østergård *et al.*, 2020). One study stated no difference in number of sessions (Hansson *et al.*, 2013), and one provided the mean of the intervention together with TAU (Rise *et al.*, 2016).

Five studies reported use of medication, of which two explicitly reported pharmacological drug of relevance for mental health condition (33% in Bratberg *et al.*, 2021; 47.4% in Rise *et al.*, 2016). One study reported drug categories (antidepressant 35%, anxiolytics 16% in Sundquist, Palmér, Johansson & Sundquist, 2017). Two studies presented percentage of patients with pharmacological drug use, without explicitly reporting if it was of relevance for their mental health condition (34.9% in Møllersen *et al.*, 2009; 41% in Werbart *et al.*, 2013).

Results of synthesis

Power analysis. For the within-group meta-analysis, 12 ESs were calculated with a mean number of treated participants of 117.8 (n = 1,413, SD = 126.6, range 28–480), and for the between-group analysis there were nine ESs with a mean number of treated participants of 103.8 (intervention n = 937, SD = 149.6, range 22–492, TAU n = 932, SD = 144.8, range 28–480). According to the formulas for power analysis in meta-analyses by Valentine, Pigott, and Rothstein (2010), there would be a 75.6% power for within-group and 57.8% power for between-group to detect a small ES (0.20), when assuming that the heterogeneity of ESs was high.

Meta analysis. Random effects model for overall within-group resulted in a significant small to moderate ES (g = 0.49, 95% CI = 0.30–0.68, $p < 0.001$; forest plot is displayed in Fig. 2). However, the heterogeneity between the studies was substantial, with a Q-value of 107.4, $df(11)$, $p < 0.001$, and I^2 = 90%. For the studies that reported follow-up data (n = 4), the follow-up ES (g = 0.64, 95% CI = 0.44–0.84, $p < 0.001$) was non-significantly different from the post ES (Q = 1.16, p = 0.281). For these, a non-significant test of heterogeneity was found, Q = 0.49, $df(3)$, p = 0.920, I^2 < 1% (see Appendix S6). Nine ESs were extracted for TAU compared to an intervention (see Fig. 2)

The random effects model for between-group resulted in a significant small ES in favor of the interventions (g = −0.21, 95% CI = −0.36 to −0.05, p = 0.010). The Q-value, Q = 15.90, $df(8)$, p = 0.044, was significant and the I^2 showed moderate heterogeneity (50%). For the few studies that reported follow-up data (n = 4), a non-significantly different ES was obtained (Q = 0.08, p = 0.782; g = −0.25, 95% CI = −0.52 to −0.02, p = 0.066), with smaller heterogeneity (Q = 3.21, $df(3)$, p = 0.361, I^2 = 6%; see Appendix S6).

Publication bias. Publication bias did not seem to be a critical issue regarding within-group analysis, with a non-significant Egger's regression intercept (t = −1.21, p = 0.253), and Duval

Table 2. *Therapists and treatment characteristics*

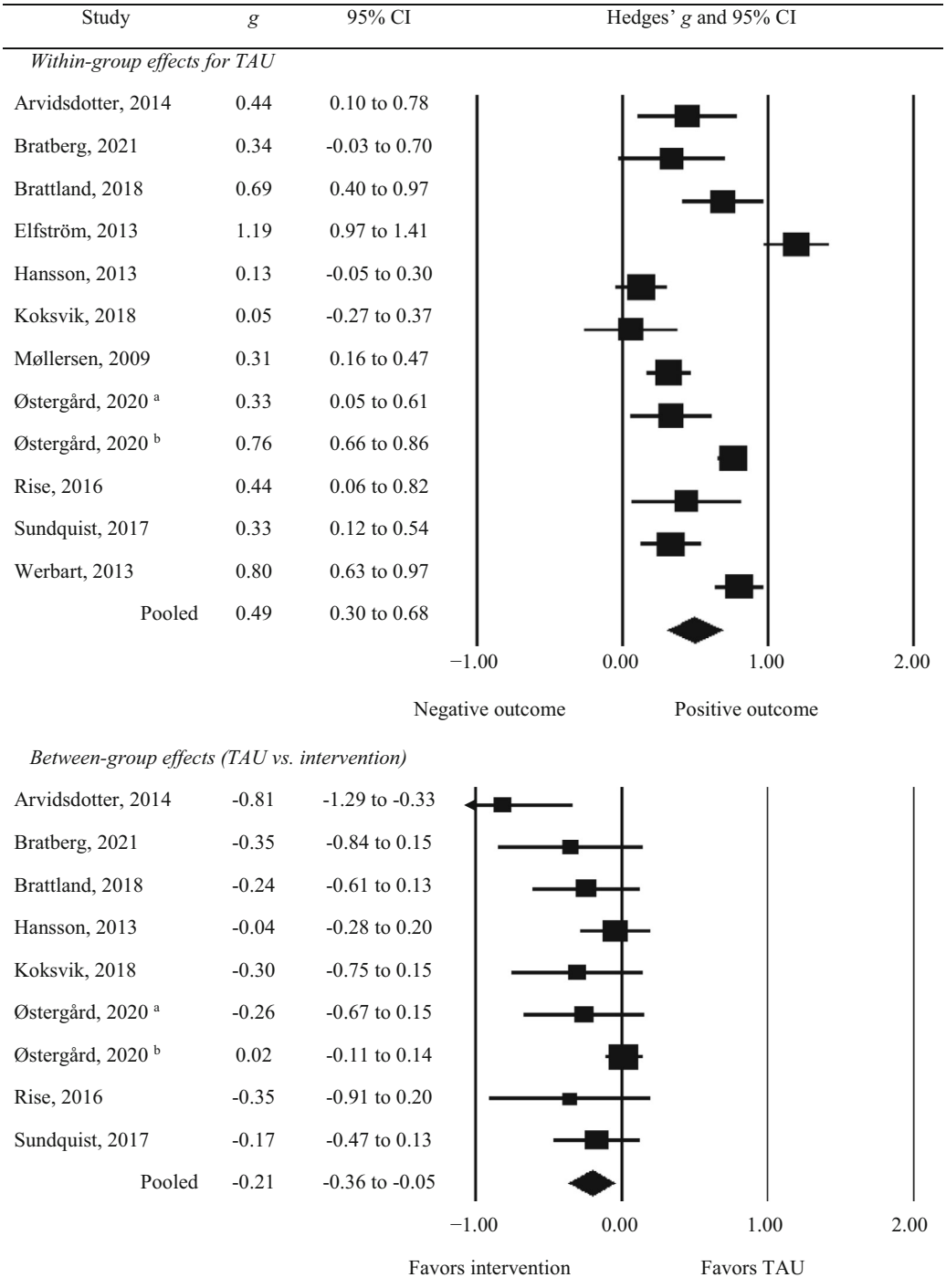
Study	Therapists profession (*n*)	Format of therapy	Psychological Intervention	Sessions *M* (*SD*)	Mean duration (Frequency)	Medication
Arvidsdotter *et al.* (2014)	Unknown (Unknown)	Psychological & pharmacological	Unknown	Unknown	Unknown	Unknown
Bratberg *et al.* (2021)	Psychologists, psychiatrists, psychiatric nurses, milieu therapists and social working clinicians (Unknown)	Psychological & pharmacological	CBT, MCT, PDT, and/or support therapy	13.8 (Unknown)	7 months (Unknown)	33%
Brattland *et al.* (2018)	Psychologists (11), psychiatrists (6), and other mental health care professions (3) (Total 20)	Psychological	PDT, humanistic/ existential, CBT	13.01 (10.92)	(Weekly/Biweekly)	Unknown
Elfström *et al.* (2013)	Psychologists (Unknown)	Psychological	Unknown	Unknown	Unknown	Unknown
Hansson *et al.* (2013)	Psychiatrists, nurses, psychologists, social workers, physiotherapists and occupational therapists (Total 56)	Psychological & pharmacological	Unknown	Unknown	10 weeks (Unknown)	Unknown
Koksvik *et al.*(2018)	Unknown (Unknown)	Not specified	Unknown	Unknown	Unknown	Unknown
Møllersen *et al.* (2009)	Psychiatrists (11), Psychologists (9), Psychiatric nurse (7), Clinical social workers (6) (Total 33)	Psychological & pharmacological	Unknown	9.0 (8.0)	7.2 months (1.4 times per month)	34.9%
Østergård *et al.* (2020)	Psychologists and social workers (Total 33)	Psychological	CBT, PDT, systemic/ humanistic	9.3 (4.78)[a] 4.12 (2.46)[b]	56 days (10 days)[a] 83 days (27 days)[b]	Unknown
Ramirez *et al.* (2008)	Unknown (Unknown)	Not specified	Unknown	Unknown	Unknown	Unknown
Rise *et al.* (2016)	Psychiatric nurse (3), psychologists (10) (Total 17)	Not specified	Unknown	Unknown	Unknown	47.4%
Sundquist *et al.* (2017)	Psychologists and social counsellors (Unknown)	Psychological & pharmacological	80% CBT, 20% unknown	6.3 (Unknown)	Unknown	35% Antidep. 16% Anx.
Werbart *et al.* (2013)	Unknown (Total 75)	Psychological & pharmacological	17% CBT, 66% PDT, 17% INT	27.4 (Unknown)	Unknown (intervention dependent)	41%

Notes: Anx = Anxiolytics/tranquilizer; Antidep = Antidepressive medication; CBT = Cognitive therapy or cognitive behavioral therapy; INT = Interpersonal therapy; MCT = Metacognitive therapy; PDT = Psychodynamic therapy.

[a]Group treatment,

[b]Individual treatment.

Forest plots depicting random effect sizes at post-treatment for TAU and comparisons with other interventions



[a] Group treatment. [b] Individual treatment.

Fig. 2. Forest plots depicting random effect sizes at post-treatment.

and Tweedie's (2000) method suggested one condition to be trimmed (adjusted $g = 0.53$). However, potential publication bias was found for between-group analysis, with significant Egger's regression intercept ($t = -3.49$, $p = 0.010$), and Duval and Tweedie's procedure indicated five conditions to be trimmed left of the mean (adjusted $g = -0.06$). Thus, the difference between the intervention and TAU groups could be close to zero when adjusted for publication bias.

Moderator analysis. Only the subgroup analysis for difference between open trials and RCTs was statistically significant (see Table 3; $Q = 4.60$, $p = 0.032$). Open trials demonstrated both

Table 3. *Subgroup analysis for clinically representative treatment (within-group random effects)*

Variable	*k*	*g*	95% CI	*Q*	*p*
Country				1.402	0.496
Denmark	2	0.563	0.144–0.982		
Norway	5	0.365	0.173–0.556		
Sweden	5	0.579	0.192–0.965		
Study design				4.598	0.032*
Open trial	5	0.680	0.407–0.953		
RCT	7	0.331	0.166–0.496		
Level of care				3.087	0.079
Primary	5	0.620	0.323–0.917		
Secondary	6	0.312	0.140–0.485		

Notes: *k* = number of comparisons.
*$p < 0.05$.

Table 4. *Meta-regression analysis of potential moderators of treatment outcome (within-group analysis)*

Variable	*k*	Point est.	*z*	*p*
Quality	12	0.003	0.09	0.926
N start	12	0.001	0.82	0.412
Female %	12	−0.175	−0.17	0.864
Age mean	12	0.003	0.16	0.876
Attrition	11	0.161	0.18	0.855
Excluded	10	0.033	0.04	0.964
Weeks after start of treatment	11	−0.001	−0.14	0.890
Publication year	12	−0.007	−0.22	0.823

Notes: *k* = number of comparisons. Publication year = years after first study, namely, 2013.

higher ES ($g = 0.68$, $p < 0.001$) and also greater heterogeneity (df[4], $Q = 53.10$, $p < 0.001$, $I^2 = 92\%$) compared to RCTs ($g = 0.33$, $p < 0.001$, df[6], $Q = 14.57$, $p = 0.024$, $I^2 = 59\%$). When one outlier was removed (Elfström, Evans, Lundgren, Johansson, Hakeberg & Carlsson, 2013) the results demonstrated the same tendency, but were non-significant (open trials $g = 0.56$, RCT $g = 0.33$, $Q = 2.17$, $p = 0.141$). Statistically significant differences were not found for country, level of care, or any of the continuous variables in the meta-regression analysis including study quality (see Table 4).

DISCUSSION

This review demonstrated a great variability in what constituted clinically representative therapy, TAU, and resulted in a small to moderate within-group ES ($g = 0.49$). It also showed high heterogeneity, which partially was explained by research design (higher effects in open trials, smaller in RCTs). Compared to a broad set of interventions, TAU was only marginally less effective ($g = -0.21$, adjusted for potential publication bias $g = -0.06$). Further, follow-up scores were not significantly different from post-treatment scores. The results should be interpreted with caution and not as the true effect of clinically representative therapy in the Nordic countries, as there is a need for more studies of higher methodological quality.

The results demonstrated a variety of methodological qualities, and a plethora of patient characteristics, professional backgrounds and treatments were reported in these, assumed to be, clinically representative therapies within the Nordic countries. However, more than half of the studies did not provide information on what intervention the psychological treatment consisted of, and half of the comparative studies did not provide mean number of sessions of both intervention and TAU. In addition, assessment for publication bias indicated potentially missing studies of TAU with higher ES. Further, high heterogeneity of ES is a finding in line with previous meta-analyses of TAU (Wampold *et al.*, 2011; Watts *et al.*, 2015). Still, no study in the present review examined TAU on its own as the primary intervention, which is noteworthy, considering the widely recognized knowledge gap on routine treatment (Hewlett & Moran, 2014). However, this review used strict inclusion criteria, which excluded articles examining other routine treatments, and this may have biased the results.

TAU in the Nordic countries demonstrated weaker within-group ES ($g = 0.49$) than EBT in IAPT ($d = 0.87$–0.88; Wakefield *et al.*, 2021). Thus, TAU seems to be less effective than evidence-based treatments, but compared to a broad set of interventions, the difference could be negligible, especially when adjusting for a potential publication bias. One excluded but highly relevant study (Nordmo, Sønderland, Havik, Eilertsen, Monsen & Solbakken, 2020) reported a larger effect size ($d = 0.85$) for patients treated between 1995–2008, who received a considerably higher number of sessions (mean of 51) than the included studies (range 4–27). This could indicate that a larger dose of psychotherapy is associated with increased effects. As noted by Shadish *et al.* (2000), study outcomes may vary depending on treatment dose as well as sample- and treatment characteristics. The effect size reported by Shadish *et al.* (2000), $d = 0.41$, which resembles the effect found in our study, may thus be representable for routine outpatient psychiatric facilities in the Scandinavian countries today, but as illustrated by the Nordmo *et al.* study, this effect could vary depending upon treatment duration.

Adding to the already acknowledged efficacy-effectiveness gap, this meta-analysis demonstrated a significant difference in ES in favor of studies conducted in open trials, in contrast to other meta-analyses (e.g., Shadish *et al.*, 2000). It has been said that the ES of effectiveness studies could easily be overestimated, for example, due to regression towards the mean and spontaneous remissions (Cuijpers, Weitz, Cristea & Twisk, 2017). However, using within-group comparisons, the same confounding factors apply to efficacy studies.

Limitations

The heterogeneity of TAU, as demonstrated both in the systematic review and meta-analysis could arguably undermine the certainty of the results. Although this meta-analysis included more studies than median number of studies found in the Cochrane library, the I^2 statistics is both prone to be imprecise and biased in small meta-analyses (von Hippel, 2015). While our results had a comparable between-size I^2 statistics compared to other TAU meta-analyses, it also indicates something else: Nordic mental health care could be very unequal, which contradicts the very presumption of egalitarian health care systems.

All Nordic countries were represented in the search strategy, but we identified no matching articles from Iceland and Finland.

Thus, only Scandinavian studies were synthesized. The database SveMed+ was only available for Scandinavian gray literature and was no longer updated as of January 2020. Although no gray literature was found eligible in the present study, there was a risk of not retrieving potential non-peer reviewed literature.

Additionally, due to low power, moderator analysis of many potentially confounding variables was not conducted, such as for instrument information (e.g., language of instruments), patient characteristics (e.g., diagnosis), or treatment information (e.g., use of pharmacotherapy or number of sessions). The latter has been demonstrated to be significantly associated with ES in one of the few meta-analyses on clinically representative therapies (Shadish *et al*., 2000). Also, many variables were presented in different ways, such as education, working status, and medication, which resulted in insufficient information for subgroup comparisons. Moreover, the subgroup analysis that was conducted may have been underpowered, which poses a risk of both alpha and beta errors. Therefore, non-significant findings of moderators in the current study must be interpreted with caution.

The present meta-analysis included studies with generic measures of symptom severity, which may have affected the results. Generic measures have been suggested justified for comparisons across diagnostic groups but are also associated with less precision and lower estimates of ES than specific measures (de Beurs, Vissers, Schoevers, Carlier, van Hemert & Meesters, 2019; Shadish *et al*., 2000).

CONCLUSION

To the best of our knowledge this is the first review to systematically assess clinically representative TAU in a restricted geographical area. Unlike previous meta-analyses, this study applied a stricter definition of TAU, and only studies with heterogeneous clinical samples and treatments were included. Although limited to the Nordic countries, this study demonstrated the ambiguity of TAU, but also its effect compared to other comparable meta-analyses.

Although with limitations, this review and meta-analysis may not only serve as a benchmarking study of clinical effect in mental health treatment for CMHP within the Nordic region, but also an in-depth examination of the nature of TAU within a region that shares many commonalities within the mental health sector. Considering the widespread use of TAU in clinical practice, and also the lack of research on it, there is a need for a pivotal change in research attitude toward routine treatment. Further research is warranted, to increase the understanding of the most commonly delivered treatment by the majority of mental health professionals, to the majority of patients.

ACKNOWLEDGMENT

Data is available on reasonable request. The study was conducted without external funding. The authors declare that they have no competing interests. MB, ML, SS, OH and AH designed the study and the search strategy. MB, ML & JL extracted studies, together with their information. MB & ML did quality assessment. MB and LGÖ were responsible for statistical analyses. All authors has been involved writing this paper, making revisions and accepted it in its final form. Special thanks to Katrine Aronsen (Head librarian at library section for Medicine and Health Sciences at NTNU University Library) who helped designing the search strategy..

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

*Arvidsdotter, T., Marklund, B., & Taft, C. (2014). Six-month effects of integrative treatment, therapeutic acupuncture and conventional treatment in alleviating psychological distress in primary care patients – follow up from an open, pragmatic randomized controlled trial. *BMC Complementary and Alternative Medicine*, *14*, 210.

Bear, H.A., Edbrooke-Childs, J., Norton, S., Krause, K.R. & Wolpert, M. (2020). Systematic review and meta-analysis: Outcomes of routine specialist mental health care for young people with depression and/or anxiety. *Journal of the American Academy of Child and Adolescent Psychiatry*, *59*(7), 810–841.

Bergmark, M., Sundberg, L.R., Markström, U. & Rosenberg, D. (2022). Evidence-based methods in rural areas – knowledge and national guideline utilization in mental health service development. *Journal of Evidence-Based Social Work*, *19*(2), 161–184.

Borenstein, M., Hedges, L., Higgins, J. & Rothstein, H. (2014). *Comprehensive meta-analysis* (Version 3.3.070) [Computer software]. Biostat. Retrieved October 3, 2023, from https://www.meta-analysis.com/

*Bratberg, G., Leira, K., Granan, L-P., Jonsbu, E., Fadnes, B. L., Thuland, S. F., & Myklebust, T. Å. (2021). Learning oriented physiotherapy (LOP) in anxiety and depression: An 18 months multicentre randomised controlled trial (RCT). *European Journal of Physiotherapy*, *23*(5), 295–304.

*Brattland, H., Koksvik, J. M., Burkeland, O., Gråwe, R. W., Klöckner, C., Linaker, O. M., Ryum, T., Wampold, B., Lara-Cabrera, M. L., & Iversen, V. C. (2018). The effects of routine outcome monitoring (ROM) on therapy outcomes in the course of an implementation process: A randomized clinical trial. *Journal of Counseling Psychology*, *65*(5), 641–652.

Cuijpers, P., Karyotaki, E., Reijnders, M. & Ebert, D.D. (2019). Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and Psychiatric Sciences*, *28*(1), 21–30.

Cuijpers, P., Quero, S., Papola, D., Cristea, I.A. & Karyotaki, E. (2021). Care-as-usual control groups across different settings in randomized trials on psychotherapy for adult depression: A meta-analysis. *Psychological Medicine*, *51*(4), 634–644.

Cuijpers, P., Weitz, E., Cristea, I. & Twisk, J. (2017). Pre-post effect sizes should be avoided in meta-analyses. *Epidemiology and Psychiatric Sciences*, *26*(4), 364–368.

de Beurs, E., Vissers, E., Schoevers, R., Carlier, I., van Hemert, A.M. & Meesters, Y. (2019). Comparative responsiveness of generic versus disorder-specific instruments for depression: An assessment in three longitudinal datasets. *Depression and Anxiety*, *36*(1), 93–102.

Downs, S.H. & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, *52*(6), 377–384.

Duval, S. & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463.

Egger, M., Davey Smith, G., Schneider, M. & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634.

*Elfström, M. L., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M., & Carlsson, S. G. (2013). Validation of the Swedish version of the clinical outcomes in routine evaluation outcome measure (CORE-OM). *Clinical Psychology & Psychotherapy*, *20*(5), 447–455.

Follmann, D., Elliott, P., Suh, I. & Cutler, J. (1992). Variance imputation for overviews of clinical trials with continuous response. *Journal of Clinical Epidemiology*, *45*(7), 769–773.

Freedland, K.E., Mohr, D.C., Davidson, K.W. & Schwartz, J.E. (2011). Usual and unusual care: Existing practice control groups in randomized controlled trials of behavioral interventions. *Psychosomatic Medicine*, *73*(4), 323–335.

*Hansson, H., Rundberg, J., Österling, A., Öjehagen, A., & Berglund, M. (2013). Intervention with feedback using outcome questionnaire 45 (OQ-45) in a Swedish psychiatric outpatient population. A randomized controlled trial. *Nordic Journal of Psychiatry*, *67*(4), 274–281.

Hewlett, E. & Moran, V. (2014). *Making mental health count: The social and economic costs of neglecting mental health care*. OECD Health Policy Studies, OECD Publishing. https://doi.org/10.1787/9789264208445-en.

Higgins, J.P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J. *et al*. (2021). *Cochrane handbook for systematic reviews of interventions*. version 6.2 (updated February 2021). Retrieved October 3, 2023, from https://training.cochrane.org/handbook.

Hooper, P., Jutai, J.W., Strong, G. & Russell-Minda, E. (2008). Age-related macular degeneration and low-vision rehabilitation: A systematic review. *Canadian Journal of Ophthalmology. Journal Canadien D'ophtalmologie*, *43*(2), 180–187.

Kazdin, A.E. (2015). Treatment as usual and routine care in research and clinical practice. *Clinical Psychology Review*, *42*, 168–178.

*Koksvik, J. M., Linaker, O. M., Gråwe, R. W., Bjørngaard, J. H., & Lara-Cabrera, M. L. (2018). The effects of a pretreatment educational group programme on mental health treatment outcomes: A randomized controlled trial. *BMC Health Services Research*, *18*(1), 665.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863.

Leichsenring, F., Steinert, C., Rabung, S. & Ioannidis, J. (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: Official Journal of the World Psychiatric Association (WPA)*, *21*(1), 133–145.

Löfholm, C.A., Brännström, L., Olsson, M. & Hansson, K. (2013). Treatment-as-usual in effectiveness studies: What is it and does it matter? *International Journal of Social Welfare*, *22*, 25–34.

Lyttkens, C.H., Christiansen, T., Häkkinen, U., Kaarboe, O., Sutton, M. & Welander, A. (2016). The core of the Nordic health care system is not empty. *Nordic Journal of Health Economics*, *4*(1), 7–27.

*Møllersen, S., Sexton, H. C., & Holte, A. (2009). Effects of client and therapist ethnicity and ethnic matching: A prospective naturalistic study of outpatient mental health treatment in northern Norway. *Nordic Journal of Psychiatry*, *63*(3), 246–255.

National Institute for Health and Care Excellence (NICE). (2011). *Common mental health problems: identification and pathways to care*. (NICE Clinical Guideline CG123). Retrieved October 3, 2023, from https://www.nice.org.uk/guidance/CG123.

Nordmo, M., Sønderland, N.M., Havik, O.E., Eilertsen, D.E., Monsen, J.T. & Solbakken, O.A. (2020). Effectiveness of open-ended psychotherapy under clinically representative conditions. *Frontiers in Psychiatry*, *11*, 384.

*Østergård, O. K., O'Toole, M. S., Svendsen, M. M., & Hougaard, E. (2020). The effect of implementing the partners for change outcome management system in a nationwide student counseling service. *Journal of Counseling Psychology*, *67*(1), 90–103.

Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210.

*Ramirez, A., Ekselius, L., & Ramklint, M. (2008). Axis V – global assessment of functioning scale (GAF), further evaluation of the self-report version. *European Psychiatry: The Journal of the Association of European Psychiatrists*, *23*(8), 575–579.

*Rise, M. B., Eriksen, L., Grimstad, H., & Steinsbekk, A. (2016). The long-term effect on mental health symptoms and patient activation of using patient feedback scales in mental health out-patient treatment. A randomised controlled trial. *Patient Education and Counseling*, *99*(1), 164–168.

Setkowski, K., Boogert, K., Hoogendoorn, A.W., Gilissen, R. & van Balkom, A. (2021). Guidelines improve patient outcomes in specialised mental health care: A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*, *144*(3), 246–258.

Shadish, W.R., Matt, G.E., Navarro, A.M. & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, *126*(4), 512–529.

Shadish, W.R., Matt, G.E., Navarro, A.M., Siegle, G., Crits-Christoph, P., Hazelrigg, M.D. *et al*. (1997). Evidence that therapy works in clinically representative conditions. *Journal of Consulting and Clinical Psychology*, *65*(3), 355–365.

*Sundquist, J., Palmér, K., Johansson, L. M., & Sundquist, K. (2017). The effect of mindfulness group therapy on a broad range of psychiatric symptoms: A randomised controlled trial in primary health care. *European Psychiatry: the Journal of the Association of European Psychiatrists*, *43*, 19–27.

Sundquist, J., Palmér, K., Memon, A.A., Wang, X., Johansson, L.M. & Sundquist, K. (2019). Long-term improvements after mindfulness-based group therapy of depression, anxiety and stress and adjustment disorders: A randomized controlled trial. *Early Intervention in Psychiatry*, *13*(4), 943–952.

Tarescavage, A.M. & Ben-Porath, Y.S. (2014). Psychotherapeutic outcomes measures: A critical review for practitioners. *Journal of Clinical Psychology*, *70*(9), 808–830.

Valentine, J.C., Pigott, T.D. & Rothstein, H.R. (2010). How many studies do you need?: A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, *35*(2), 215–247.

von Hippel, P.T. (2015). The heterogeneity statistic I(2) can be biased in small meta-analyses. *BMC Medical Research Methodology*, *15*, 35.

Wakefield, S., Kellett, S., Simmonds-Buckley, M., Stockton, D., Bradbury, A. & Delgadillo, J. (2021). Improving access to psychological therapies (IAPT) in the United Kingdom: A systematic review and meta-analysis of 10-years of practice-based evidence. *The British Journal of Clinical Psychology*, *60*(1), 1–37.

Wampold, B.E., Budge, S.L., Laska, K.M., Del Re, A.C., Baardseth, T.P., Fluckiger, C. *et al*. (2011). Evidence-based treatments for depression and anxiety versus treatment-as-usual: A meta-analysis of direct comparisons. *Clinical Psychology Review*, *31*(8), 1304–1312.

Watts, S.E., Turnell, A., Kladnitski, N., Newby, J.M. & Andrews, G. (2015). Treatment-as-usual (TAU) is anything but usual: A meta-analysis of CBT versus TAU for anxiety and depression. *Journal of Affective Disorders*, *175*, 152–167.

*Werbart, A., Levin, L., Andersson, H., & Sandell, R. (2013). Everyday evidence: Outcomes of psychotherapies in Swedish public health services. *Psychotherapy (Chicago, Ill.)*, *50*(1), 119–130.

World Health Organization. (2021). *Mental health atlas 2020*. Geneva: World Health Organization. Retrieved October 3, 2023, from. https://apps.who.int/iris/handle/10665/345946.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Appendix S1. Search strategy in PsycINFO with an Ovid interface.

Appendix S2. PRISMA checklist.

Appendix S3. Excluded after full text reading.

Appendix S4. Study design, attrition rate, and methodological quality of included studies

Appendix S5. Downs and Black checklist of quality assessment.

Appendix S6. Forest plots depicting random effect sizes at follow-up.

Received 15 November 2022, Revised 3 October 2023, accepted 5 October 2023