
Article:
Collins, C., Broglia, E. orcid.org/0000-0003-2137-8187 and Barkham, M. orcid.org/0000-0003-1687-6376 (2025) Evaluating the evidence base for university counseling services and their effectiveness using core measures: a systematic review and meta-analysis. Journal of Affective Disorders, 372. pp. 451-462. ISSN 0165-0327

https://doi.org/10.1016/j.jad.2024.12.022

Journal of Affective Disorders 372 (2025) 451–462

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Review Article

Evaluating the evidence base for university counseling services and their effectiveness using CORE measures: A systematic review and meta-analysis

Caitlin Collins, Emma Broglia *, Michael Barkham

School of Psychology, The University of Sheffield, United Kingdom of Great Britain and Northern Ireland

ARTICLE INFO

Keywords:
CORE-OM
University
Counseling
Meta-analysis
Systematic review
Effectiveness

ABSTRACT

Background: The Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM) is a pantheoretical measure of psychological functioning. Use of routine outcome measures (ROMs) assessing university counseling outcomes (i.e., clinical effectiveness) are essential to demonstrate evidence-based practice. But the evidence in higher education is limited as data usage and adoption of ROMs are not inherent within services, restricting knowledge of their effectiveness. This systematic review and meta-analysis document the evidence-base for in-house university counseling services (UCSs) and use of ROMs.

Methods: The review protocol was pre-registered on PROSPERO doi:tinyurl.com/2vw5464d and reported following PRISMA guidelines. Scopus, PsycInfo, Opengrey, reference lists, and Sheffield Star Plus were searched March 22nd, 2021 and May 29th, 2024. Subgroup analyses explored the effectiveness of services using multiple therapies compared to a single therapy.

Results: 15 studies ($N = 28,237$) were included in the narrative synthesis and 13 studies ($N = 14,795$) included in a meta-analysis. CORE-OM data showed a large pre-post effect size ($g = 1.19$) demonstrating a reduction in students' psychological distress. The effect size for services using a range of therapies was significantly larger ($g = 1.34$) than those using a single therapy ($g = 0.90$).

Limitations: The small number of papers demonstrates the limited body of evidence evaluating the effectiveness of UCSs. Lack of follow-up rates, likely due to limited ROM adoption, also hindered the evaluation.

Conclusions: Quality of studies was good and UCSs are more effective than comparative CORE-OM benchmark data of no treatment derived from a large dataset of psychological therapy services.

1. Introduction

The annual growth in students presenting to mental health and university counseling services (UCSs) with complex needs is a global issue and is cause for concern among researchers and practitioners (Auerbach et al., 2016; Rückert, 2015). UCSs provide short-term, in-house student support, including interventions such as counseling delivered by accredited trained practitioners (e.g., psychotherapists and clinical psychologists), who offer a range of support options including face-to-face, online, one-to-one, or group sessions to address students' needs (Prince, 2015). Student mental health within UK Higher Education Institutions (HEIs) has been at the center of the political agenda, with the UK Higher Education Policy Institute (HEPI) supporting the collection of institutional data on mental health services (Brown, 2016), and the UK University Mental Health Charter (Hughes and Spanner, 2019) encouraging HEIs to offer evidence-based interventions. However, the student population is growing and there has been a rapid increase in the demand to support students with mental health needs (Byrom, 2015; Royal College of Psychiatrists, 2011). Further challenges emerge due to difficulties in establishing how to accurately measure student mental health (Broglia et al., 2021; Kreß et al., 2015; Prince, 2015).

Being uniquely placed to provide support to students, universities are urged to demonstrate the effectiveness of their mental health services to address this concern (Broglia et al., 2018; Hughes and Spanner, 2019; Randall and Bewick, 2016). The importance of demonstrating clinical effectiveness is outlined in the English National Health Service Talking Therapies for anxiety and depression program, previously known as the Improving Access to Psychological Therapies (IAPT) program (Barkham et al., 2021; Clark, 2011), which aims to make the treatment for common mental health conditions more accessible (Department of Health,

* Corresponding author at: School of Psychology, Interdisciplinary Centre of the Social Sciences (ICOSS), 219 Portobello, Sheffield S1 4DP, United Kingdom of Great Britain and Northern Ireland.
E-mail address: e.l.broglia@sheffield.ac.uk (E. Broglia).

https://doi.org/10.1016/j.jad.2024.12.022
Received 16 June 2024; Received in revised form 12 November 2024; Accepted 2 December 2024
Available online 9 December 2024

2012). The development of good practice guidelines for psychological disorders maintains that clinicians must demonstrate service effectiveness (Cooper and Reeves, 2012), as ineffective interventions waste resources, place unnecessary burdens on patients, and may cause harm (Lilienfeld, 2007). Nonetheless, there is heterogeneity among mental health and UCSs, with many services experiencing increased demand without the necessary resources to expand and creating difficulties in demonstrating their effectiveness (Thorley, 2017). Services that do not use an outcome measure, or are limited to only pre- and post- measurements, are also difficult to sustain (Barkham et al., 2019; Clark, 2011).

The Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM; Evans et al., 2002) is a 34-item pantheoretical self-report measure of psychological functioning completed at the start and end of therapy to assess clients' experiences over the past week. Research indicates that the CORE-OM is a popular instrument in UCSs, including the UK, (Broglia et al., 2018; Murray et al., 2020). In addition, there is a shorter 10-item version derived from items in the CORE-OM, named the CORE-10 (Barkham et al., 2013). The psychometric properties of both measures have been reported (Barkham et al., 2010).

A survey of 113 UK HEIs found that 39 % of services used the CORE-OM, 47 % did not use a validated clinical measure, and 15 % used their own assessment or feedback form (Broglia et al., 2018). The current evidence for mental health and UCS effectiveness is therefore limited, as routine outcome measures (ROMs) are essential to demonstrate evidence-based practice (Barkham et al., 2019). Set against this backdrop, the University Mental Health Charter provides an evidence-informed framework to guide the work of universities in adopting a whole-university approach to mental health and emphasizes the need for universities to provide evidence-based interventions (Hughes and Spanner, 2019). Therefore, this systematic review and meta-analysis will help to inform the evidence base of university counseling and mental health services.

1.1. Aims

There are discrepancies in the data collected across services which inhibit benchmarking and the recognition of fields of development across sectors (Broglia et al., 2018). Thus, the comparison of UCSs can potentially advance them, determine their effectiveness, and provide evidence to strengthen institutional funding (Murray et al., 2016). Currently there are no published systematic reviews that provide an overview of the effectiveness of UCSs based on evidence using the CORE measures.

The significance of using ROMs, and thus conducting this review, was to (1) identify the evidence-base for the effectiveness of in-house counseling according to the CORE-OM and CORE-10 (a popular measure used by UCSs, as referenced by Broglia et al. (2018)) (2) review the quality of the available evidence, (3) establish what the mental health profiles are of students accessing UCSs according to the CORE measures, and (4) provide an overview of the different types of psychological interventions being delivered. The meta-analysis aimed to (1) assess whether services using multiple therapies are more effective than those using a single type of therapy, and (2) investigate whether the types of psychotherapy used within services affects counseling effectiveness. The goals of this review and meta-analysis were to integrate and cover the available literature globally whilst maintaining a neutral perspective (Cooper, 2003).

2. Methods

This systematic review adheres to the preferred reporting items for systematic reviews (PRISMA) and was pre-registered and published on PROSPERO on 15th March 2021 available at doi:tinyurl.com/2vw5464d. The pre-registration form was updated on March 22nd, 2021, to include search terms relating to the CORE-OM and CORE-10 and was re-run on May 29th, 2024.

The CORE-OM (Evans et al., 2002) taps four domains: Problems (12 items), Subjective Well-being (4 items), Functioning (12 items), and Risk (6 items: 4 risk to self-items and 2 risk-to-others items). Items are assessed on a 5-point Likert scale (0 = not at all, 4 = most or all of the time) and higher scores indicate higher symptom severity. The measure has good psychometric properties (Barkham et al., 2010), with at least 30 translations (CORE System Trust; doi:www.coresystemtrust.org.uk/home/translations/). The CORE-10 (10 items) was developed for ease of use and scoring by routine practice practitioners, to assess a client's problems, functioning, and risk (Barkham et al., 2013).

2.1. Date sources and search strategies

A comprehensive search of Scopus, PsycINFO, Opengrey, and Sheffield Star Plus was conducted on the 22nd of March 2021. A grey literature search enabled the exploration of unpublished theses to reduce publication bias (Rosenthal, 1979). The reference lists of relevant studies were searched using backward snowballing to identify further studies. Experts in the field were consulted (the SCORE consortium see doi:score-consortium.sites.sheffield.ac.uk/) to identify additional studies that were not retrieved from searching the databases. The search strategy was formed in consultation with an experienced librarian at Sheffield University. The key MeSH terms and Boolean operators to join related terms together are presented in Table 1. A final search of the databases was performed on May 29th, 2024. However, Opengrey was no longer available and therefore was not searched.

The initial searches were developed by mapping the terms to subject headings in each database. The search terms included: universities, students, counseling effectiveness, clinical outcomes in routine evaluation, and the CORE-OM or CORE-10 measures. Studies between 1990 and the date the searches were run in 2021 were sought, based on the development of the CORE measures in the 1990s (Connell et al., 1997). The searches were limited to the English language.

2.2. Study selection

Articles were exported to Zotero and screened by author CC against stringent inclusion and exclusion criteria (Table 2). The methods sections of papers were screened to establish whether the CORE measures were used, and the full texts were retrieved and assessed for review eligibility. Screening the results established whether the papers included the necessary summary statistics, and the full papers were read. Authors were contacted via email if their paper did not report on student outcomes to retrieve the post-counseling data. A total of 15 papers met the eligibility criteria. A record of the inclusion decisions, data collections, and adjustment are reported in a PRISMA flow diagram (see Fig. 1).

Table 1
Literature search strategy.

Search engine	Search strategy
PsycINFO	Clinical outcomes in routine evaluation OR CORE-10 OR CORE-OM OR counseling effectiveness OR counseling effectiveness AND Universit* OR higher education OR HEI OR College graduates or student* OR Universit* student
Scopus	TITLE-ABS-KEY (“clinical outcomes in routine evaluation” OR core-10 OR core-om OR “counseling effectiveness” OR “counseling effectiveness” AND universit* OR “higher education” OR hei OR “College graduates” OR student* OR “Universit* student”) AND (LIMIT-TO (LANGUAGE, “English”)).
OpenGrey	(“clinical outcomes in routine evaluation” OR core-10 OR core-om OR “counseling effectiveness” OR “counseling effectiveness” AND universit* OR “higher education” OR hei OR “College graduates” OR student* OR “Universit* student”) lang:”en” discipline:(05Q - Psychology) doctype:(U - Thesis)

Table 2
Inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Population	Adult samples (ages 18+) Clinical samples University students Postgraduate and undergraduate students Students in any year of university study Students referred to and accessing support from their university counseling service	Studies involving children/ adolescents under 18. Studies exclusively including non-clinical samples. Students attending a different higher education institution e.g. sixth form college. Studies where we are unable to differentiate between students and other adult samples, as they will hinder our ability to evaluate the effectiveness of UCSs for students. Students attending wellbeing services or academic support.
Intervention	Any type of counseling offered via a university counseling service	Counseling offered through another service, e.g. primary or secondary care. College or school counseling
Outcomes	Studies where the methods section reports using the CORE-OM or CORE-10	Other outcome measures e.g. the GP-CORE or the CCAPS, where the CORE measures are not reported
Measure	The means and SDs of students' pre- and post-counseling scores are reported	Studies that do not report on students' post counseling outcomes as this will not inform us of the effectiveness of therapy
Setting	Any type of setting that counseling is delivered in, including face to face, telephone, or online counseling. Any university counseling service offered to students, including those in different countries, to ensure a thorough search of the available evidence.	
Study design	Cohort studies Studies that are reported in the English language. Grey literature, including dissertation theses which are not published in peer reviewed scientific journals. Databases including the benchmarks of UCSs using the CORE-OM or CORE-10 reported online. Quantitative studies	Randomised control trials will not be included as these will not contain participants who are referred to and accessing UCSs. Sources that are not published in English. Newspaper articles, editorials, and popular media Qualitative studies

2.3. Quality assessment

Author CC independently assessed the quality of studies using the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (National Heart, Lung, and Blood Institute, 2013), which informed the critique of the literature. Articles were rated and 33 % of the papers were reviewed by an independent researcher. Disagreements over the risk of bias in studies were resolved by discussion.

2.4. Meta-analysis

A random effects meta-analysis was performed using Meta-Essentials software (Field, 2003; Suurmond et al., 2017) to quantitatively synthesise counseling effectiveness. This was selected as it was anticipated that there would be heterogeneity between studies in terms of intervention delivery, populations, and university settings. The random effects model therefore allows for within-study and between-study variation, providing a more generalizable estimate of an intervention's effect across diverse contexts (Borenstein, 2009). The following data were extracted for the meta-analysis: pre- and post-mean CORE-OM scores, SDs, sample size, and post-counseling statistics for one paper (Firth et al., 2020). The following data were created: the weighted average effect sizes using the means and SDs using Meta-essentials software (Suurmond et al., 2017), and the means and SDs for one paper (McKenzie et al., 2015) using the Cochrane handbook formula (Higgins et al., 2021).

The *r* values were imputed at 0.5 into Meta-essentials software using Workbook 4, as recommended in previous literature (Balk et al., 2012). Hedges *g* (Hedges, 1981), a corrected effect size, was chosen to account for unequal sample sizes across groups. Publication bias was assessed via visual inspection of the symmetry in a funnel plot (Liu, 2011) and by using Egger's regression (Egger et al., 1997) and Trim and Fill (Duval and Tweedie, 2000). Heterogeneity was assessed using the Q statistic (Patil, 1975), with a *p*-value, I^2 statistic and T^2 statistic. I^2 was used to quantify the percentage of observed variation due to heterogeneity, providing a more intuitive understanding of heterogeneity, whilst T^2 provides a more precise estimation of the variance across studies. T^2 was used as a measure of heterogeneity to account for between study variance, particularly given the small number of studies in the analysis. In this review, T^2 was interpreted alongside I^2, with an I^2 value above 50 % indicating moderate to high heterogeneity (Borenstein, 2009).

Subgroup analyses were conducted by assessing the type of therapy reported and study region (Hak et al., 2016). A mixed effects model (random within and fixed across subgroups) was employed for categorical subgroup analyses due to suspected substantial variability within each subgroup, including study characteristics such as participant demographics and intervention types. A fixed across subgroups effects model was used to compare the effects of region and therapy type (Borenstein, 2009). Pseudo R^2 was calculated automatically using the formulas within Meta-essentials software, Workbook 4. Finally, sensitivity analyses were performed according to study quality by removing those with low quality ratings (Bani et al., 2020; Strepparava et al., 2016).

3. Results

Following the initial search on March 22nd, 2021, and the removal of duplicates, 395 records were screened via their titles and abstracts. Of these, 355 were excluded as they did not meet the eligibility criteria, leaving 40 articles for full-text assessment. An additional search of Sheffield Star plus resulted in 301 papers, with 5 papers meeting the eligibility criteria and 296 were excluded. A final search of the databases, excluding OpenGrey, was run on May 29th, 2024, resulting in no additional papers. The process of selection resulted in 15 studies included in this review.

3.1. Study characteristics

The sample sizes in the studies ranged from 15 to 28,237 participants and all studies included more females. All studies used the CORE-OM except one study that used the CORE-10 (Broglia et al., 2021). Some studies also used additional measures to assess client symptoms including the Global Assessment Scale (GAS; Endicott et al., 1976) (Biocalti et al., 2017; Vescovelli et al., 2017) and the Symptoms Checklist-90-R (SC-90-R; Derogatis, 1994) (Bani et al., 2020; Strepparava et al., 2016). Measures such as the international classification of disease ninth edition (ICD-9-CM; World Health Organization, 1978) (Biocalti et al., 2017) and the Emotion Regulation Questionnaire (ERQ; Balzarotti et al., 2010) (Strepparava et al., 2016) were also used. Academic distress measures including the Counseling Impact on Academic Outcomes measure (CIAO; Wallace, 2012; Scruggs et al., 2023) were used (McKenzie et al., 2015), and Wright (2006) used an '*indicators of academic performance'* measure. Most studies did not report students' reasons for referral. However, Murray et al. (2014) and Murray et al. (2016) used the Association of University and College Counseling (Association of University and College Counseling (AUCC), 2009)

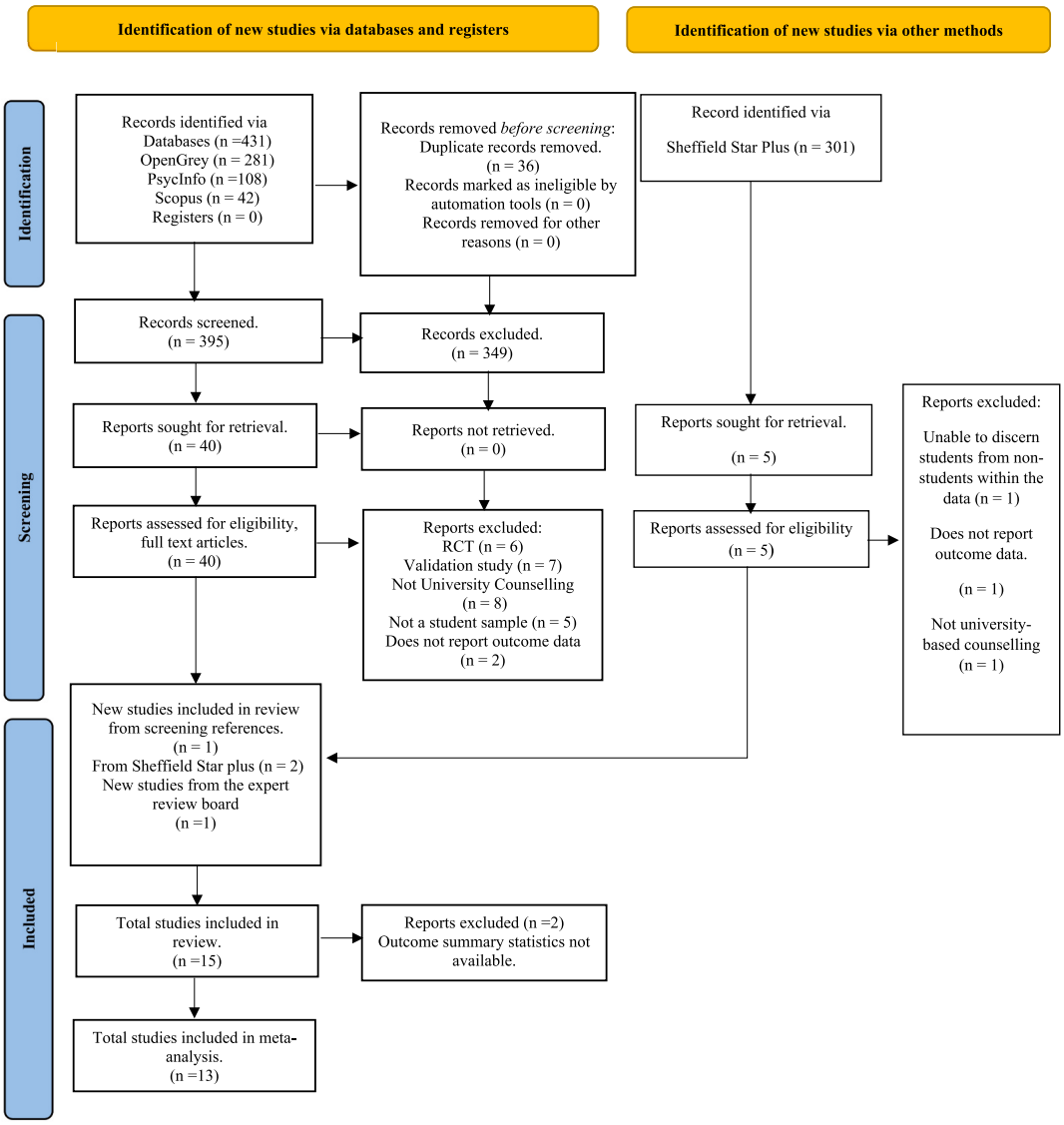

Fig. 1. PRISMA flow diagram based on Page et al., (2021).

presenting issues checklist.

3.2. Quality assessment of reviewed studies

Most of the studies reviewed had a low risk of bias. Over half (57 %) were rated as "good" quality and four (29 %) were "fair". Two (14 %) were rated as "poor" with high risk of bias. The most prominent sources of bias included low follow-up and small samples. Other sources included not assessing potential confounding variables or the number of sessions relative to therapy outcomes.

3.3. University counseling outcomes and quantifying the effectiveness

The results of the systematic review suggested university counseling effectively reduced problems/symptoms and improved wellbeing among students (Bani et al., 2020; Biocalti et al., 2017; Evans et al., 2002; Strepparava et al., 2016; Vescovelli et al., 2017). Studies that used additional measures further demonstrated an improvement in paranoid ideation (according to the SC-90-R; Strepparava et al., 2016), and significant improvement for phobic and general anxiety (Bani et al., 2020). Connell et al. (2008) found that cognitive behavioral therapy (CBT), person centered therapy, or psychodynamic therapy were effective for anxiety, interpersonal problems, depression, and work/academic problems. A recent study found that the combined rate for severe and moderately severe distress fell from 60 % pre-counseling to 27 % post-counseling for all clients with planned and unplanned endings (Broglia et al., 2021). Alternatively, three studies reported that UCSs effectively reduced psychological distress and improved academic performance (McKenzie et al., 2015; Murray et al., 2016; Wright, 2006). Studies in the UK reported rates for reliable and clinically significant improvement to be between 49 and 60 %, and studies in Italy reported rates of approximately 23 % (Jacobson and Truax, 1992). Authors of the selected studies reported using various analytical methods to assess UCS effectiveness, including descriptive analyses, *t*-tests, analysis of covariance, correlational analyses, multi-level modeling, analysis of variance, and logistic regression.

3.4. Mental health profiles of clients presenting to services

Anxiety and depression were common presenting problems in research (Murray et al., 2016), in addition to academic distress (Bani et al., 2020; Broglia et al., 2021; Murray et al., 2014, 2016). Students also primarily presented to counseling with anxiety, depression, or interpersonal problems (Connell et al., 2008). Using the ICD-9 and ICD-

10 respectively, Biocalti et al. (2017) and Vescovelli et al. (2017) found that the largest percentage of students presented to counseling with neurotic disorders. Significant research by the CORE IMS (2010) reported that anxiety and stress were the most prevalent for HEI clients (see Supplemental Materials).

3.5. CORE-OM subscale scores

Many studies reported that students presented to counseling with the highest scores on the well-being and problems (anxiety and depression combined) subscales of the CORE-OM (Bani et al., 2020; Biocalti et al., 2017; Evans et al., 2002; Strepparava et al., 2016; Vescovelli et al., 2017; Wright, 2006), with other research demonstrating the highest scores for well-being, and the lowest for risk (Broglia et al., 2021). Conversely, the pre-therapy non-risk CORE score was the highest in one study (Firth et al., 2020). A number of studies did not provide CORE-OM subscale scores, and instead provided an overall CORE-OM pre-counseling score (Connell et al., 2008; McKenzie et al., 2015; Murray et al., 2016; Stiles et al., 2015).

4. Meta-analysis

A meta-analysis was conducted using effect size data from 13 studies comprising 14,797 participants. The outcome data could not be obtained for two studies (CORE IMS, 2010; Murray et al., 2020) and were excluded from the analysis. The weighted mean effect size was $g = 1.19$ with 95 % confidence intervals not overlapping zero (95 % CI [0.95, 1.42]), $p <$. 001), indicating that across studies, university counseling was effective in reducing student psychological distress as determined by the CORE-OM. A forest plot is shown in Fig. 2 in which studies are ordered from smallest to largest effect sizes. There was significant heterogeneity in the sample, determined by Cochran's Q statistic ($Q = 1133.44, p < 0.001$). The I^2 value was 98.94 %, indicating a high amount of heterogeneity (Higgins et al., 2003). Such results indicate that subgroup analyses are appropriate, due to variance that was unaccounted for in the data.

Egger's regression (Egger et al., 1997) did not reveal significant asymmetry ($p = 0.71$) and according to the Trim and Fill (Duval and Tweedie, 2000) there were no missing studies (Fig. 3). Therefore, there is no evidence of publication bias. Following sensitivity analyses and removal of low-quality studies, the weighted mean effect size was $g = 1.24$ with 95 % confidence intervals not overlapping zero (95 % CI [0.84, 1.64]), $p <$. 001), confirming the stability of the results. Significant heterogeneity remained in the sample, determined by Cochran's Q statistic ($Q = 1096.1$ $p < 0.001$). The I^2 value was 99.36 %, indicating a high amount of heterogeneity (Higgins et al., 2003).

4.1. A comparison of services using multiple therapies with services using a single therapy

4.1.1. Moderator analyses

Subgroup analyses were conducted to explore the effectiveness of services using a range of therapies compared to those using a single therapy type. The analysis of therapy type involved $N = 13{,}307$ subjects, 10 studies and 2 subgroups. Three studies were excluded as they did not report the type of therapy used. Subgroup analyses indicated that there was a significant effect of using multiple therapies in services on post-counseling CORE-OM scores, demonstrating psychological improvement ($Q = 1.81, p = 0.179$). Studies that used mixed therapies ($g = 1.34$,

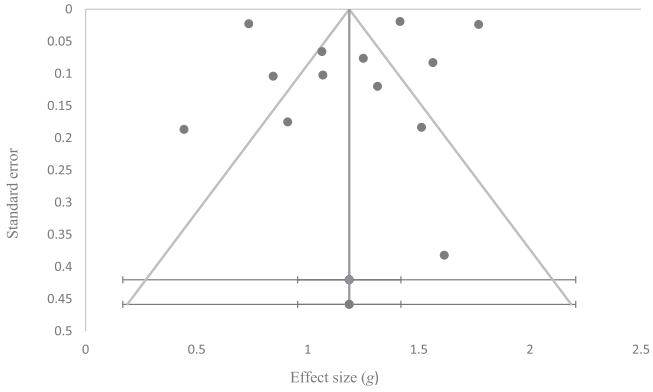

Fig. 3. Funnel plot.

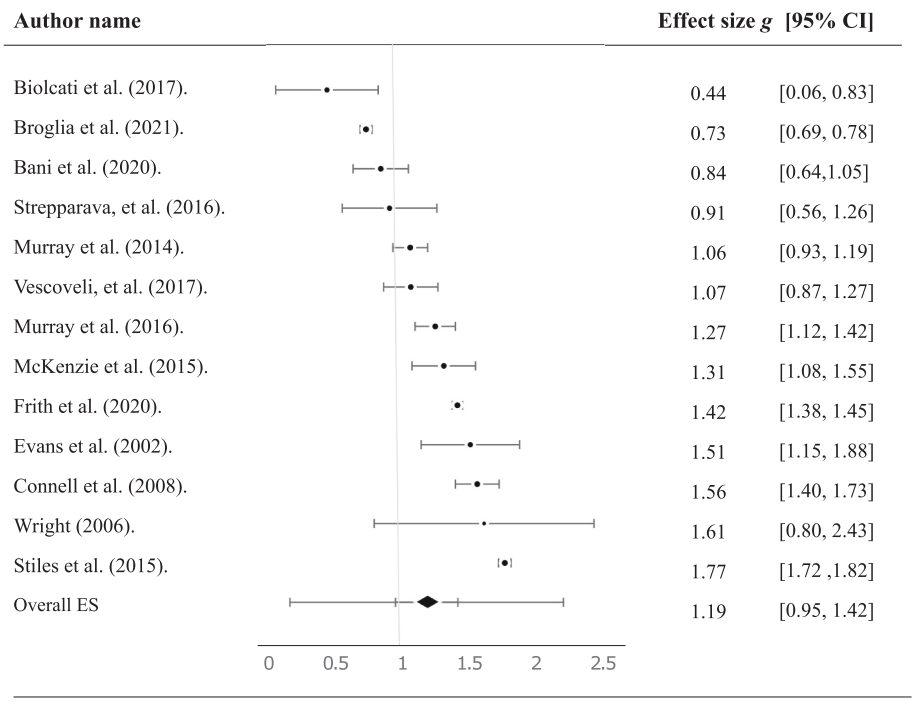

Fig. 2. Random effects meta-analysis: Post-treatment effect size for the effectiveness of university counseling.

95 % CI [1.04, 1.65]) were more effective than those using one therapy type within their service (g = 0.90, 95 % CI [0.46, 1.34]) (Table 3). Calculation of R^2 indicated that 25.7 % of the heterogeneity was accounted for by this moderator. The I^2 value was 99.53 % for studies that used mixed therapies, and 65.59 % for those that used one therapy type.

Further subgroup analyses were conducted to assess the impact of region by comparing studies in the UK to Italy. This analysis involved N = 14,797 subjects, 13 studies and 2 subgroups. The analyses indicated that there was no significant effect of region on the effectiveness of counseling determined by the CORE-OM post-counseling scores (Q = 1.56, $p = 0.67$). However, studies conducted in the UK ($g = 1.35$, 95 % CI [1.14, 1.55]) had a larger mean effect size than non-UK countries (g = 0.82, 95 % CI [0.57, 1.25]). Calculation of R^2 indicated that 18.95 % of the heterogeneity was accounted for by this moderator. The I^2 value was 99.26 % for UK countries and 98.94 % for non-UK countries, indicating a high amount of heterogeneity (Higgins et al., 2003).

4.1.2. The effectiveness of interventions

The largest effect sizes were for integrative therapy ($g = 1.77$; Stiles et al., 2015), followed by psychodynamic therapy (g = 1.56; Connell et al., 2008). A large effect size was also reported for using either psychodynamic, CBT or behavioral therapy ($g = 1.51$; Evans et al., 2002), or using a combination of psychodynamic therapy and CBT (g = 1.07; Vescovelli et al., 2017). The other most prominent effect sizes in smaller services were for CBT ($g = 0.84$; Bani et al., 2020; $g = 0.91$; Strepparava et al., 2016), with the smallest effect size for analytical psychodrama (g = 0.44; Biocalti et al., 2017) (see Table 4 for further information).

Studies also used the SCL-90-R in addition to the CORE-OM (Derogatis, 1994) to assess clients' post-therapy, finding the most significant improvements among those with phobic and general anxiety (Bani et al., 2020). Alternatively, Strepparava et al. (2016) found that clients with paranoid ideation demonstrated the most improvement post-therapy.

5. Discussion

This systematic review and meta-analysis provide robust evidence for the effectiveness of UCSs as determined by the CORE-OM and CORE-10. The findings contribute to the limited evidence and support evidence-based practice by demonstrating the value in clinicians using ROMs within UCSs (Stiles et al., 2015). There was a low risk of bias in most studies, which were of good methodological quality, and no evidence of publication bias. The overall pre-post effect size for studies using multiple therapies within their services was significantly larger than those using one therapy type.

In high quality studies, a statistically significant improvement in CORE-OM scores post- counseling was found (Broglia et al., 2021; Connell et al., 2008; Evans et al., 2002; Firth et al., 2020; McKenzie et al., 2015), supporting the effectiveness of UCSs for students. Across studies, counseling effectively reduced problems/symptoms and improved well-being among students (Bani et al., 2020; Biocalti et al., 2017; Evans et al., 2002; Strepparava et al., 2016; Vescovelli et al., 2017), and was particularly effective in reducing depression, anxiety, and improving student well-being (Broglia et al., 2021).

Connell et al. (2008) presented reliable evidence for the effectiveness of a UCS for anxiety, interpersonal problems, depression, and work/academic problems; though counseling was less effective for eating disorders, living/welfare problems, addictions, and personality problems. However, only 38.2 % of the final sample had follow-up data, which can bias the results (Mallinckrodt et al., 2001). Some studies used the SCL-90-R (Derogatis, 1994), and the CORE-OM (Evans et al., 2002) to assess client change (Bani et al., 2020; Strepparava et al., 2016), which can allow clinicians to identify any subtle client changes post-therapy (Moreau and Wiebels, 2021).

Other researchers used academic distress measures within their services. Wright (2006) found a correlation between psychological distress and academic impairment reduction. Despite the small student sample, their work is supported by two other high-quality studies. McKenzie et al. (2015) and Murray et al. (2016) found that UCSs effectively reduced psychological distress and increased resilience, which positively impacted academic performance. Hence, UCSs are not only effective for reducing psychological distress but may also enhance academic performance.

Integrative, psychodynamic, and CBT were equally effective in reducing psychological distress according to the CORE-OM (Evans et al., 2002). Two studies reported exclusive use of CBT within services large effect sizes observed (Bani et al., 2020; Strepparava et al., 2016), however the completion rates fell below 60 %, posing a risk to research validity (Chambless and Hollon, 2023). Similarly, Vescovelli et al. (2017) found no differences in effectiveness between CBT and psychodynamic therapy, reporting a large effect size and supporting the effectiveness of psychodynamic therapy. Connell et al. (2008), analyzing a larger dataset, found psychodynamic therapy was the most common in UCS (32 %) and had a large effect size, although the therapy type for other clients was unspecified.

Stiles et al. (2015) found integrative therapy to be the most common approach (41.2 %) in a large student sample, followed by person-centered therapy (36.4 %), which encourages clients to lead sessions and bring their own issues forward (British Association for Counselling and Psychotherapy, 2021). Both approaches showed large effect sizes, although multiple therapy types were used in 41.6 % of cases. Biocalti et al. (2017) also found analytical psychodrama effective based on the premise that enacting conflicts allows expression of repressed feelings (Moreno, 1985). However, the small effect size suggests it be less effective overall.

Subgroup analyses indicated that services offering a range of therapies are significantly more effective than those offering one therapy. Pseudo R^2 indicated that 25.7 % of the heterogeneity in counseling outcomes was accounted for by this moderator, suggesting that clinical effectiveness may be partly due to whether services offer a range of therapies and thereby better able to match individual needs to specific therapeutic approaches. However, there was considerable heterogeneity ($I^2 = 99.53$ %) (Deeks et al., 2021) within the subgroup of studies which report using various types of therapies to treat students. Therefore, further analyses are required to investigate other factors that may influence therapy effectiveness.

Anxiety, depression, stress, and academic distress were the most prevalent problems among students followed by trauma (Bani et al., 2020; CORE IMS (2010); Broglia et al., 2021; Murray et al., 2014, 2016), with other studies identifying interpersonal problems (Connell et al., 2008; Stiles et al., 2015). Despite indicating elevated mental health difficulties, studies have also shown high levels of positive constructs like wellbeing and functioning at intake to psychological therapies (Bani et al., 2020; Biocalti et al., 2017; Broglia et al., 2021; Strepparava et al., 2016; Vescovelli et al., 2017). Therefore these studies demonstrate a pattern of students presenting to counseling services with high-wellbeing and/or functioning alongside elevated symptomology of poor mental health.

Services that used the CCAPS (Locke et al., 2011) reported academic distress as the source of highest concern (Broglia et al., 2021). Murray et al. (2016) found a positive correlation between the CORE-OM total scores and academic impairment. Although the measure used by Wright

Table 3
The effectiveness of multiple therapies vs. a single therapy type on student outcomes.

Moderator	k	g (95 % CI)	I^2	Q	p value
Therapy type					
Multiple therapies	6	1.34 [1.04, 1.65]	99.53 %	1074.25	<0.01
Single therapy	4	0.90 [0.46, 1.34]	65.59 %	8.72	0.03

Table 4
Study characteristics and findings. *Note.* Studies are presented in alphabetical order.

Author, year of publication, study	*N*, sample size, gender	Study design and context	Measures	Intervention details	Outcomes	Sessions	Completion rate	CORE-OM baseline vs follow-up scores	Findings and effect size (ES)
Bani et al. (2020). January 2012–June 2015.	124 F[1] = 86 M = 38	University of Milano-Biococca. Italy	Symptoms Checklist 90-Revised; CORE-OM	Cognitive behavioral counseling service.	43.4 % RC, 37.7 % CSC	Total sample sessions 7.46 (SD = 3.45) Analysed participants 8.14 (SD = 2.86). Not analysed participants 6.39 (SD =4.00).	59 %	1.37 (0.62) vs 0.88 (0.52).	ES = 0.51. Significant reduction of distress and symptom scores. Significant improvement in academic self-efficacy.
Biolcati et al. (2017). January 2013–December 2016	30 F = 22 M = 8	Analytical psychodrama. University of Bologna Italy	The Italian translations of the Global Assessment scale (GAS), ICD-9-CM; The Italian validation of the CORE-OM	Psychotherapy by psycho-dramatist group psychotherapy involving 8–10 members. The work method: based on free associations, *role playing & reversal playing*	23.3 % RCSI. 13.5 % CSI. 9.1 % RI. 66 % remained stable.	40 sessions, 1 session per week, 90 mins. The intervention lasts one year.	83 %	1.77 (0.52) vs 1.51 (0.61).	Significant decrease in CORE-OM total score. Well-being and problems/ symptoms domains improved.
Broglia et al. (2021). 2017/2018.	2491 referred to counseling. F = 1544 M = 822 Not specified =108 Non-binary = 18	Data from four UK UCSs. Services embedded into UK Universities.	Two services used the CORE-OM, (*N* = 1131); CORE-10; Two services used the CCAPS (*N* = 1350).	Short-term one to one face to face in house counseling.	CORE-OM & CORE-10 (All clients). 44 % RI 52 % no RC 4 % RD	4.6 sessions (SD = 3.03). 13.4 weeks. Students received a minimum of two counseling sessions.	96 % CCAPS = 54 %; CORE = 46 %; CORE-OM = 55 %; CORE-10 = 45 %	CORE-OM & 10 = (all, *n* = 1131) pre-mean = 1.98 (SD = 0.6), post-mean = 1.46 (SD = 0.78).	CORE ES = 0.87 CCAPS ES = 0.65

Author, year of publication and study.	*N*, sample size, gender	Study design and context	Measures	Intervention details	Outcomes	Sessions	Completion rate	CORE-OM baseline vs follow-up scores	Findings and effect size (ES)
Connell et al. (2008). 1999–2001.	Total = 846. Final sample = 323 (ODA).	Assessed effectiveness of 11 UK student counseling services. UK	CORE-OM; CORE-A: TAF & EOT forms	Psychodynamic therapy (32 %). Structured brief (27 %) and/or integrative therapy (23 %).	14.2 % RI. 28.2 % no RC 56.3 % RCSI 1.2 % RD	Outcome data = 4.25 no outcome data = 1.35	38.2 %	1.78 (SD = 0.56) vs 0.91(SD = 0.55).**	ES = 1.57. The pre-post effect sizes clients with completed data for all services were high, ranging from 1.35 to 2.04.
CORE IMS (July 2010). Benchmarks for Higher Education Counseling Services.	30,519 (total). 28,237 clients. Clients. 16, 805 accepted into therapy. 751 long consultations. Female to male ratio is 70:30.	Data was contributed by 18 Universities, ranging from 16 to 10,381. UK	CORE-OM	HE sectors service-level benchmarks for assessment outcomes.	57 % recovered. 18 % improved. 75 % recovered or improved 24 % no RC 1 % RD	Overall (M = 6) With planned endings (M = 7) Unplanned endings (M = 5)	Pre-post therapy 41 % (19.8). Post-therapy 42 % 20.0)	Mean pre-therapy scores were 1.8 (SD = 0.1).	University counseling was effective for over half of the students attending the services.
Evans et al. (2002).	Total = 890 Students accessing a university counseling service (*N* = 63).	Data from 23 sites (Evans et al., 2002), included 3 UCSs UK	CORE-OM	Theoretical orientation of therapists: behavioral, cognitive behavioral or psychodynamic.	RCSI = 49 % No CSC = 51 %		Not assessed.	1.82 (SD 0.56) vs 0.97 (SD, 0.55).	ES = 1.52*.

(*continued on next page*)

Table 4 (*continued*)

Author, year of publication and study.	*N*, sample size, gender	Study design and context	Measures	Intervention details	Outcomes	Sessions	Completion rate	CORE-OM baseline vs follow-up scores	Findings and effect size (ES)
Author, year of publication and study.	*N*, sample size, gender	Study design and context	Measures	Intervention details	Outcomes	Sessions	Completion rate	CORE-OM baseline vs follow-up scores	Findings and effect size (ES)
Firth et al. (2020). The CORE National data base (2011).	466 therapists. 5472 University students, 26.2 years (SD 8.7), 75 therapists and 10 clinics	The CORE National data base (2011; Stiles et al., 2015) provided the sample. UK	CORE-OM; CORE- A	Integrative (41.2 %), Person-centred (36.4 %), Psychodynamic (22.8 %), Cognitive–behavioral (14.9 %), Structured/brief (14.6 %) Supportive (14.0 %).	NR	8.3 sessions (6.3, median and mode 6 sessions).	25.7 %	1.76 (SD = 0.61). ** vs 0.91(SD = 0.59). **	Therapist effects = 2.1 % (1.1 % - 3.7 %).
McKenzie et al. (2015).	129 F = 91 M = 37	Anonymised, pre-existing data. UK	CORE-OM; Counseling impact on academic outcomes (CIAO; Wallace, 2012).	No further details provided.	92 % experienced academic issues. RC for 67 % with academic issues CSC for 40 % with academic issues	With academic issues, M = 4.9 (SD = 1.7). Without academic issues M = 5 (SD = 1.5)	99 %	With academic issues, M = 1.8 (SD = 1.65) vs 1.0 (SD = 0.56). Without academic issues, M = 1.2 (SD = 0.055) vs M = 0.72 (SD = 0.42).	Combined scores for both groups: Pre-therapy: M = 1.74 (SD = 0.61) Post-therapy: M = 0.97 (SD = 0.55).
Murray et al. (2014). Data Collection in 2012.	360 M = 108 F = 249 *T* = 1 Missing = 2	Pre-existing CORE-OM forms, 2 years prior to the study UK	CORE-OM; AUCC	No further details provided.	NR	Not reported.	99.4 %.	1.73 (SD = 0.79) vs 0.99 (SD = 0.51)	ES =0.48 *
Murray et al. (2016). 2012–2013.	305, 2 M = 89 F = 214 T = 1	Large counseling service. Retrospective case notes; UK	CORE-OM; AUCC	No further details provided.	63 % RI 49 % CSC 2 % RD	5.4 (SD = 1.63)	84 %	1.73 (SD = 0.59) vs 0.98 (SD = 0.58) **	ES = 1.24*
Murray et al. (2020).	359 F = 249 M = 108 T = 1	No further details provided. UK	CORE-OM	No further details provided.	NR	5.36 (SD = 1.61)	NR	1.31 vs 1.15.	Evidence of response shifts associated with treatment determined by the CORE-OM.
Strepparava et al. (2016).	45 F = 24 M = 8	Cognitive relational intervention. North Italy	SC-90-R; ERQ; CORE-OM	Cognitive behavior therapy.	48.9 % RC 37.8 % CSC.	7.49, last 60 mins. 10 weekly/ fortnightly sessions.	54 %	1.29 (0.55) vs 0.82 (0.45)	Total score, Pearson, *r* = 0.66
Stiles et al. (2015). April 1999–November 2011.	69.3 % women. 4595 University students.	Data from the CORE National Research Database 2011 UK	CORE-OM; CORE-A	Integrative (41.2 %), Person-centred (36.4 %), Psychodynamic (22.8 %), Cognitive–behavioral (14.9 %) Structured/brief (14.6 %) Supportive (14.0 %)	77.2 % RI 59.6 % RCSI	8.3 sessions	25.3 %	1.85 (SD = 0.51) vs 0.90 (SD = 0.56).	ES = 1.81 Patients seen for many sessions achieve no greater improvement than patients seen for fewer sessions
Vescovelli et al. (2017). April 2012–January 2013.	149 F = 102 M = 47	Large University Northern Italy	CORE-OM; GAS	Cognitive Behavior Therapy (*N* = 73) Psychodynamic (*N* = 96)	NR	32.24 ± 12.80 50-min per week. PDT: M = 33.82 ±	74.5 %	1.84 (0.61) vs 1.21 (0.56).	ES = 0.48

(*continued on next page*)

Table 4 (*continued*)

Author, year of publication and study.	*N*, sample size, gender	Study design and context	Measures	Intervention details	Outcomes	Sessions	Completion date	CORE-OM baseline vs follow-up scores	Findings and effect size (ES)
						11.60. CBT: M = 30.66 ± 13.94			

Author, year of publication and study.	*N*, sample size, gender	Study design and context	Measures	Intervention details	Outcomes	Sessions	Completion rate	CORE-OM baseline vs follow-up	Findings and effect size (ES)
Wright (2006).	15 F = 8 M = 7	Postgraduate research students registered at a university counseling service. UK	CORE-OM; Indicators of academic performance measure	Time conscious psychological therapy Individual vs group counseling.	Time 1[2] Above clinical cut-off = 87 %, below 13 %. Time 3: Above clinical cut off = 27 %, below = 73 %.	8 sessions	80 %	Time 1, M = 1.75 (SD = 0.45). Time 3, M = 0.96 (SD = 0.47).	ES = 1.75* Link between psychological distress and academic impairment.

[1] Note: M: Male, F: Female, T: Transgender. CORE-OM: Clinical Outcomes in Routine Evaluation - Outcome Measure. *Denotes effect size calculation.
[2] Note** papers that provided clinical scores are marked with **. The scores for each paper were divided by 10 to provide the participants' mean score rather than their clinical scores. RCSC = reliable and clinically significant change, CSC = clinically significant change, RC = reliable change, RD = reliable deterioration. M = Mean, SD = Standard Deviation.

(2006) was not validated, this research supports a link between students' psychological and academic distress (Broglia et al., 2021; Murray et al., 2016; Wright, 2006), encouraging academic distress assessments within UCSs (Broglia et al., 2021).

Murray et al. (2020) found evidence of response shifts, caused by re-administering a measure (Lievens et al., 2007) among university students as some CORE-OM items were better markers of functioning prior to therapy. Clinicians must therefore ensure that scores have the same meaning and scale pre- and post- treatment (McLeod, 2001) to guarantee that self-report scores are due to symptom change (Oort et al., 2009). Additionally, most studies did not assess therapist effects, which may predict therapy outcomes (Wampold and Brown, 2005). Firth et al. (2020) found evidence of therapist effects in UCSs (2.1 %), which were smaller than primary and secondary care sectors (8.4 %, 4.1 %, respectively). UCSs should therefore consider therapist variability and the potential impact on client outcomes (Saxon et al., 2017).

5.1. Strengths and limitations

To our knowledge, this is the first systematic and meta-analytic review to evaluate the effectiveness of UCSs using data from the CORE-OM and CORE-10, covering relevant and global research. The meta-analysis enabled an analysis of the overall effectiveness of counseling and facilitated comparisons between different types of therapies while the inclusion of grey literature helped reduce publication bias (Rosenthal, 1979). The quality assessment enabled consideration of the quality and reliability of the current evidence-base and reviewed research.

The review findings are limited due to the small number of papers that were evaluated. Nevertheless, this demonstrates the limited evidence base for the effectiveness of UCS determined by the CORE measures. Sub-group analyses were conducted to compare the results of studies in the UK compared to Italy. However, UCS data could not be compared to other countries due to small sample sizes. Furthermore, recent guidelines state that all papers should be reviewed for quality by two independent researchers (Kolaski et al., 2023), though 33 % of the papers were reviewed by another researcher. In addition, low-follow up rates highlight the need for services to ensure a more rigorous approach to data collection (Barkham et al., 2019).

Finally, in terms of limitations, it is appropriate to highlight the fact that none of the studies reported in the meta-analysis employed a control condition comprising no treatment. Hence, effect sizes reflect within-group rather than between-group change and, by themselves, can only reflect the extent of student change in their mental wellbeing.

5.2. Implications

By utilizing comparative CORE-OM data derived from the English National Health Service, we are able to estimate the effect of no treatment from a study taking repeated measurement of the CORE-OM over multiple months with no intervention as an effect size of 0.24 (i.e., equivalent to the effect of natural change; Barkham et al., 2007). By comparison, the average effect in all studies reported in the current meta-analysis far exceeded this natural extent of change, thereby supporting the interpretation that the interventions were more effective than no treatment and were contributing to the reported change. Further, the actual range of reported effect sizes (i.e., from 0.44 to 1.77) are not dissimilar to those obtained from various subsamples (>10,000 to >25,000) of NHS patients based on the CORE-OM, comprising clinical and non-clinical patients at intake, and yielding effect sizes ranging from 0.60 to 1.47. Hence, services and researchers should consider benchmarking the results of university counseling effectiveness data against NHS primary care samples to add value to the impact of UCSs (Connell et al., 2008).

The limited number of papers in this review supports research urging a more coordinated approach to data collection (Barkham et al., 2019). Without routine effectiveness data the ability to ensure public confidence in the quality of counseling services is restricted (Evans et al., 2002). As this review demonstrates the value in using methods such as meta-analyses and systematic reviews to evaluate routine effectiveness data, the cost implications of this should be evaluated in future research. Additionally, although the studies within this review did not contain data gathered during the COVID pandemic, investigating the effects of the COVID pandemic on mental health interventions is an important avenue for future research.

This review supports the use of the CORE measures to detect client change and the benefits in using ROMs to identify where clients require

support (Youn et al., 2012). However, further efforts are necessary to acquire the self-report outcome data for clients, as those with unplanned endings may represent a different population to those who stay in therapy and provide outcome data (Connell et al., 2008).

The effect sizes for services using CBT and psychodynamic therapy may encourage practitioners to use these more. The significant effect size difference between the subgroups of services using various therapies compared to one therapy type may inform policy decisions for inclusive services and interventions (Center for Collegiate Mental Health, 2020). This may encourage practitioners to offer students a choice of therapy type, to potentially enhance clinical effectiveness. The review findings therefore suggest a need to continue to train and provide more jobs for counseling and psychotherapists from CBT and other backgrounds. It is important to recognise however that services are constrained to the workforce and resources that they have access to. Clinicians should also consider academic distress assessment as part of standard practice within UCSs, as many students begin counseling with high levels of academic distress (Broglia et al., 2021). Support staff should also ensure that students attending counseling understand the assessment measures used, to reduce inaccuracies in participant responding (Börjesson and Boström, 2020).

Finally, this review supports the use of bona fide outcome measures within UCSs (Broglia et al., 2018). A number of measures of psychological distress have been validated to assess students, including the GP-CORE (Sinclair et al., 2005), the CORE-OM (Evans et al., 2002), the CORE-10 (Barkham et al., 2013) and the CCAPS (Locke et al., 2011), and can be used routinely within services to ensure high-quality evidence-based practice. Using ROMs within services such as the CORE-OM should be encouraged, to enable further research to report on and publish their outcomes. Thus, this review highlights the value in conducting meta-analyses to look at the evidence for service effectiveness through use of validated outcome measures.

6. Conclusion

This systematic review and meta-analysis demonstrate that UCSs are effective. The quality of the evidence is good; however, evidence is limited due to small samples and shortfalls in the collection of outcome data due to client dropout rates. This review supports the use of ROMs including the CORE-OM and CORE-10 to assess clients attending UCSs. However, clinicians should use additional measures to assess academic distress and consider offering a range of therapies to students. The former will enable a more accurate assessment of students' psychological distress, whilst the latter may potentially reduce client drop-out. In turn, this will increase the amount of outcome data for students, which will improve the current evidence base demonstrating UCS effectiveness.

CRediT authorship contribution statement

Caitlin Collins: Writing – review & editing, Writing – original draft, Visualization, Project administration, Investigation, Formal analysis, Data curation. **Emma Broglia:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Funding acquisition. **Michael Barkham:** Supervision, Conceptualization.

Role of the funding source

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

Declaration of competing interest

Michael Barkham was a co-developer of the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) and the Clinical Outcomes in Routine Evaluation–10 (CORE-10).

Acknowledgement

We thank the SCORE Consortium https://score-consortium.sites.sheffield.ac.uk/ for offering their expertise in shaping the review and their ongoing support throughout the development of this research. We also thank Dr. Mel Simmonds-Buckley for advising on data analysis and curation, as well as John Mellor-Clark, Alex Curtis Jenkins, and Rosina Hollings for providing the 2010 benchmarking data and archived sources from CORE IMS.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2024.12.022.

References

Association of University and College Counselling (AUCC), 2009. Categorisation System, 2009. http://iaucc.ie/wp-content/uploads/2013/09/IAUCC-Categories-3rdEdnSept09.pdf.

Auerbach, R.P., Alonso, J., Axinn, W.G., Cuijpers, P., Ebert, D.D., Green, J.G., Hwang, I., Kessler, R.C., Liu, H., Mortier, P., Nock, M.K., Pinder-Amaker, S., Sampson, N.A., Aguilar-Gaxiola, S., Al-Hamzawi, A., Andrade, L.H., Benjet, C., Caldas-de-Almeida, J. M., Demyttenaere, K., Florescu, S., de Girolamo, G., Gureje, O., Haro, J.M., Karam, E. G., Kiejna, A., Kovess-Masfety, V., Lee, S., McGrath, J.J., O'Neill, S., Pennell, B.E., Scott, K., Ten Have, M., Torres, Y., Zaslavsky, A.M., Zarkov, Z., Bruffaerts, R., 2016. Mental disorders among college students in the World Health Organization World Mental Health Surveys. Psychol. Med. 46 (14), 2955–2970. https://doi.org/10.1017/S0033291716001665.

Balk, E.M., Earley, A., Patel, K., Trikalinos, T.A., Dahabreh, I.J., 2012. Empirical assessment of within-arm correlation imputation in trials of continuous outcomes. In: Methods Research Report. AHRQ Publication 12, (13). EHC141-EF. Agency for Healthcare Research and Quality.

Balzarotti, S., John, O.P., Gross, J.J., 2010. An Italian adaptation of the emotion regulation questionnaire. Eur. J. Psychol. Assess. 26 (1), 61–67. https://doi.org/10.1027/1015-5759/a000009.

*Bani, M., Zorzi, F., Corrias, D., Strepparava, M., 2020. Reducing psychological distress and improving student's well-being and academic self-efficacy: the effectiveness of a cognitive university counselling service for clinical and non-clinical situations. Brit. J. Guid. Couns. 50 (5), 757–767. https://doi.org/10.1080/03069885.2020.1840512.

Barkham, M., Mullin, T., Leach, C., Stiles, W.B., Lucock, M., 2007. Stability of the CORE-OM and BDI-I prior to therapy: evidence from routine practice. Psychol. Psychother. Theory Res. Pract. 80, 269–278. https://doi.org/10.1348/147608306X148048.

Barkham, M., Mellor-Clark, J., Connell, J., Evans, R., Evans, C., Margison, F., 2010. The CORE measures & CORE system: measuring, monitoring, and managing quality evaluation in the psychological therapies. In: Barkham, M., Hardy, G.E., Mellor-Clark, J. (Eds.), Developing and Delivering Practice-based Evidence: A Guide for the Psychological Therapies. Wiley, pp. 175–219.

Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, D., Unsworth, G., Evans, C., 2013. The CORE-10: a short measure of psychological distress for routine use in the psychological therapies. Couns. Psychother. Res. 13 (1), 3–13. https://doi.org/10.1080/14733145.2012.729069.

Barkham, M., Broglia, E., Dufour, G., Fudge, M., Knowles, L., Percy, A., Turner, A., Williams, C., SCORE Consortium, 2019. Towards an evidence-base for student wellbeing and mental health: definitions, developmental transitions, and data sets. Couns. Psychother. Res. 19 (4), 351–357. https://doi.org/10.1002/capr.12227.

Barkham, M., Saxon, D., Hardy, G.E., Bradburn, M., Galloway, D., Wickramasekera, N., Keetharuth, A.D., Bower, P., King, M., Elliott, R., Gabriel, L., Kellett, S., Shaw, S., Wilkinson, T., Connell, J., Harrison, P., Ardern, K., Bishop-Edwards, L., Ashley, K., Ohlsen, S., Pilling, S., Waller, G., Brazier, J.E., 2021. Clinical and cost-effectiveness of person-centred experiential therapy vs. cognitive behavioural therapy for moderate and severe depression as delivered in the English Improving Access to Psychological Therapies (IAPT) national programme: a pragmatic randomised non-inferiority trial [PRaCTICED]. The Lancet Psychiat. 8, 487–499. https://doi.org/10.1016/S2215-0366(21)00083-3.

*Biolcati, R., Agostini, F., Mancini, G., 2017. Analytical psychodrama with college students suffering from mental health problems: preliminary outcomes. Res. Psychother: Psychopath. Proc. Outcome. 20 (3), 201–209. https://doi.org/10.4081/ripppo.2017.272.

Borenstein, M., 2009. Introduction to Meta-analysis. John Wiley & Sons.

Börjesson, S., Boström, P.K., 2020. "I want to know what it is used for" Clients' perspectives on completing a routine outcome measure (ROM) while undergoing psychotherapy. Psychother. Res. 30 (3), 337–347. https://doi.org/10.1080/10503307.2019.1630780.

British Association for Counselling and Psychotherapy, 2021. What is person-centred counselling? www.bacp.co.uk/about-therapy/types-of-therapy/person-centred-counselling/.

Broglia, E., Millings, A., Barkham, M., 2018. Challenges in addressing student mental health in embedded counselling services: a survey of UK higher and further

education institutions. Brit. J. Guid. Couns. 46, 441–455. https://doi.org/10.1080/03069885.2017.1370695.
*Broglia, E., Ryan, G., Williams, C., Fudge, M., Knowles, L., Turner, A., Dufour, G., Percy, A., Barkham, M., on behalf of the SCORE Consortium, 2021. Profiling student mental health and counselling effectiveness: lessons from four UK services using complete data and different outcome measures. Brit. J. Guid. Couns. 51 (2), 204–222. https://doi.org/10.1080/03069885.2020.1860191.
Brown, P., 2016. The invisible problem? Improving students' mental health. Higher Education Policy Institute: HEPI Report 88 doi:www.hepi.ac.uk/2016/09/22/3592/.
Byrom, N., 2015. A summary of the report to HEFCE: understanding provision for students with mental health problems and intensive support needs. www.studentminds.org.uk/uploads/3/7/8/4/3784584/summary_of_the_hefce_report.pdf.
Center for Collegiate Mental Health, 2020. 2019 Annual Report (Publication No. STA 20-244) doi:files.eric.ed.gov/fulltext/ED602859.pdf.
Chambless, D.L., Hollon, S.D., Panter, A., 2023. Treatment validity for intervention studies. In: Cooper, H., Coutanche, M.N., McMullen, L.M., Rindskopf, D., Sher, K.J. (Eds.), APA Handbook of Research Methods in Psychology: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological, 2nd ed. American Psychological Association, pp. 673–701. https://doi.org/10.1037/0000319-030.
Clark, D.M., 2011. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. Int. Rev. Psychiatr. 23 (4), 318–327. https://doi.org/10.3109/09540261.2011.606803.
Connell, J., Barkham, M., Evans, C., Margison, F., McGrath, G., Milne, D., 1997. Clinical Outcomes in Routine Evaluation (CORE): Outcome measure - Guidelines for use (Version 1.0). PTRC Memo 333. Psychological Therapies Research Centre, University of Leeds.
*Connell, J., Barkham, M., Mellor-Clark, J., 2008. The effectiveness of UK student counselling services: an analysis using the CORE System. Brit. J. Guid. Couns. 36 (1), 1–18. Scopus. https://doi.org/10.1080/03069880701715655.
Cooper, H., 2003. Editorial. Psychol. Bull. 129, 3–9. https://doi.org/10.1037/0033-2909.129.1.3.
Cooper, M., Reeves, A., 2012. The role of randomised controlled trials in developing an evidence-base for counselling and psychotherapy. Couns. Psychother. Res. 12 (4), 303–307. https://doi.org/10.1080/14733145.2012.687388.
CORE IMS, 2010. CORE-OM recovery and improvement rates: benchmarks for higher education counselling services. Available from CORE IMS. coreims.co.uk.
Deeks, J.J., Higgins, J.P.T., Altman, D.G., 2021. Analysing data and undertaking meta-analyses. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., Welch, V.A. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 6.2. Cochrane.
Department of Health, 2012. IAPT three-year report: the first million patients. http://www.mhinnovation.net/sites/default/files/downloads/innovation/reports/Three-year-report.pdf.
Derogatis, L.R., 1994. Symptom Checklist-90-R: Administration, Scoring & Procedure Manual for the Revised Version of the SCL-90. National Computer Systems.
Duval, S.J., Tweedie, R.L., 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., Smith, G.D., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315 (7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629.
Endicott, J., Spitzer, R.L., Fleiss, J.L., Cohen, J., 1976. The global assessment scale: a procedure for measuring overall severity of psychiatric disturbance. Arch. Gen. Psychiatry 33 (6), 766–771. https://doi.org/10.1001/archpsyc.1976.01770060086012.
Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., Audin, K., 2002. Towards a standardised brief outcome measure: psychometric properties and utility of the CORE–OM. Brit. J. Psychiat. 180 (1), 51–60.
Field, A.P., 2003. The problems in using fixed-effects models of meta-analysis on real-world data. Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences 2 (2), 105–124. https://doi.org/10.1207/S15328031US0202_02.
Firth, N., Saxon, D., Stiles, W.B., Barkham, M., 2020. Therapist effects vary significantly across psychological treatment care sectors. Clin. Psychol. Psychother. 27, 770–778. https://doi.org/10.1002/cpp.2461.
Hak, T., Van Rhee, H.J., Suurmond, R., 2016. How to Interpret Results of meta-Analysis (Version 1.3). Erasmus Rotterdam Institute of Management, Rotterdam, The Netherlands. www.erim.eur.nl/researchsupport/meta-essentials/downloads.
Hedges, L.V., 1981. Distribution theory for Glass's estimator of effect size and related estimators. J. Educ. Stat. 6 (2), 107–128. https://doi.org/10.2307/1164588.
Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. BMJ 327 (7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.
Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A., 2021. Cochrane Handbook for Systematic Reviews of Interventions, Version 6.2 (Updated February 2021), 2021. Cochrane. Available from. www.training.cochrane.org/handbook.
Hughes, G., Spanner, L., 2019. The University Mental Health Charter. Student Minds, Leeds. Available from doi: universitymentalhealthcharter.org.uk/themes/.
Jacobson, N.S., Truax, P., 1992. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J. Consult. Clin. Psychol. 59, 12–19. https://doi.org/10.1037//0022-006x.59.1.12.
Kolaski, K., Logan, L.R., Ioannidis, J.P.A., 2023. Guidance to best tools and practices for systematic reviews. Syst. Rev. 12, 96. https://doi.org/10.1186/s13643-023-02255-9.
Kreß, V., Sperth, M., Hofmann, F.H., Holm-Hadulla, R.M., 2015. Psychological complaints of students: a comparison of field samples with clients of a counseling service at a typical German university. Ment. Health Prevent. 3, 41–47. https://doi.org/10.1016/j.mhp.2015.04.002.
Lievens, F., Reeve, C.L., Heggestad, E.D., 2007. An examination of psychometric bias due to retesting on cognitive ability tests in selection settings. J. Appl. Psychol. 92 (6), 1672–1682. https://doi.org/10.1037/0021-9010.92.6.1672.
Lilienfeld, S.O., 2007. Psychological treatments that cause harm. Perspect. Psychol. Sci. 2 (1), 53–70. https://doi.org/10.1111/j.1745-6916.2007.00029.x.
Liu, J., 2011. The role of the funnel plot in detecting publication and related biases in meta-analysis. Evid. Based Dent. 12, 121–122. https://doi.org/10.1038/sj.ebd.6400831.
Locke, B.D., Buzolitz, J.S., Lei, P.-W., Boswell, J.F., McAleavey, A.A., Sevig, T.D., Dowis, J.D., Hayes, J.A., 2011. Development of the Counseling Center Assessment of Psychological Symptoms-62 (CCAPS-62). J. Couns. Psychol. 58 (1), 97–109. https://doi.org/10.1037/a0021282.
Mallinckrodt, C.H., Clark, W.S., David, S.R., 2001. Accounting for dropout bias using mixed-effects models. J. Biopharm. Stat. 11 (1–2), 9–21. https://doi.org/10.1081/BIP-100104194.
*McKenzie, K., Murray, K.R., Murray, A.L., Richelieu, M., 2015. The effectiveness of university counselling for students with academic issues. Couns. Psychother. Res. 15 (4), 284–288. https://doi.org/10.1002/capr.12034.
McLeod, J., 2001. An administratively created reality: some problems with the use of self-report questionnaire measures of adjustment in counselling/psychotherapy outcome research. Couns. Psychother. Res. 1 (3), 215–226. https://doi.org/10.1080/14733140112331385100.
Moreau, D., Wiebels, K., 2021. Assessing change in intervention research: the benefits of composite outcomes. Adv. Methods Pract. Psychol. Sci. 4 (1). https://doi.org/10.1177/2515245920931930.
Moreno, J.L., 1985. Manuale di psicodramma: il teatro come terapia. Astrolabio, Rome.
Murray, A.L., McKenzie, K., Murray, K.R., Richelieu, M., 2014. Mokken scales for testing both pre- and postintervention: an analysis of the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) before and after counseling. Psychol. Assess. 26 (4), 1196–1204. https://doi.org/10.1037/pas0000015.
Murray, A.L., McKenzie, K., Murray, K.R., Richelieu, M., 2016. An analysis of the effectiveness of university counselling services. Brit. J. Guid. Couns. 44 (1), 130–139. https://doi.org/10.1080/03069885.2015.1043621.
Murray, A.L., McKenzie, K., Murray, K., Richelieu, M., 2020. Examining response shifts in the Clinical Outcomes in Routine Evaluation- Outcome Measure (CORE-OM). Brit. J. Guid. Couns. 48 (2), 276–288. https://doi.org/10.1080/03069885.2018.1483007.
National Heart, Lung, and Blood Institute, 2013. Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies. Retrieved from: www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools.
Oort, F.J., Visser, M., Sprangers, M.A., 2009. Formal definitions of measurement bias and explanation bias clarify measurement and conceptual perspectives on response shift. J. Clin. Epidemiol. 62 (11), 1126–1137. https://doi.org/10.1016/j.jclinepi.2009.03.013.
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., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. bmj 372.
Patil, K.D., 1975. Cochran's Q test: exact distribution. J. Am. Stat. Assoc. 70 (349), 186–189.
Prince, J.P., 2015. University student counseling and mental health in the United States: trends and challenges. Ment. Health Prevent. 3, 5–10. https://doi.org/10.1016/j.mhp.2015.03.001.
Randall, E.M., Bewick, B.M., 2016. Exploration of counsellors' perceptions of the redesigned service pathways: a qualitative study of a UK university student counselling service. Brit. J. Guid. Couns. 44, 86–98. https://doi.org/10.1080/03069885.2015.1017801.
Rosenthal, R., 1979. The "file drawer problem" and the tolerance for null results. Psychol. Bull. 86 (3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638.
Royal College of Psychiatrists, 2011. Mental Health of Students in Higher Education. College Report CR166.. Royal College of Psychiatrists.
Rückert, H.-W., 2015. Students' mental health and psychological counselling in Europe. Ment. Health Prevent. 3, 34–40. https://doi.org/10.1016/j.mhp.2015.04.006.
Saxon, D., Firth, N., Barkham, M., 2017. The relationship between therapist effects and therapy delivery factors: therapy modality, dosage, and on-completion. Admin. Pol. Ment. Health 44, 705–715. https://doi.org/10.1007/s10488-016-0750-5.
Scruggs, R., Broglia, E., Barkham, M., Duncan, C., 2023. The impact of psychological distress and university counselling on academic outcomes: analysis of a routine practice-based dataset. Couns. Psychother. Res. 23 (3), 781–789. https://doi.org/10.1002/capr.12640.
Sinclair, A., Barkham, M., Evans, C., Connell, J., Audin, K., 2005. Rationale and development of a general population well-being measure: psychometric status of the GP-CORE in a student sample. Br. J. Guid. Couns. 33 (2), 153–173. https://doi.org/10.1080/03069880500132581.
*Stiles, W.B., Barkham, M., Wheeler, S., 2015. Duration of psychological therapy: relation to recovery and improvement rates in UK routine practice. Brit. J. Psychiat. 207 (2), 115–122. https://doi.org/10.1192/bjp.bp.114.145565.
Strepparava, M.G., Bani, M., Zorzi, F., Corrias, D., Dolce, R., Rezzonico, G., 2016. Cognitive counselling intervention: treatment effectiveness in an Italian university Centre. Brit. J. Guid. Couns. 44 (4), 423–433. https://doi.org/10.1080/03069885.2015.1110561.

Suurmond, R., van Rhee, H., Hak, T., 2017. Introduction, comparison, and validation of Meta-essentials: a free and simple tool for meta-analysis. Res. Synth. Methods 8 (4), 537–553. https://doi.org/10.1002/jrsm.1260.

Thorley, C., 2017. Not by Degrees: Improving Student Mental Health in the UK's Universities. Institute for Public Policy Research. http://www.ippr.org/research/publications/not-by-degrees.

Vescovelli, F., Melani, P., Ruini, C., Ricci Bitti, P.E., Monti, F., 2017. University counseling service for improving students' mental health. Coll. Couns. Serv. 14 (4), 470–480. https://doi.org/10.1037/ser0000166.

Wallace, P., 2012. The impact of counselling on academic outcomes: the student perspective. AUCC J. 7, 6–11.

Wampold, B.E., Brown, G.S.J., 2005. Estimating variability in outcomes attributable to therapists: a naturalistic study of outcomes in managed care. J. Consult. Clin. Psychol. 73 (5), 914–923. https://doi.org/10.1037/0022-006X.73.5.914.

World Health Organization, 1978. International Classification of Diseases: [9th] Ninth Revision, Basic Tabulation List with Alphabetic Index. World Health Organization doi:iris.who.int/handle/10665/39473.

Wright, T., 2006. Issues in brief counselling with postgraduate research students. Couns. Psychol. Q. 19 (4), 357–372. https://doi.org/10.1080/09515070601043203.

Youn, S.J., Kraus, D.R., Castonguay, L.G., 2012. The treatment outcome package: facilitating practice and clinically relevant research. Psychother 49 (2), 115. https://doi.org/10.1037/a0027932.