



VU Research Portal

Internet interventions for mental health in university students

Harrer, Mathias; Adam, Sophia H.; Baumeister, Harald; Cuijpers, Pim; Karyotaki, Eirini; Auerbach, Randy P.; Kessler, Ronald C.; Bruffaerts, Ronny; Berking, Matthias; Ebert, David D.

published in

International Journal of Methods in Psychiatric Research
2019

DOI (link to publisher)

[10.1002/mpr.1759](https://doi.org/10.1002/mpr.1759)

document version

Publisher's PDF, also known as Version of record

document license

Article 25fa Dutch Copyright Act

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Harrer, M., Adam, S. H., Baumeister, H., Cuijpers, P., Karyotaki, E., Auerbach, R. P., Kessler, R. C., Bruffaerts, R., Berking, M., & Ebert, D. D. (2019). Internet interventions for mental health in university students: A systematic review and meta-analysis. *International Journal of Methods in Psychiatric Research*, 28(2), 1-18. [e1759]. <https://doi.org/10.1002/mpr.1759>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl



Internet interventions for mental health in university students: A systematic review and meta-analysis

Mathias Harrer¹ | Sophia H. Adam¹ | Harald Baumeister² | Pim Cuijpers³ | Eirini Karyotaki³ | Randy P. Auerbach⁴ | Ronald C. Kessler⁵ | Ronny Bruffaerts⁶ | Matthias Berking¹ | David D. Ebert³

¹Clinical Psychology and Psychotherapy, Friedrich–Alexander University Erlangen–Nuremberg, Erlangen, Germany

²Clinical Psychology and Psychotherapy, University of Ulm, Ulm, Germany

³Department of Clinical, Neuro, and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

⁴Department of Psychiatry, Columbia University, New York, New York

⁵Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts

⁶Universitair Psychiatrisch Centrum, Katholieke Universiteit Leuven, Leuven, Belgium

Correspondence

Mathias Harrer, Clinical Psychology and Psychotherapy, Department of Psychology, Friedrich–Alexander University Erlangen–Nürnberg, Nögelsbachstr. 25a, Erlangen 91052, Germany.

Email: mathias.harrer@fau.de

Abstract

Objectives: Mental health disorders are highly prevalent among university students. Universities could be an optimal setting to provide evidence-based care through the Internet. As part of the World Mental Health International College Student initiative, this systematic review and meta-analysis synthesizes data on the efficacy of Internet-based interventions for university students' mental health.

Method: A systematic literature search of bibliographical databases (CENTRAL, MEDLINE, and PsycINFO) for randomized trials examining psychological interventions for the mental health (depression, anxiety, stress, sleep problems, and eating disorder symptoms), well-being, and functioning of university students was performed through April 30, 2018.

Results: Forty-eight studies were included. Twenty-three studies (48%) were rated to have low risk of bias. Small intervention effects were found on depression ($g = 0.18$, 95% confidence interval [CI; 0.08, 0.27]), anxiety ($g = 0.27$, 95% CI [0.13, 0.40]), and stress ($g = 0.20$, 95% CI [0.02, 0.38]). Moderate effects were found on eating disorder symptoms ($g = 0.52$, 95% CI [0.22–0.83]) and role functioning ($g = 0.41$, 95% CI [0.26, 0.56]). Effects on well-being were non-significant ($g = 0.15$, 95% CI [–0.20, 0.50]). Heterogeneity was moderate to substantial in many analyses. After adjusting for publication bias, effects on anxiety were not significant anymore.

Discussion: Internet interventions for university students' mental health can have significant small-to-moderate effects on a range of conditions. However, more research is needed to determine student subsets for which Internet-based interventions are most effective and to explore ways to increase treatment effectiveness.

KEYWORDS

college, Internet, mental disorders, meta-analysis, psychotherapy

1 | INTRODUCTION

The university years are a decisive developmental phase and mark the transition from late adolescence to emerging adulthood (Arnett, 2004). Although often conceptualized as a time of positive personal

development (Evans, Forney, Guido, Patton, & Renn, 2009), post-secondary education also represents a peak onset period for the occurrence of mental disorders (Ibrahim, Kelly, Adams, & Glazebrook, 2013). It is estimated that 12–46% of all university students are affected by mental health disorders in any given year (Auerbach

et al., 2016; Auerbach et al., 2018; Blanco et al., 2008; Eisenberg, Hunt, & Speer, 2013; Verger, Guagliardo, Gilbert, Rouillon, & Kovess-Masfety, 2009). Mental disorders account for about half of the disease burden of young adults in high-income countries (WHO, 2008) and are associated with long-standing negative outcomes for both the individual and society, including lowered academic achievement (Eisenberg, Golberstein, & Hunt, 2009; Hysenbegasi, Hass, & Rowland, 2005), college dropout (Ishii et al., 2018; Kessler, Foster, Saunders, & Stang, 1995), and worse functioning in later life (Goldman-Mellor et al., 2014; Niederkroenthaler et al., 2014).

Despite the availability of effective treatment, research documents a substantial treatment gap in university students suffering from mental illness, with only one in five receiving minimally adequate treatment (Auerbach et al., 2016). The average duration of untreated mental disorders stretches between 4 to 23 years (Wang et al., 2005) and is associated with worse clinical outcomes (Ricky & O'Donnell Siobhan, 2017). Reaching students through early intervention is therefore of paramount importance.

In recent years, the potential of the Internet to facilitate help-seeking and address mental health issues in post-secondary education has become increasingly evident (Davies, Morriss, & Glazebrook, 2014; Ebert, Cuijpers, Muñoz, & Baumeister, 2017). Contents can be easily and anonymously accessed through the Internet, and Internet-delivered programs provide high cost-effectiveness and scalability (Ebert et al., 2018). In 2011, the United Kingdom's Royal College of Psychiatrists recommended to increase the availability of evidence-based Internet interventions among university students (Royal College of Psychiatrists, 2011).

Two previous meta-analyses have synthesized the effects of technology and computer-delivered interventions on university students' mental health. Davies, Morriss, and Glazebrook (2014) included 14 studies, primarily evaluating Internet-based cognitive behavioural therapy (CBT; 93%), and found moderate to large effects on depression, anxiety, and stress (standardized mean difference [SMD] = 0.43–0.73) compared with inactive control groups but no superiority of these interventions compared with active controls. Conley, Durlak, Shapiro, Kirsch, and Zahniser (2016) conducted a systematic search in 2014 targeting preventive interventions, but included somewhat outdated technology (e.g., VCR and audiotape player), and reported effects from SMD = 0.20–0.34 on depression, anxiety, and stress outcomes in non-clinical student samples compared with control groups. A systematic review on Internet-based eating disorder prevention was published in 2008 (Yager & O'Dea, 2008), with no meta-analysis performed. Given the rapid pace in technological development and the proliferation of Internet-based treatments in the last years, more updated, comprehensive knowledge on the effects of Internet-based approaches to address common mental health disorders in university students is needed. There is also a lack of synthesized information on the impact of Internet interventions on the academic functioning of students.

This systematic review aims to assess existing evidence regarding the effectiveness of Internet interventions on symptoms of common mental health disorders, well-being, and functioning outcomes among university students when compared with control groups.

2 | METHODS

This study was carried out as part of the WHO World Mental Health International College Student (WMH-ICS) initiative (WMH-ICS, 2018). The WMH-ICS aims to obtain accurate cross-national information on the prevalence, incidence, and correlates of mental, substance, and behavioural problems among college students worldwide; to describe patterns of service use, barriers to treatment, and unmet need for treatment; to investigate the associations of these disorders with role function in academic and other life domains; to evaluate the effects of a wide range of preventive and clinical interventions on student mental health, functioning, and academic performance; and to develop precision medicine clinical decision support tools to help select the right interventions for the right students (see Cuijpers et al., 2018). The WMH-ICS's meta-analysis initiative and this specific study have both been registered with PROSPERO (CRD42017068758; CRD42018090259). The procedures and results of this systematic review are outlined in accordance to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberati et al., 2009).

2.1 | Eligibility criteria

We included (a) randomized controlled trials (RCTs) in which (b) participants were enrolled at a tertiary education facility (university, college, or comparable post-secondary higher education) at the time of randomization and (c) had self-selected themselves to participate in the trial. Studies had to compare (d) a psychological intervention delivered via the Internet to (e) a control condition (wait list, no treatment, psychoeducation, and placebo) in terms of effects on (f) symptoms of common mental health problems (depression, anxiety, (di)stress, sleep problems, and disordered eating) or well-being as a (g) target outcome of the study using (h) standardized symptom measures.

Only studies (a) published in English or German were considered for inclusion. We included studies (b) published in peer-review journals, dissertations that were indexed in bibliographical databases, and unpublished full manuscripts.

Interventions were defined as eligible when the Internet was used as delivery mode, irrespective of the platform or device used (computer, tablet, mobile, and app). Technology-supported interventions with no involvement of the Internet were excluded. We defined a study's outcome as its target condition when (a) it was declared the primary outcome of the study, or (b) the article stated that the intervention was primarily aimed at this outcome (e.g., "the intervention aimed to reduce depression and anxiety" and "the program was designed to help students deal with academic stress"). When two or more articles were found to report analyses of the same study sample, only the one reporting the primary analysis of an outcome of interest was included. Studies reporting secondary analyses and studies published after the primary analysis was published were excluded. We only focused on studies in which participants actively decided to participate in the intervention (e.g., by responding to a recruitment e-mail or on-campus advertisements). We believe this best reflects the

routine practices of many universities, at which interventions are offered to all students interested to use them.

2.2 | Search strategy

Publications were identified by searching three major electronic databases, the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, and PsycINFO on April 27, 2017. There were no restrictions on publication date or status. The search was based on a string combining terms (text words, MeSH terms, and subject headings) indicative of psychological interventions in tertiary education settings and included filters for RCTs (see Table S1). The string did not contain terms restricting the search to disorders or delivery modes targeted in this analysis, thus accepting a high number of references to screen but minimizing the risk of missing relevant studies. In a second step, references in identified studies and previous systematic reviews of overlapping topics were checked for earlier publications. To identify grey literature, the WHO international clinical trials registry platform was searched for unpublished trials. Authors of study protocols without published results were contacted to determine the eligibility of potentially unpublished data. The initial search was updated on April 30, 2018.

2.3 | Study selection

Study selection began with titles and abstracts of all articles being screened for overall fit for this analysis. We then retrieved and independently assessed full texts of all selected articles for eligibility. Both steps were performed independently by two researchers (M. H. and S. A.) supervised by a senior researcher (D. E.). Discussion between researchers was initiated in case of assessor disagreement; two senior researchers (D. E. and H. B.) were consulted when disagreement could not be resolved.

2.4 | Data extraction and classification

The following data were extracted for each article if reported or applicable: (a) bibliographical data (first author and year of publication); (b) study design features (sample size, study flow, recruitment method, cut-offs for inclusion, type of control group, primary outcome/target condition and functioning outcomes, time and point of assessments, duration of the intervention, name of the intervention, and compensation) and sample characteristics (mean age, gender, and studies majors); (c) therapeutic content of the intervention(s); (d) setting (country and type of tertiary education facility); (e) treatment modality (e.g., discussion forum, website, app, or e-mail); (f) type of human guidance (intervention reminders, individual feedback, or unguided); (g) study dropout rate and missing data handling; (h) intervention attrition rates; and (i) data needed to calculate effect sizes. Functioning outcomes were defined as measures assessing either the (a) academic productivity or the (b) social/work impairment due to mental health problems of students. When relevant information could not be extracted, corresponding authors were contacted a maximum of two times to attain or clarify information. When the authors did not

respond, and information given in the article was insufficient to perform meta-analysis, the article was excluded.

Extracted data were used to classify studies into pre-specified categories for subgroup analyses. A comprehensive list of all categories along with their coding criteria can be found in Table S2.

2.5 | Quality assessment

The validity of included studies was evaluated by two researchers (M. H. and S. A.). Assessment followed the approach described by Furlan, Pennick, Bombardier, and van Tulder (2009), using the domains of the "Risk of Bias" assessment tool in RCTs developed by the Cochrane collaboration (Higgins & Green, 2011). There were 12 criteria: (a) random sequence generation; (b) allocation concealment; blinding of (c) participants, (d) personnel, or (e) outcome assessors; (f) dropout rate and (g) intention-to-treat analysis (incomplete outcome data); (h) selective outcome reporting; and other threats to validity: (i) similar groups at baseline, (j) no or similar co-interventions between intervention and control groups, (k) compliance, and (l) identical timing for outcome assessment. Typically, blinding of patients and treatment providers is difficult to achieve and maintain for non-pharmacological interventions (Boutron et al., 2007), resulting in a high risk of bias on this domain. Studies were rated as showing either "low," "high," or "unclear" risk of bias on each of these criteria. If at least six criteria were rated as "low" in one study, and if no serious flaw was detected, a study was declared to show an overall low risk of bias.

2.6 | Meta-analytic procedure

Data analysis was performed using the *meta* (Schwarzer, Carpenter, & Rücker, 2015) and *metafor* (Viechtbauer, 2010) packages in R version 3.4.3 (R Core Team, 2017). For each included study, between-group effect sizes between intervention and control groups at post-treatment were calculated. We recoded the calculated effect size when higher scores of an outcome indicated better results (e.g., well-being and academic productivity). This was done to ensure that positive effect sizes always mean that the results favour the intervention group. As many included studies had small sample sizes, we calculated Hedges' *g* as an effect size because this adjusts for small sample bias (Hedges & Olkin, 2014). The SMD (Hedges' *g*) being difficult to interpret from a clinical standpoint, we also report pooled effect sizes transformed into the numbers-needed-to-treat (NNT) or the numbers-needed-to-harm (NNH), respectively. NNTs indicate the number of patients to be treated to generate one additional positive outcome compared with no treatment. NNHs are used for negative effects to indicate the number of individuals to be treated to generate one additional negative outcome. NNTs/NNHs were calculated using the formulae by Kraemer and Kupfer (2006):

$$\text{NNT} = \frac{1}{2 \times \Phi\left(\frac{\text{SMD}}{\sqrt{2}}\right) - 1}$$

As we anticipated considerable heterogeneity among studies, we applied a random-effects pooling model using the Hartung-Knapp-

Sidik–Jonkman method (IntHout, Ioannidis, & Borm, 2014; Sidik & Jonkman, 2002) in all analyses. Prediction intervals (95%) were calculated around the pooled effect sizes, indicating the range in which the true effects of a similar future trial are expected to fall based on present evidence (Borenstein, Higgins, Hedges, & Rothstein, 2017). A significance level of $\alpha = 0.05$ (two-sided) was used for all analyses. Effect sizes were interpreted according to Cohen (1988). Using this convention, an effect size of 0.2 can be considered a small effect, 0.5 a moderate effect, and 0.8 a large effect. Heterogeneity was evaluated using the I^2 statistic (Higgins & Thompson, 2002). I^2 heterogeneity of 25% can be regarded as low, 50% as moderate, and 75% as substantial heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003).

2.7 | Analyses

2.7.1 | Main analyses

Studies with the same primary/target outcome were pooled to generate a mean effect size for each outcome. If a study reported more than one target outcome of interest, it was included in all analyses for which it provided fitting outcomes. For the analysis of functioning outcomes, we pooled all available outcome data, irrespective of having been declared a target or secondary outcome in eligible studies. Target outcomes for which less than four studies were available were not synthesized.

Some sensitivity analyses were performed. When considerable heterogeneity ($I^2 \geq 50\%$) was present in an analysis, analyses without statistical outliers were conducted. Outliers were controlled for by removing studies when their 95% confidence interval (CI) lay entirely outside the one of the pooled effect size. For all outcomes, influence analyses, also referred to as “leave-one-out” analyses (i.e., omitting one study at a time when calculating the pooled effect size; Viechtbauer & Cheung, 2010), were performed to evaluate the influence of individual studies on the overall effect. The outlying study1000 exerting the greatest influence on the overall results was then removed. For some analyses, studies with several intervention arms were included, resulting in two or more interventions being compared with the same control condition. Such comparisons are not independent, which may distort pooled effect sizes by artificially reducing heterogeneity (Borenstein, Hedges, Higgins, & Rothstein, 2009). As a sensitivity analysis, we therefore recalculated results including only one effect size per study, starting with the comparison with the smallest effect size and then the one with the largest effect size. In another approach, we combined the effects of all intervention groups in one study to create a single comparison and then recalculated the results (Higgins & Green, 2011). Lastly, we estimated the average effect using only the studies with low risk of bias.

Two approaches were used to evaluate publication bias. First, we inspected contour-enhanced funnel plots (Peters, Sutton, Jones, Abrams, & Rushton, 2008) and performed Egger's test of the intercept (Egger, Smith, Schneider, & Minder, 1997) to assess funnel plot asymmetry. When we found indications of publication bias, we used the Duval and Tweedie trim and fill procedure (Duval & Tweedie, 2000) to adjust for possible bias. These methods assume that publication

bias primarily operates through effect size. It has been argued that this assumption may not be true (Simonsohn, Nelson, & Simmons, 2014a) and that in the social sciences, publication bias is often driven by statistical significance instead (p levels; Fanelli, 2012).

We therefore additionally used p curve, a novel method to detect selective reporting bias in meta-analysis (Simonsohn et al., 2014a), which assumes that bias is caused because only statistically significant results ($p < 0.05$) are considered for publication. A p curve was attained by plotting the percentage of exact p values of all significant effects ($p < 0.05$) in an analysis. A significant test of right-skewness of the p curve indicates the presence of evidential value in an analysis. R syntax provided by Simonsohn, Nelson, and Simmons (2018; <http://www.p-curve.com>) was used to conduct the p curve analysis. To estimate the adjusted true average effect, R syntax by Simonsohn, Nelson, and Simmons (2014b) was used. It has been shown that p curve leads to accurate effect size estimates in the presence of publication bias and outperforms the trim and fill procedure (Simonsohn et al., 2014b). However, a lack of robustness has been noted for analyses with substantial heterogeneity (van Aert, Wicherts, & van Assen, 2016). Following the recommendations by van Aert, Wicherts, and van Assen (2016), we only report the adjusted effect size estimate for analyses with $I^2 < 50\%$.

These two approaches are based on different assumptions on the origin of publication bias. As it cannot be ultimately decided which assumption better reflects the field of this meta-analysis, and as both methods have evident shortcomings, we report both here.

2.7.2 | Subgroup analyses

To examine possible sources of heterogeneity, subgroup analyses were conducted for outcomes with enough studies available ($k > 10$; depression, anxiety, and stress). Subgroup analyses were performed for treatment technique (CBT interventions, interventions training one particular mental health-related skill, or other interventions), sample type (unselected, preselected through risk factors, standardized symptom cut-offs, or clinical diagnostic interview), control group (active or passive), study compensation (yes or no), type of guidance (unguided, reminders, or feedback), risk of bias rating (high or low), convenience sample rating (yes or no), and recruitment type (i.e., on campus, subject pool, online, or mixed).

3 | RESULTS

The database search yielded 44,839 records. A total of 90 studies remained for full-text analysis after duplicate removal and exclusion based on title and abstract. Forty-eight studies, with 54 comparisons between intervention and control groups, fulfilled all eligibility criteria and were included. After contacting authors of published trial protocols, we were provided with one manuscript not yet published at the time (Noone & Hogan, 2016), which was included. The study selection process and reasons for exclusion are depicted in Figure 1. References for the included studies are given in Table S3.

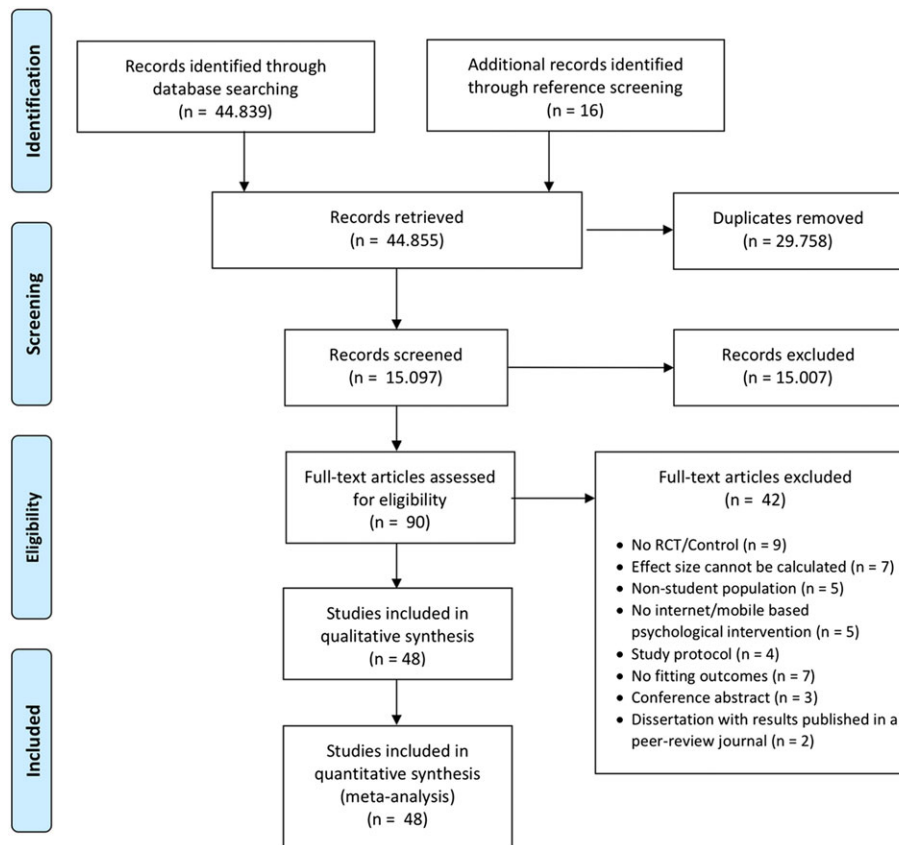


FIGURE 1 PRISMA flowchart

3.1 | Study characteristics

Detailed study characteristics can be found in Table 1. In sum, 10,583 participants were included in the trials. Sample sizes ranged from $N = 38$ to 2,638. The mean age ranged from 18.37 to 29. Seventy-four per cent ($n = 7,831$) of all participants were female. A total of 25 studies (52.08%; 27 comparisons) were conducted in general university samples in which no preselection of participants based on symptom measure or risk factors was performed. Eighteen studies (37.5%; 22 comparisons) were conducted in samples preselected through standardized cut-offs or risk factors. Only five (10.42%; five comparisons) were conducted in confirmed clinical samples.

Depression was the target or primary outcome of 26 studies, followed by anxiety ($k = 24$), stress ($k = 16$), disordered eating ($k = 9$), well-being ($k = 4$), and sleep ($k = 2$). Student functioning outcomes (e.g., work productivity and clinical impairment) were assessed in nine studies.

In 37 (77.08%) trials, an intervention was compared with a passive control group (42 comparisons), whereas 11 (22.92%) used active control conditions (12 comparisons). Among passive controlled studies, most employed wait lists ($k = 27$, 56.25%), followed by no intervention controls ($k = 6$, 12.5%) and provision of psychoeducational material without instructions for behaviour change ($k = 4$, 8.33%). Sham placebos, diaries, or recommendations for behaviour change were used as active control conditions.

Of the 53 intervention programs, 24 (45.28%) were fully unguided, nine (16.98%) included reminder mechanisms only, and 20 (37.74%) were guided interventions in which human feedback was

given to participants. Thirty-five interventions (66.04%) were CBT programs, of which 10 (18.86%) were based on third-wave CBT techniques. Eleven interventions (20.75%) were skill trainings focusing on one specific mental health-related skill (e.g., relationship or acculturation skills). Other interventions used emotional disclosure (two interventions; 3.77%; e.g., peer support and discussion groups), personalized symptom and coping-related feedback (two interventions; 3.77%), or bias modification procedures (two interventions; 3.77%) as their main therapeutic strategy.

In 43 studies (89.58%), interventions were delivered through a website. In three studies (6.25%), intervention content was provided via e-mail. Mobile-based components were used in five studies (10.42%), of which three (6.25%) employed mobile apps.

3.2 | Risk of bias

In total, 23 articles (47.92%) received a low risk of bias rating on ≥ 6 criteria. These were coded as studies of higher quality. The overall quality was suboptimal for some of the included RCTs. Five studies (10.42%) met ≤ 3 criteria, and no study met all analysed criteria. Figure 2 presents overall percentages of studies with high, low, or unclear risk of bias on each of the criteria.

3.3 | Main analyses

The pooled effects for each outcome and sensitivity analysis are summarized in Table 2. Forest plots for all outcome analyses are presented

TABLE 1 Study characteristics

Study	Type	Recruitment	Student sample	Inclusion criterion	Target condition (instrument)	Functioning outcome	Conditions	n	Technique	Age (M)	Female (%)	Guidance	FU, weeks	Dropout (post; %)	ITT	Country
Botella et al. (2010)	Cli	Campus	General	Social phobia diagnosis	Social anxiety (SAD)	Social and work impairment (MS)	1. ICBT (Talk to Me) 2. Face-to-Face CBT 3. Wait list	62 36 29	CBT	24.4	79.20	n.s. (no)	8	39.37	Yes	ESP
Cavanagh et al. (2013)	Gen	Online	General	—	Anxiety/depression (PHQ-4), stress (PSS)	—	1. Mindfulness-based intervention 2. Wait list	54 50	3rd wave CBT	24.7	88.50	No	2	44.23	Yes	UK
Cello et al. (2000)	Gen	Campus	Undergraduate	—	Eating disorder symptoms (EDE-Q weight/shape concerns)	—	1. Eating disorder prevention (StudentBodies) 2. Classroom eating disorder program (BodyTraps) 3. Wait list	27 25 24	CBT	19.6	100	Discussion group, feedback	8, 24	11.84	Yes	USA
Day et al. (2013)	Sel	Online	General	DASS ≥ 15 , 10, 18	Depression, anxiety, stress (DASS)	—	1. ICBT 2. Wait list	33 33	CBT	23.55	89.30	Feedback	6, 24	19.75	Yes	CAN
Ellis et al. (2011)	Gen	n.s.	Psychology	—	Depression, anxiety (DASS)	—	1. ICBT (Moodgym) 2. Peer support (Moodgarden) 3. No treatment	13 13 13	CBT	19.67	77.00	No	3	n.s.	n.s.	AUS
Frazier et al. (2014)	Gen	n.s.	Psychology	—	Stress (DASS)	—	1. Present Control Intervention 2. Psychoeducation	92 82	Skills training	n.s.	75.00	No	2, 5	18.68	Yes	USA
Freeman et al. (2017)	Cli	Online	General	SCI ≤ 16	Insomnia (SCI-8)	Social and work impairment (WSAS)	1. ICBT (Sleepio) 2. Wait list	1152 1486	CBT	24.7	71.50	Discussion group	10, 22	50.10	Yes	UK
Freeman (2008)	Gen	Online	General	—	Well-being (CORE-OM)	—	1. Peer support 2. Psychoeducation	51 82	Emotional disclosure	n.s.	70.00	No	10	44.12	No	UK
Gaffney (2013)	Gen	Subject pool	General	—	"Distress": depression, anxiety, stress (DASS)	—	1. Problem-solving simulator (MYLO) 2. Placebo therapist simulator (Eliza)	22 20	Skills training	21.4	78.57	No	2	12.50	No	UK
Geisner (2006)	Sel	Subject pool	Psychology	BDI ≥ 14	Depression (BDI)	—	1. Personalized feedback and coping strategies 2. Wait list	89 88	Personalized feedback	19.28	70.00	No	4	5.65	n.s.	USA
Geisner (2015)	Sel	Subject pool	Psychology	BDI ≥ 14 AUDIT ≥ 8	Depression (BDI)	—	1. Personalized feedback and coping strategies (depression) 2. PF and coping strategies (depression and alcohol) 3. Wait list	85 85 85	Personalized feedback	20.14	62.40	No	4	8.30	Yes	USA
Gibbel (2010)	Sel	n.s.	Psychology	CES-D ≥ 10 , <25	Depression (CES-D)	—	1. Spiritual intervention 2. ICBT (Moodgym) 3. No treatment	24 19 22	1. 3rd wave 2. CBT	20.45	83.00	Reminder	5, 7	27.70	No	USA
Greer (2015)	Gen	n.s.	Psychology	—	Depression, anxiety, stress (DASS)	—	1. Present control + mindfulness int. 2. Mindfulness intervention 3. Stress self-help readings	121 122 122	1. Skills training 2. 3rd wave	n.s.	75.00	Feedback	4, 12	6.51	Yes	USA
Harrer (2018)	Sel	Online	General	PSS-4 ≥ 8	Stress (PSS)	Academic impairment (PS-5)	1. ICBT for stress (StudiCare Stress) 2. Wait list	75 75	CBT	24.1	74.70	Feedback	7, 12	7.35	Yes	GER

(Continues)

TABLE 1 (Continued)

Study	Type	Recruitment	Student sample	Inclusion criterion	Target condition (instrument)	Functioning outcome	Conditions	n	Technique	Age (M)	Female (%)	Guidance	FU, weeks	Dropout (post; %)	ITT	Country
Hintz (2015)	Sel	n.s.	Psychology	PCOSES ≤ 3	Depression, anxiety, stress (DASS)	—	1. Present control intervention 2. Present control + feedback 3. Psychoeducation	97	Skills training	n.s.	70.00	Feedback (2. only)	4, 12	6.51	Yes	USA
Hoppitt (2014)	Gen	Online	Undergraduate	—	Anxiety (STAI)	—	1. Cognitive bias modification 2. Placebo	35	CBM	n.s.	79.71	No	2, 4	8	n.s.	UK
Jacobi (2007)	Gen	Mixed	General	—	Eating disorder symptoms (EDE-Q weight concerns)	—	1. Eating disorder prevention (StudentBodies) 2. Wait list	47	CBT	22.3	100	No	8, 12	3	No	GER
Kanekar (2009)	Gen	n.s.	Asian students	—	Distress (K-10)	—	1. Acculturation intervention 2. Wellness information and strategies	30	Skills training	24.6	12.80	No	8	35	No	USA
Kattelmann (2014)	Gen	Campus	Undergraduate	—	Stress (PSS)	—	1. Eating and stress management (YEAH) 2. Wait list	618	Skills training	19.3	67	No	12, 60	24.28	No	USA
Kenardy (2003)	Sel	n.s.	Psychology	ASI ≥ 21	Anxiety (ASI)	—	1. ICBT 2. Wait list	36	CBT	19.92	68.40	No	6, 24	10.84	n.s.	AUS
Kvilemo (2016)	Gen	Campus	General	—	Depression (CES-D), Well-being (PWB)	—	1. Mindfulness-based intervention 2. Expressive writing	40	3rd wave	29	75	No	8	15.56	Yes	SWE
Lee (2018)	Gen	Campus	General	—	Stress (PSS-10), Anxiety (STAI), Depression (QIDS-SR)	Work productivity (WPAI)	1. Mindfulness app (DeStressify) 2. Wait list	86	3rd wave	20.62	63	No	4	22.18	No	CAN
Levin (2014)	Gen	Campus	Undergraduate	—	"Distress": depression, anxiety, stress (DASS)	—	1. ACT-based intervention (ACT on college life) 2. Wait list	37	3rd wave	18.37	53.90	Reminder	3, 6	3.95	Yes	USA
Levin (2016)	Gen	Subject pool	Undergraduate	—	n.s., depression, anxiety, stress (DASS)	—	1. ACT-based intervention (ACT on college life) 2. Mental health education website	114	3rd wave	21.61	76.90	Reminder	3, 12	22	Yes	USA
Levin (2017)	Gen	Subject pool	Undergraduate	—	n.s., depression, anxiety (CCAPS)	—	1. ACT-based intervention (ACT on college life) 2. Wait list	40	3rd wave	20.51	66	Reminder	4	19	Yes	USA
Lintvedt (2011)	Sel	Online	General	K-10 ≥ 20	Depression (CES-D)	—	1. ICBT (Moodgym) + psychoeducation (BluePages) 2. Wait list	81	CBT	28.7	76.60	No	8	37.40	Yes	NOR
Mailey (2010)	Cli	Campus	General	Mental health disorder, n.s.	Depression (BDI), anxiety (STAI)	—	1. Physical activity intervention (IPAC5) + TAU 2. TAU	24	Skills training	25	68.10	Feedback	10	8.51	No	USA
McCall (2018)	Sel	Subject pool	General/psychology	Mini-SPIN, DSM criteria (≥ 1)	Social anxiety (SIAS)	—	1. Web-based CBT (Overcome social anxiety) 2. Wait list	51	CBT	21.86	72	Reminder	16	35.64	Yes	CAN
Melnyk (2015)	Gen	Campus	General	—	Depression (PHQ-9), anxiety (GAD-7)	Grade point average	1. ICBT skills training (COPE) 2. No treatment	61	CBT	18.67	86.40	No	7, 10	23.14	No	USA

(Continues)

TABLE 1 (Continued)

Study	Type	Recruitment	Student sample	Inclusion criterion	Target condition (instrument)	Functioning outcome	Conditions	n	Technique	Age (M)	Female (%)	Guidance	FU, weeks	Dropout (post; %)	ITT	Country
Mogase (2013)	Sel	Online	Undergraduate	BDI \geq 12	Depression (BDI)	—	1. Cognitive bias modification 2. Wait list	20	CBM	22.87	95.24	No	1	2.30	No	ROM
Morris (2015)	Gen	Online	General	—	Anxiety (STAI), insomnia (PSQI)	—	1. iCBT for anxiety 2. iCBT for insomnia 3. Wait list	43 48 47	CBT CBT	20.5	67.39	Reminder	6	18.80	Yes	UK
Mullin (2015)	Sel	Mixed	General	Depression, anxiety symptoms (self-identified)	Depression (PHQ-9), anxiety (GAD-7)	Mental health-related disability (SDS)	1. iCBT (UniWell-being Course) 2. Wait list	30 24	CBT	27.86	61.81	Feedback	6, 12	19.50	Yes	AUS
Musiat (2014)	Gen	Online	General	—	Depression (PHQ-9), anxiety (GAD-7)	—	1. Trait-focused iCBT 2. Psychoeducation	519 528	CBT	n.s.	70.50	No	6, 12	50.33	n.s.	UK
Nguyen-Feng (2015)	Gen, Sel	Campus	Psychology	IPV history	"Distress": depression, stress (DASS)	—	1. Present control intervention 2. Wait list	329 171	Skills training	n.s.	62	No	6	14	Yes	USA
Noone (2018)	Gen	n.s.	Psychology	—	Well-being (WEMWBS)	—	1. Meditation app (Headspace) 2. Sham meditation placebo	43 48	3rd wave	20.92	75.82	Reminder	6	22	Yes	IRL
Orbach (2007)	Gen	Online	General	—	Test anxiety (TAI)	—	1. iCBT 2. Placebo	30 28	CBT	23.67	72.55	F2F introduction	6	32.56	No	UK
Räsänen (2016)	Gen	Online	General	—	Anxiety, depression, stress (DASS), Well-being (MHC)	—	1. ACT-based intervention (Student compass) 2. Wait list	33 35	3rd wave	24.29	85.30	Feedback	7	6	Yes	FIN
Richards (2016)	Cli	Online	General	GAD-7 \geq 10	Anxiety (GAD-7)	Social amd work impairment (WASA)	1. iCBT (Calming anxiety) 2. Wait list	70 67	CBT	23.82	77.40	Feedback	6	18.25	Yes	IRL
Saekow (2015)	Sel	Mixed	General	WCS \geq 47	Eating disorder symptoms (EDE-Q)	Clinical impairment (CIA)	1. Eating disorder prevention (StudentBodies) 2. Wait list	31 34	CBT	23	100	Feedback	10	36.92	Yes	USA
Sánchez-Ortiz (2011)	Cli	Online	General	Bulimia, ED NOS diagnosis	Eating disorder symptoms (EDE)	—	1. iCBT (Overcome bulimia online) 2. Wait list	36 31	CBT	23.9	100	Feedback	12, 24	11.84	Yes	UK
Sarniak (2009)	Gen	Subject pool	Psychology	—	Depression (GES-D)	—	1. Expressive writing, positive events 2. No treatment	47 44	Emotional disclosure	19.6	77	Reminder	4	9	No	USA
Sethi (2010)	Sel	Campus	Undergraduate	DASS, "severe" level	Depression, anxiety (DASS)	—	1. iCBT (Moodgym) 2. Therapist-delivered CBT 3. Blended CBT 4. No treatment	9 10 9 10	CBT CBT CBT	19.74	65.79	Feedback	3	n.s.	n.s.	AUS
Stallman (2016)	Sel	Mixed	General	K-10 \geq 16	Depression, anxiety, stress (DASS)	—	1. Low-intensity iCBT (referral to Moodgym, VirtualClinic, TheDesk, AnxietyOnline) 2. Personalized feedback	52 55	CBT	23	92	Feedback	8, 24, 48	18.69	Yes	AUS
Stice 2012 (2014)	Sel	Mixed	General	Body image concerns (interview)	Eating disorder symptoms (eating disorder interview based on DSM)	—	1. Dissonance online intervention (eBody) 2. Dissonance group intervention 3. Brochure with body image tips	19 39 20	CBT CBT	21.6	100	No	6, 52, 104	1.90	Yes	USA

(Continues)

TABLE 1 (Continued)

Study	Type	Recruitment	Student sample	Inclusion criterion	Target condition (instrument)	Functioning outcome	Conditions	n	Technique	Age (M)	Female (%)	Guidance	FU, weeks	Dropout (post; %)	ITT	Country
Taylor (2016)	Sel	Mixed	General	WCS \geq 47	Eating disorder symptoms (EDE-Q)	Clinical impairment (CIA)	1. Eating disorder prevention (image and Mood) 2. Wait list	106	CBT	20	100	Feedback, discussion group	12.52, 104	9.70	Yes	USA
Taylor (2006)	Sel	Mixed	General	WCS \geq 50	Eating disorder symptoms (EDE-Q)	—	1. Eating disorder prevention (StudentBodies) 2. Wait list	206	CBT	20.8	100	Feedback, discussion group	8, 52	12.29	Yes	USA
Winzelberg (2000)	Gen	Campus	General	—	Eating disorder symptoms (EDE-Q weight concerns)	—	1. Eating disorder prevention (StudentBodies) 2. Control, n.s.	24	CBT	20	100	Feedback, discussion group	8, 12	13.30	Yes	USA
Zabinski (2001)	Sel	Campus	Psychology	BSQ \geq 110	Eating Disorder symptoms (EDE-Q)	—	1. Eating disorder prevention (StudentBodies) 2. Wait list	27	CBT	19.3	100	Feedback, discussion group	4, 10	7.35	Yes	USA

Note. AUS: Australia; CAN: Canada; CBM: cognitive bias modification; CBT: cognitive behavioural therapy; Cli: clinical sample; ESP: Spain; FIN: Finland; FU: follow-up; Gen: general sample; GER: Germany; iCBT: Internet-based cognitive behavioural therapy; IRL: Ireland; ITT: intention-to-treat analysis; n.s.: not specified; NED: Netherlands; NOR: Norway; ROM: Romania Sel: selected sample; SWE: Sweden; TAU: treatment as usual; UK: United Kingdom; USA: United States of America. ASI = Anxiety Sensitivity Index; BDI = Beck Depression Inventory; CCAPS = Counseling Center Assessment of Psychological Symptoms; SCI = Sleep Condition Indicator; CES-D = Center for Epidemiological Studies' Depression Scale; CIA = Clinical Impairment Assessment Questionnaire; CORE-OM = Clinical Outcomes in Routine Evaluation - Outcome Measure; DASS = Depression Anxiety Stress Scale; DSM = Diagnostic and Statistical Manual of Mental Disorders; ED NOS = Eating Disorder not otherwise specified; EDE = Eating Disorder Examination; EDE-Q = Eating Disorder Examination - Self Report Questionnaire; GAD-7 = Generalized Anxiety Disorder 7; IPV = Interpersonal Violence; PHQ-4 = Patient Health Questionnaire - 4 item version; PSS = Perceived Stress Scale; PS-5 = Presenteeism Scale for Students; QIDS-SR = Quick Inventory of Depressive Symptomatology - Self Report; SAD = Social Avoidance and Distress Scale; STAI = Spielberger State-Trait Anxiety Inventory; WCS = Weight Concern Scale; WEMWBS = Warwick-Edinburgh Mental Well-being Scale; WPAL = Work Productivity and Activity Impairment questionnaire.

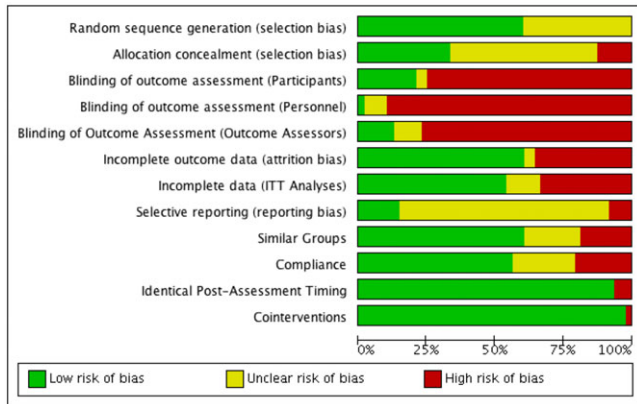


FIGURE 2 Risk of bias summary. ITT: intention-to-treat analysis

TABLE 2 Pooled effects of Internet interventions on target outcomes compared with control groups

Target outcome	n_c	Effect size			Heterogeneity			95% PI	NNT
		g	95% CI	p	I^2	95% CI	p		
Depression	31	0.18	[0.08, 0.27]	0.001	44	[15, 64]	0.002	-0.26-0.62	9.80
Influence analysis ("leave-one-out") ^a	30	0.20	[0.11, 0.30]	<0.001	38	[3, 60]	0.020	-0.21-0.62	8.93
One ES/study (lowest)	26	0.17	[0.06, 0.28]	0.003	48	[18, 67]	0.004	-0.27-0.61	10.42
One ES/study (highest)	26	0.19	[0.08, 0.31]	0.002	52	[26, 70]	0.001	-0.29-0.68	9.43
One ES/study (combined)	26	0.19	[0.08, 0.29]	0.001	52	[24, 69]	0.001	-0.26-0.63	9.43
Only low risk of bias	13	0.21	[0.03, 0.40]	0.025	59	[24, 78]	0.004	-0.36-0.79	8.47
Anxiety	27	0.27	[0.13, 0.40]	<0.001	51	[24, 68]	<0.001	-0.36-0.90	6.58
Outliers removed ^b	25	0.31	[0.19, 0.43]	<0.001	34	[0, 59]	0.059	-0.20-0.82	5.75
Influence analysis ("leave-one-out") ^c	26	0.30	[0.17, 0.43]	<0.001	45	[12, 65]	0.010	-0.30-0.91	5.95
One ES/study (lowest)	24	0.27	[0.13, 0.42]	<0.001	54	[27, 71]	<0.001	-0.37-0.92	6.58
One ES/study (highest)	24	0.29	[0.15, 0.43]	<0.001	45	[12, 66]	0.009	-0.32-0.89	6.17
One ES/study (combined)	24	0.28	[0.14, 0.42]	<0.001	55	[29, 72]	<0.001	-0.34-0.91	6.41
Only low risk of bias	11	0.22	[0.01, 0.43]	0.041	46	[0, 73]	0.050	-0.42-0.86	8.06
Stress	18	0.20	[0.02, 0.38]	0.030	72	[56, 83]	<0.001	-0.50-0.90	8.93
Outliers removed ^d	16	0.18	[0.05, 0.32]	0.010	57	[24, 75]	0.003	-0.27-0.64	9.80
Influence analysis ("leave-one-out") ^e	17	0.15	[0.01, 0.29]	0.038	64	[39, 78]	<0.001	-0.36-0.66	11.90
One ES/study (lowest)	16	0.22	[0.02, 0.42]	0.034	75	[59, 85]	<0.001	-0.53-0.97	8.06
One ES/study (highest)	16	0.24	[0.06, 0.42]	0.014	71	[52, 83]	<0.001	-0.45-0.93	7.46
One ES/study (combined)	16	0.23	[0.04, 0.42]	0.024	75	[59, 85]	<0.001	-0.50-0.95	7.69
Only low risk of bias	9	0.30	[-0.05, 0.66]	0.084	80	[62, 89]	<0.001	-0.75-1.36	5.95
Well-being	4	0.15	[-0.20, 0.50]	0.259	3	[0, 85]	0.378	-0.64-0.94	11.90
Influence analysis ("leave-one-out") ^f	3	0.25	[0.11, 0.39]	0.016	0	[0, 0]	0.930	-0.18-0.68	7.14
Only low risk of bias	3	0.12	[-0.55, 0.79]	0.526	31	[0, 93]	0.237	-2.98-3.21	14.71
Eating disorders	9	0.52	[0.22, 0.83]	0.004	61	[18, 81]	0.009	-0.32-1.37	3.50
Influence analysis ("leave-one-out") ^g	8	0.61	[0.35, 0.86]	<0.001	39	[0, 73]	0.120	-0.04-1.25	2.99
Only low risk of bias	5	0.63	[0.14, 1.12]	0.023	59	[0, 85]	0.046	-0.50-1.77	2.96
Functioning	9	0.41	[0.26, 0.56]	<0.001	53	[1, 78]	0.029	0.02-0.81	4.39
Influence analysis ("leave-one-out") ^h	8	0.45	[0.10, 0.81]	<0.001	31	[0, 69]	0.180	0.10-0.81	4.00
Only low risk of bias	5	0.41	[0.22, 0.60]	0.004	54	[0, 83]	0.070	-0.05-0.87	4.39

Note. ES: effect size; n_c : number of comparisons; NNT: number-needed-to-treat; PI: prediction intervals.

^aRemoved in leave-one-out-analysis: Greer, 2015 (Mindfulness).

^bRemoved as outliers: Greer, 2015 (Mindfulness); Gaffney et al., 2013.

^cRemoved in leave-one-out-analysis: Greer, 2015 (Mindfulness).

^dRemoved as outliers: Greer, 2015 (Mindfulness); Day et al., 2013.

^eRemoved in leave-one-out-analysis: Day et al., 2013.

^fRemoved in leave-one-out-analysis: Kvillemo et al., 2016.

^gRemoved in leave-one-out-analysis: Zabinski et al., 2000.

^hRemoved in leave-one-out-analysis: Lee et al., 2018.

in Figures 3–8. Detailed results of the influence analyses are displayed in Figure S1. Funnel plots and p curves are presented in Figures S2 and S3. Detailed results on the publication bias analyses are reported in Tables S4 and S5.

3.3.1 | Depression

We could compare the effect of Internet interventions on symptoms of depression with control groups in 31 comparisons. The overall effect size was $g = 0.18$ (95% CI [0.08, 0.27]), which corresponds to an NNT of 9.80. Heterogeneity was moderate ($I^2 = 44\%$; 95% CI [15, 64]). The prediction interval ranged from $g = -0.26$ to 0.62. Similar effects emerged in all sensitivity analyses, including analyses in which outliers were removed, when only the highest, lowest, and combined effect of

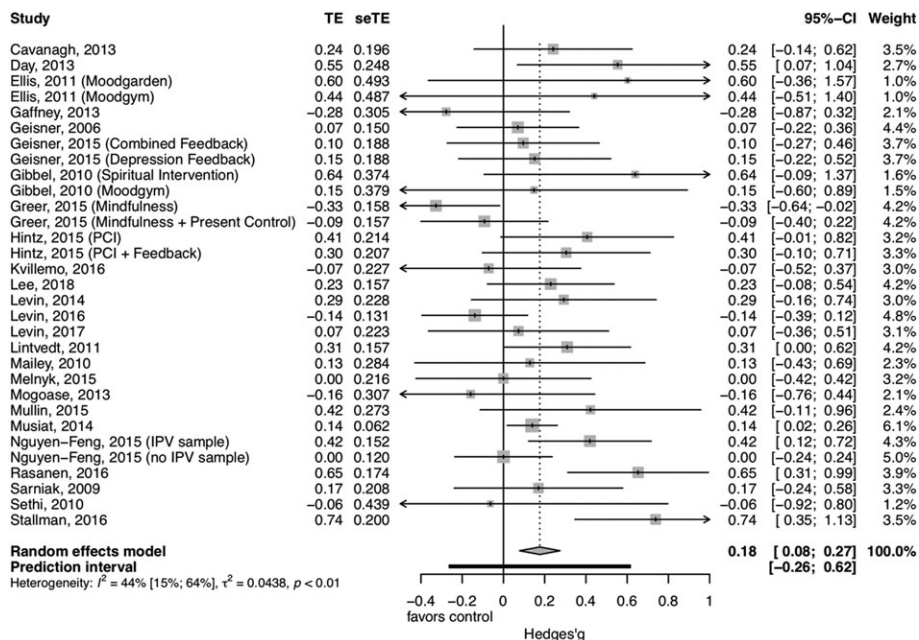


FIGURE 3 Forest plot for depression outcomes

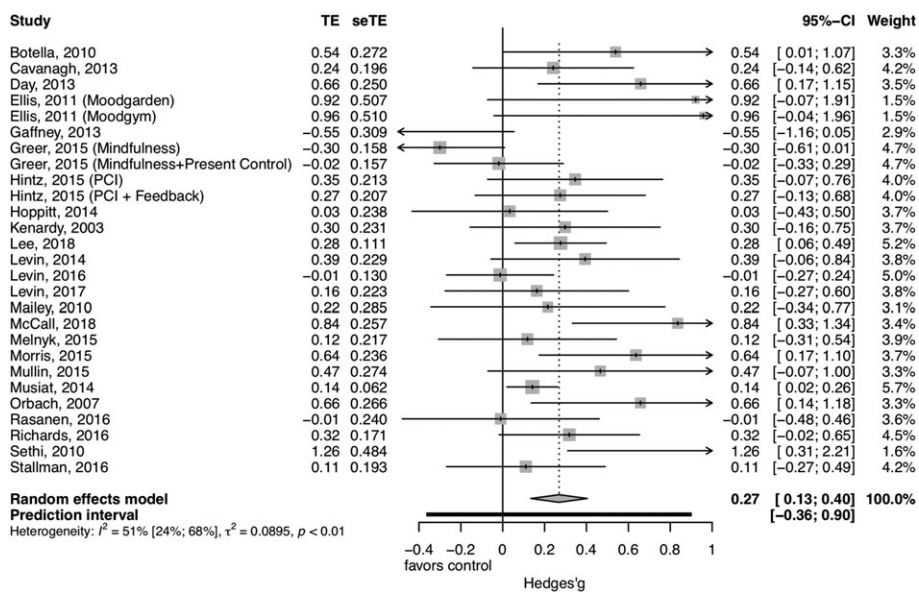


FIGURE 4 Forest plot for anxiety outcomes

multiple comparisons was considered and when only studies with a low risk of bias rating were included, and in the influence analysis. This supports the robustness of this finding. Funnel plot and Egger's test did not hint at publication bias (see Figure S2 and Table S4). The results of the p curve analysis were inconclusive. The test for right-skewness did not indicate the presence of evidential value ($p_{Full} = 0.269$, $p_{Half} = 0.427$, $k = 8$; see Table S5). However, the test for flatness was not significant ($p_{Full} = 0.149$, $p_{Half} = 0.981$, and $p_{Binomial} = 0.423$); p curve's estimate of the average true effect size was Cohen's $d = 0.09$.

3.3.2 | Anxiety

The overall effect size for anxiety (27 comparisons) was $g = 0.27$ (95% CI [0.13, 0.40]), which corresponds with an NNT of 6.58. Heterogeneity was moderate ($I^2 = 51\%$; 95% CI [24, 68]). The prediction interval

ranged from $g = -0.36$ to 0.90. When two outliers were removed, the between-study heterogeneity became non-significant ($I^2 = 34\%$; 95% CI [0-59]; $p = 0.059$) and a similar effect size of $g = 0.31$ (95% CI [0.19, 0.43], NNT = 5.75) resulted. Results of all the other sensitivity analyses were in line with the main finding. We found strong indications for publication bias. Egger's test was significant (intercept: 1.34; 95% CI [0.24, 2.43]; $p = 0.024$). Duvall and Tweedie's trim and fill procedure imputed seven missing studies. The adjusted average effect size declined to $g = 0.15$ (95% CI [-0.01, 0.31], NNT = 11.90), which was not statistically significant anymore ($p = 0.066$); p curve's right-skewness test also indicated that evidential value was not present ($p_{Full} = 0.052$, $p_{Half} = 0.572$, $k = 9$), but the existence of a small effect could not be rejected ($p_{Full} = 0.420$, $p_{Half} = 0.990$, $p_{Binomial} = 0.957$); p curve's effect size estimate was $d = 0.18$ when an outlier was removed.

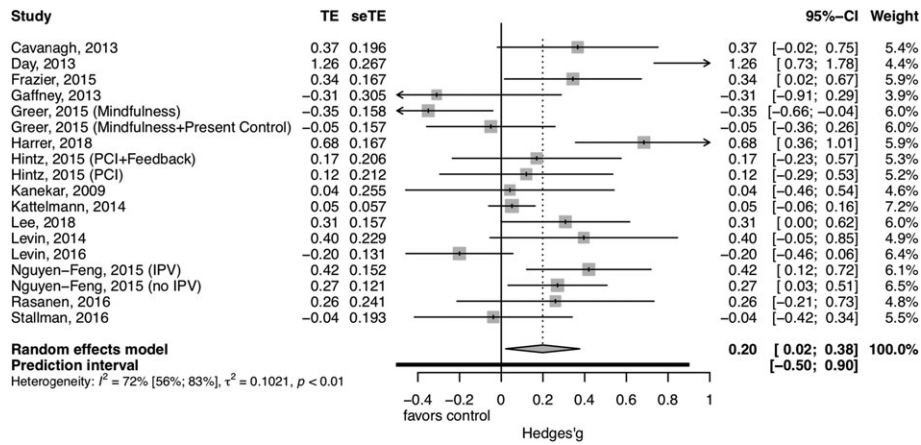


FIGURE 5 Forest plot for stress outcomes

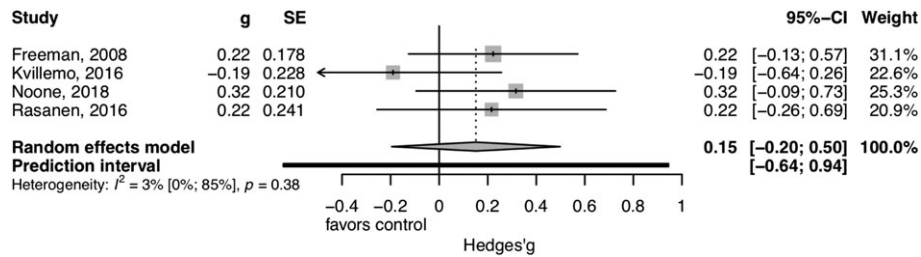


FIGURE 6 Forest plot for well-being outcomes

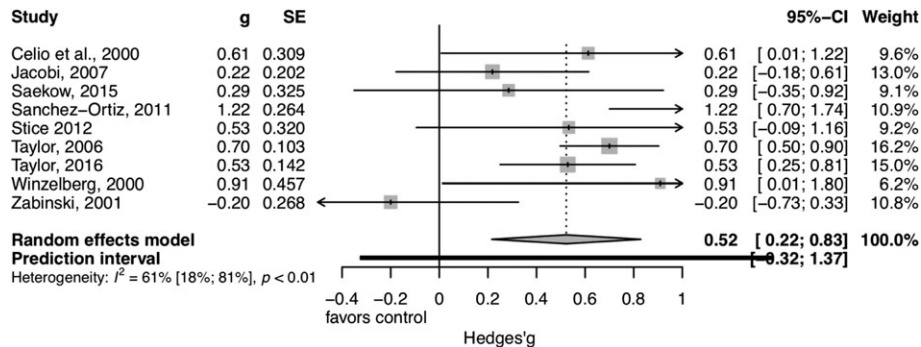


FIGURE 7 Forest plot for eating disorder symptom outcomes

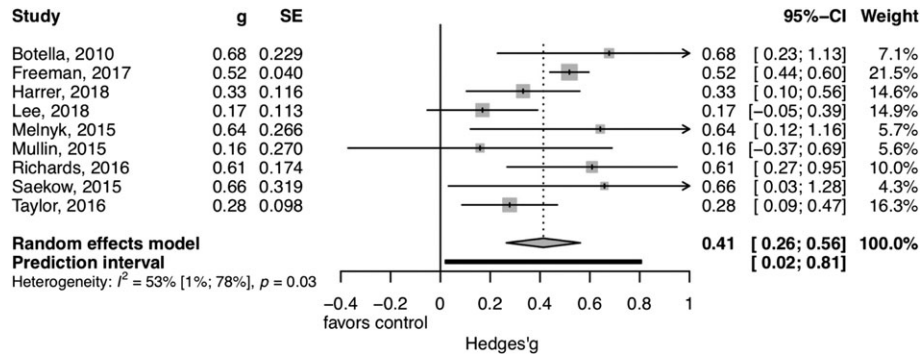


FIGURE 8 Forest plot for functioning outcomes

3.3.3 | Stress

The overall effect of 18 comparisons on stress was $g = 0.20$ (95% CI [0.02, 0.38]). This equals an NNT of 8.93. Heterogeneity was high ($I^2 = 72\%$; 95% CI [56–83]). The prediction interval for future trials ranged from $g = -0.50$ to 0.90. The pooled effect estimate

was slightly higher but not significant when only low risk of bias studies were included ($g = 0.30$; 95% CI: [-0.05, 0.66], NNT = 5.95; nine comparisons). In all the other sensitivity analyses, results in line with the main finding emerged. We found no indications of publication bias. The p curve indicated the presence of evidential value.

TABLE 3 Subgroup analyses

Subgroup	Depression						Anxiety						Stress					
	Effect size			Heterogeneity			Effect size			Heterogeneity			Effect size			Heterogeneity		
	<i>n_c</i>	<i>g</i>	95% CI	<i>I²</i>	95% CI	<i>p^a</i>	<i>n_c</i>	<i>g</i>	95% CI	<i>I²</i>	95% CI	<i>p^a</i>	<i>n_c</i>	<i>g</i>	95% CI	<i>I²</i>	95% CI	<i>p^a</i>
Technique						0.027						0.018						0.123
CBT	19	0.28	[0.15, 0.40]	45	[6, 68]		21	0.36	[0.23, 0.50]	43	[5, 66]		9	0.33	[0.02, 0.65]	78	[58, 88]	
Skills training	7	0.04	[-0.23, 0.30]	63	[16, 84]		5	-0.06	[-0.46, 0.35]	59	[0, 85]		9	0.08	[-0.12, 0.28]	60	[16, 81]	
Other	5	0.10	[-0.01, 0.21]	0	[0, 15]		1	0.03	[-0.43, 0.50]	-	-							
Guidance						0.651						0.825						0.865
Feedback	9	0.28	[-0.02, 0.57]	76	[54, 87]		11	0.27	[0.02, 0.52]	62	[27, 80]		7	0.26	[-0.23, 0.74]	85	[71, 92]	
Reminder	6	0.14	[-0.13, 0.41]	33	[0, 73]		5	0.37	[-0.07, 0.80]	67	[17, 88]		2	0.07	[-3.70, 3.83]	80	-	
None	16	0.15	[0.06, 0.25]	0	[0, 48]		11	0.25	[0.02, 0.49]	40	[0, 70]		9	0.21	[0.06, 0.35]	39	[0, 72]	
Treatment length						0.027						0.544						0.101
≤4 weeks	18	0.09	[-0.02, 0.21]	28	[0, 60]		14	0.21	[-0.03, 0.46]	62	[31, 79]		10	0.10	[-0.02, 0.21]	28	[0, 60]	
4-8 weeks	12	0.31	[0.13, 0.49]	59	[23, 78]		11	0.32	[0.16, 0.47]	31	[0, 66]		7	0.30	[0.13, 0.48]	55	[14, 76]	
≥8 weeks	1	0.13	[-0.43, 0.69]	-	-		2	0.54	[-3.40, 4.47]	62	-		1	0.13	[-0.43, 0.69]	-	-	
Recruitment						0.003						0.924						<0.001
Online	6	0.30	[0.25, 0.57]	57	[0, 83]		8	0.30	[0.09, 0.52]	43	[0, 75]		4	0.63	[-0.05, 1.31]	69	[10, 89]	
Mixed	2	0.62	[-1.33, 2.57]	0	-		2	0.25	[-1.95, 2.44]	11	-		1	-0.04	[-0.42, 0.34]	-	-	
Campus	8	0.14	[-0.01, 0.30]	0	[0, 68]		6	0.37	[0.06, 0.70]	12	[0, 78]		5	0.23	[0.03, 0.43]	57	[0, 84]	
Subject pool	7	0.04	[-0.10, 0.17]	0	[0, 57]		4	0.12	[-0.76, 0.99]	78	[40, 92]		2	-0.22	[-0.70, 0.27]	0	-	
Other/n.s.	8	0.20	[-0.09, 0.49]	64	[22, 83]		7	0.27	[-0.13, 0.66]	69	[32, 86]		6	0.04	[-0.21, 0.30]	53	[0, 81]	
Sample						0.026						0.054						0.110
Unselected	16	0.09	[-0.05, 0.23]	51	[12, 72]		16	0.18	[-0.01, 0.38]	57	[26, 76]		12	0.09	[-0.07, 0.26]	58	[20, 78]	
Preselected	15	0.29	[0.16, 0.42]	17	[0, 54]		11	0.42	[0.24, 0.59]	10	[0, 50]		6	0.41	[-0.06, 0.89]	78	[51, 90]	
Control group						<0.001						0.004						<0.001
Active	7	-0.06	[-0.23, 0.11]	50	[0, 79]		8	0.02	[-0.25, 0.29]	59	[9, 81]		5	-0.19	[-0.37, 0.01]	0	[0, 70]	
Passive	24	0.27	[0.17, 0.36]	23	[0, 54]		19	0.39	[0.25, 0.52]	17	[0, 52]		13	0.33	[0.15, 0.51]	68	[43, 82]	
Risk of bias						0.603						0.452						0.264
High	18	0.16	[0.04, 0.28]	37	[0, 64]		16	0.32	[0.12, 0.52]	60	[31, 77]		9	0.11	[-0.06, 0.28]	56	[6, 79]	
Low	13	0.21	[0.03, 0.40]	59	[24, 78]		11	0.22	[0.01, 0.44]	46	[0, 73]		9	0.30	[-0.05, 0.66]	80	[62, 89]	
Compensation						0.006						0.689						0.076
Yes	19	0.08	[-0.05, 0.20]	35	[0, 63]		12	0.25	[-0.06, 0.56]	68	[42, 83]		10	0.08	[-0.12, 0.27]	66	[33, 82]	
No	12	0.31	[0.18, 0.45]	42	[0, 71]		15	0.31	[0.18, 0.43]	23	[0, 58]		8	0.37	[0.04, 0.70]	69	[35, 85]	
Convenience sample						0.913						0.107						0.604
Yes	17	0.19	[0.06, 0.32]	17	[0, 53]		13	0.40	[0.14, 0.65]	53	[11, 75]		4	0.13	[-0.33, 0.59]	44	[0, 81]	
No	14	0.18	[0.02, 0.34]	65	[38, 80]		14	0.18	[0.03, 0.33]	45	[0, 71]		14	0.22	[0.01, 0.44]	77	[61, 86]	

Note. *n_c*: number of comparisons; n.s.: not specified.

^aThe *p* values in this column indicate whether the differences between the effect sizes in the subgroups are significant.

3.3.4 | Well-being

A pooled effect of $g = 0.15$ (95% CI [-0.20, 0.50]) was found for the effect of Internet interventions on well-being (four comparisons). This corresponds with an NNT of 11.90, but the effect was not statistically significant ($p = 0.259$). Heterogeneity was low ($I^2 = 3\%$; 95% CI [0–85]). The prediction interval ranged from $g = -0.64$ to 0.94. A significant effect ($g = 0.25$; 95% CI [0.11, 0.39], NNT = 7.14; $p = 0.016$) was found when leaving out one study in the influence analysis, and the between-study heterogeneity remained low ($I^2 = 0\%$; 95% CI [0–0]). Results in line with the main finding emerged when only studies with a low risk of bias rating were analysed. We found no indications of publication bias. The p curve indicated the presence of a true effect.

3.3.5 | Eating disorder symptoms

A total of nine comparisons on symptoms of disordered eating were analysed. The pooled effect was $g = 0.52$ (95% CI [0.22, 0.83]). This equals an NNT of 3.50. Heterogeneity was moderate ($I^2 = 61\%$; 95% CI [18–81]). The prediction interval ranged from $g = -0.32$ to 1.37. A slightly higher effect of $g = 0.61$ (95% CI [0.35, 0.68], NNT = 2.99) was found when leaving out one study in the influence analysis, and heterogeneity became non-significant ($I^2 = 39\%$; 95% CI [0–73], $p = 0.120$). We found results in line with the main finding when only studies with a low risk of bias rating were included. No indications of publication bias were found. The p curve analysis indicated the presence of evidential value.

3.3.6 | Sleep

Only two studies evaluated the effect of Internet interventions for insomnia relative to controls. These studies were not pooled in meta-analysis. Both interventions had a low risk of bias rating. The calculated effects on sleep were $g = 0.73$ (95% CI [0.63, 0.82], NNT = 2.54; Freeman et al., 2017) and $g = 0.50$ (95% CI [0.05, 0.94], NNT = 3.62; Morris et al., 2016).

3.3.7 | Functioning

The pooled effect of nine comparisons for functioning outcomes was $g = 0.41$ (95% CI [0.26, 0.56]), which corresponds with an NNT of 4.39. Heterogeneity was moderate ($I^2 = 53\%$; 95% CI [1–78]). The prediction interval only included positive values and ranged from $g = 0.02$ to 0.81. A similar effect emerged when leaving one study out in the influence analysis ($g = 0.45$; 95% CI [0.10, 0.81], NNT = 4), but heterogeneity became non-significant ($I^2 = 31\%$; 95% CI [0–69], $p = 0.180$). We also found a similar effect ($g = 0.41$; 95% CI [0.22, 0.60], NNT = 4.39) with non-significant between-study heterogeneity ($I^2 = 54\%$; 95% CI [0–83], $p = 0.07$) when we only pooled the effects of studies with a low risk of bias. We found no indications of publication bias. The p curve indicated the presence of evidential value.

3.4 | Subgroup analyses

Results of subgroup analyses for depression, anxiety, and stress are summarized in Table 3. We found several significant differences between

subgroups. For depression ($p = 0.026$), effects were higher in samples that were preselected through standardized cut-offs ($g = 0.29$, 95% CI [0.16, 0.21], NNT = 6.17) than in unselected samples ($g = 0.09$, 95% CI [-0.05, 0.23], NNT = 20). The pooled effect for interventions in unselected samples did not attain statistical significance ($p = 0.182$).

For all outcomes, effects were significantly higher (all $p < 0.01$) when interventions were compared with passive controls (depression: $g = 0.27$, 95% CI [0.17, 0.36], NNT = 6.58; anxiety: $g = 0.39$, 95% CI [0.25, 0.52], NNT = 4.59; stress: $g = 0.33$, 95% CI [0.15, 0.51], NNT = 5.43) than active control groups (depression: $g = -0.06$, 95% CI [-0.23, 0.11], NNH = 29.41; anxiety: $g = 0.02$, 95% CI [-0.25, 0.29], NNT = 83.33; stress: $g = -0.19$, 95% CI [-0.37, 0.01], NNH = 9.43). The pooled effects of interventions compared with active controls were not significant (all $p > 0.05$).

Intervention technique moderated effects on depression ($p = 0.027$) and anxiety ($p = 0.018$). For both target outcomes, effects were higher for interventions based on CBT principles (depression: $g = 0.28$, 95% CI [0.15, 0.40], NNT = 6.41; anxiety: $g = 0.36$, 95% CI [0.23, 0.50], NNT = 5). Effects were lower and non-significant (all $p > 0.05$) for skill trainings (depression: $g = 0.04$, 95% CI [-0.23, 0.30], NNT = 45.45; anxiety: $g = -0.06$, 95% CI [-0.46, 0.35], NNH = 29.41) and other techniques (depression: $g = 0.10$, 95% CI [-0.01, 0.21], NNT = 17.86; anxiety: $g = 0.03$, 95% CI [-0.43, 0.50], NNT = 62.5).

For depression, effects were highest for interventions between 4 and 8 weeks in length ($g = 0.31$, 95% CI [0.13, 0.49], NNT = 5.75) compared with shorter ($g = 0.09$, 95% CI [-0.02, 0.21], NNT = 20) or longer ($g = 0.13$, 95% CI [-0.43, 0.69], NNT = 13.51) programs ($p = 0.027$). The pooled effect was not significant for shorter programs ($p = 0.099$).

For depression, compensation was also an effect moderator ($p = 0.006$). The effect was higher in studies in which no compensation was provided ($g = 0.31$, 95% CI [0.18, 0.45], NNT = 5.75) compared with studies that compensated participants ($g = 0.08$, 95% CI [-0.05, 0.20], NNT = 21.74). The effect size for studies with compensation was not significant ($p = 0.209$).

Lastly, type of recruitment was a significant effect moderator for depression and stress outcomes (both $p < 0.01$). Effects were lowest when participants were recruited through a study subject pool ($g = 0.04$, 95% CI [-0.10, 0.17], NNT = 45.45; stress: $g = -0.22$, 95% CI [-0.70, 0.27], NNH = 8.06). These effects were not significant (both $p > 0.05$). Effects were higher for web-based recruitment ($g = 0.30$, 95% CI [0.25, 0.57], NNT = 5.95; stress: $g = 0.63$, 95% CI [-0.05, 1.31], NNT = 2.91).

We found no indication that guidance, risk of bias, or employment of convenience samples were significantly related to effect size (all $p \geq 0.05$).

4 | CONCLUSIONS

In this meta-analysis on Internet interventions for mental health and well-being in university students, we found small effects on depression, anxiety, and stress symptoms, as well as moderate-sized effects on eating disorder symptoms and students' social and academic functioning. No significant effects were found for interventions targeting student's well-being. Heterogeneity of effect sizes was moderate to substantial for anxiety, stress, eating disorder, and functioning outcomes. The small effect on depression as well as the moderate effects on eating disorder symptoms and student functioning found in the main analysis also emerged when accounting for potential publication

bias and when only studies with a low risk of bias were included. In subgroup analyses, we found that effects were higher in samples that were preselected through symptom cut-offs or risk factors, as well as for interventions that were of medium length (4–8 weeks), based on CBT principles, and were compared with passive control groups. Higher effects were also found when participants were not given any compensation and were not recruited through a study subject pool.

The effects on depression ($g = 0.18$, 95% CI [0.08, 0.27]), anxiety ($g = 0.27$, 95% CI [0.13, 0.40]), and stress ($g = 0.20$, 95% CI [0.02, 0.38]) in this meta-analysis are much smaller than found for such interventions in other target groups (depression: SMD = 0.90, 95% CI [0.73, 1.04]; Königbauer, Letsch, Doebler, Ebert, & Baumeister, 2017; anxiety: SMD = 0.80, 95% CI [0.42, 1.19]; Olthuis, Watt, Bailey, Hayden, & Stewart, 2015; stress: SMD = 0.43, 95% CI [0.31, 0.54]; Heber et al., 2017). This might be explained due to differences in intervention or sample characteristics, such as baseline symptom severity. Two recent meta-analyses report much smaller effects for Internet interventions aiming to prevent depression in subclinical populations (SMD = 0.25–0.35; Sander, Rausch, & Baumeister, 2016; Deady et al., 2017). However, these effects are still considerably larger than the one we found in unselected samples ($g = 0.09$, 95% CI [−0.05, 0.23]). It is also possible that Internet interventions for depression, anxiety, and stress are less effective in university students than in other target groups. The estimated effect size for depression adjusted for publication bias ($d = 0.09$) is considerably lower than the minimally important difference of SMD = 0.24 reported for depression outcomes (Cuijpers, Turner, Koole, Van Dijke, & Smit, 2014). This questions the clinical usefulness of treating depressive symptoms in students using Internet-based approaches. For anxiety, controlling for potential publication bias lead to a non-significant overall effect. Nevertheless, effect heterogeneity was moderate to substantial in many analyses, which is also reflected by the broad prediction intervals. Predictions for future trials based on present evidence ranged from negative effects to moderate and even large positive effects.

Given that previous research has clearly documented the enormous potential of Internet-based interventions for other target groups and areas of application (Andersson & Titov, 2014; Ebert et al., 2018), more research is clearly needed into how Internet interventions should be designed and delivered to exploit these capacities in university students. Results from our subgroup analyses indicate that effects are higher for interventions of moderate length (1–2 months), which is in line with previous research (Heber et al., 2017; Richards & Richardson, 2012). Findings in this meta-analysis also suggest that CBT programs were superior to other types of interventions. Although previous research suggests that guided Internet interventions have higher effect sizes than unguided interventions (Baumeister, Reichler, Munzinger, & Lin, 2014; Cowpertwait & Clarke, 2013), we did not find that guidance significantly moderated intervention efficacy. Apart from guidance, interventions in this analysis also varied considerably in terms of their length, intensity, and rationale, which may have impeded us from detecting the benefits of adding guidance to an intervention. However, it is also possible that provision of guidance could play a less crucial role in university students, and other factors are more important.

It is also very much possible that some students are more likely than others to respond to Internet interventions due to a range of

prescriptive predictors of treatment response that remain to be determined. It is noteworthy, in the latter regard, that substantial evidence exists for heterogeneity of treatment effects of standard psychotherapies and medications for the treatment of common mental disorders, based on a wide range of patient characteristics (e.g., childhood experiences, personality, coping style, symptom profiles and comorbidity, exposure to chronic stressors, and access to supportive social networks; Cohen & DeRubeis, 2018; Kessler et al., 2017; Lutz, Zimmermann, Müller, Deisenhofer, & Rubel, 2017). One main aim of the WMH-ICS is to carry out comparable analyses with the Internet interventions implemented as part of the initiative. If the results are in any way comparable with those found for face-to-face psychotherapies, we might be able to find subgroups of students among whom the effect sizes of certain Internet interventions are much higher than those in the total population, as well as other students for whom Internet interventions are likely to have no positive effects. Our finding in the subgroup analyses that effects were larger when participants were preselected through symptom cut-offs or risk factors points at this direction. If so, we hope to develop reliable clinical decision support systems based on artificial intelligence methods (e.g., Luedtke & van der Laan, 2017) to help match students in need of treatment with optimal Internet interventions and to determine which students need referrals to other types of treatment.

This study has several limitations. About half of the included studies were determined to show a high risk of bias. Therefore, results should be interpreted with caution. Furthermore, as long-term effect data were only reported for a small proportion of the included studies, and follow-up periods varied considerably, we were not able to pool such outcomes. Heterogeneity was substantial in some analyses and remained moderate even after outliers were removed. We also found evidence that some analyses could be biased by selective reporting. Given the shortcomings of the trim and fill procedure and p curve described before, there is currently no adequate method to accurately estimate effect sizes in the presence of both substantial heterogeneity and publication bias (van Aert et al., 2016). Results of analyses in which both these criteria were met should therefore be interpreted with caution. Lastly, we used Cohen's criteria to assess the magnitude of effects in this meta-analysis. Although these guidelines are commonly used in psychological research, it should be noted that there are no iron-clad criteria to assess the importance of an effect (Durlak, 2009). Effect sizes should thus be interpreted within the context of previous research, which we presented before.

Despite these limitations, we conclude that Internet-based mental health interventions for university students can be a potentially effective mean for a range of conditions and can have a beneficial impact on university students' functioning. Nevertheless, more research is needed to determine which types of interventions best fit which students, and in which context, to optimize their effects and thus fully exploit the potential of Internet-based interventions in improving university students' mental health.

ACKNOWLEDGEMENTS

We would like to thank Dr. Helen Stallman, Dr. Laura A. Szalacha, Dr. Bernadette Mazurek Melnyk, and Dr. Chris Noone for providing us with original data used in this analysis.

CONFLICT OF INTEREST

D. D. E. reports to have received consultancy fees or served in the scientific advisory board from several companies such as Minddistrict, Sanofi, Lantern, Schön Kliniken, and German health insurance companies (BARMER and Techniker Krankenkasse). D. D. E. and M. B. are also stakeholders of the institute for health trainings online (GET.ON), which aims to implement scientific findings related to digital health interventions into routine care. H. B. reports to have received consultancy fees and fees for lectures or workshops from chambers of psychotherapists and training institutes for psychotherapists. In the past 3 years, R. C. K. received support for his epidemiological studies from Sanofi-Aventis, was a consultant for Johnson & Johnson Wellness and Prevention, Sage Pharmaceuticals, Shire, Takeda, and served on an advisory board for the Johnson & Johnson Services Inc., and Lake Nona Life Project. R. C. K. is a co-owner of DataStat, Inc., a market research firm that carries out healthcare research. All other authors report no potential conflict of interest.

ORCID

Mathias Harrer  <https://orcid.org/0000-0001-7016-2687>

Sophia H. Adam  <https://orcid.org/0000-0003-4824-8163>

Harald Baumeister  <https://orcid.org/0000-0002-2040-661X>

Pim Cuijpers  <https://orcid.org/0000-0001-5497-2743>

Eirini Karyotaki  <https://orcid.org/0000-0002-0071-2599>

Randy P. Auerbach  <https://orcid.org/0000-0003-2319-4744>

Ronald C. Kessler  <https://orcid.org/0000-0003-4831-2305>

Ronny Bruffaerts  <https://orcid.org/0000-0002-0330-3694>

Matthias Berking  <https://orcid.org/0000-0001-5903-4748>

David D. Ebert  <https://orcid.org/0000-0001-6820-0146>

REFERENCES

- Andersson, G., & Titov, N. (2014). Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry, 13*(1), 4–11. <https://doi.org/10.1002/wps.20083>
- Arnett, J. J. (2004). *Emerging adulthood: The winding road from the late teens through the twenties*. New York: Oxford University Press. Retrieved from <https://books.google.de/books?id=CrGq7n4830kC>
- Auerbach, R. P., Alonso, J., Axinn, W. G., Cuijpers, P., Ebert, D. D., Green, J. G., ... Bruffaerts, R. (2016). Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychological Medicine, 46*(14), 1–16. <https://doi.org/10.1017/S0033291716001665>
- Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., ... Hasling, P. (2018). WHO World Mental Health Surveys International College Student project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology, 127*, 623–638. <https://doi.org/10.1037/abn0000362>
- Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2014). The impact of guidance on Internet-based mental health interventions—A systematic review. *Internet Interventions, 1*, 205–215. <https://doi.org/10.1016/j.invent.2014.08.003>
- Blanco, C., Okuda, M., Wright, C., Hasin, D. S., Grant, B. F., Liu, S.-M., & Olfson, M. (2008). Mental health of college students and their non-college-attending peers. *Archives of General Psychiatry, 65*(12), 1429–1437. <https://doi.org/10.1001/archpsyc.65.12.1429>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470743386.refs>
- Borenstein, M., Higgins, J., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods, 8*(1), 5–18. <https://doi.org/10.1002/jrsm.1230>
- Boutron, I., Guttet, L., Estellat, C., Moher, D., Hróbjartsson, A., & Ravaud, P. (2007). Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Medicine, 4*(2), e61. <https://doi.org/10.1371/journal.pmed.0040061>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed., Vol. 2nd). Mahwah: Erlbaum Associates.
- Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment selection in depression. *Annual Review of Clinical Psychology, 14*, 209–236. <https://doi.org/10.1146/annurev-clinpsy-050817-084746>
- Conley, C. S., Durlak, J. A., Shapiro, J. B., Kirsch, A. C., & Zahniser, E. (2016). A meta-analysis of the impact of universal and indicated preventive technology-delivered interventions for higher education students. *Prevention Science, 17*(6), 659–678. <https://doi.org/10.1007/s11121-016-0662-3>
- Cowpervait, L., & Clarke, D. (2013). Effectiveness of web-based psychological interventions for depression: A meta-analysis. *International Journal of Mental Health and Addiction, 11*(2), 247–268. <https://doi.org/10.1007/s11469-012-9416-z>
- Cuijpers, P., Auerbach, R. P., Benjet, C., Bruffaerts, R., Ebert, D. D., Karyotaki, E., & Kessler, R. C. (2018). The WHO World Mental Health International College Student (WMH-ICS) initiative: An overview. Submitted for publication.
- Cuijpers, P., Turner, E. H., Koole, S. L., Van Dijke, A., & Smit, F. (2014). What is the threshold for a clinically relevant effect? The case of major depressive disorders. *Depression and Anxiety, 31*(5), 374–378. <https://doi.org/10.1002/da.22249>
- Davies, E. B., Morriss, R., & Glazebrook, C. (2014). Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: A systematic review and meta-analysis. *Journal of Medical Internet Research, 16*(5), 1–22. <https://doi.org/10.2196/jmir.3142>
- Deady, M., Choi, I., Calvo, R. A., Glozier, N., Christensen, H., & Harvey, S. B. (2017). eHealth interventions for the prevention of depression and anxiety in the general population: A systematic review and meta-analysis. *BMC Psychiatry, 17*(1), 310. <https://doi.org/10.1186/s12888-017-1473-1>
- Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology, 34*(9), 917–928. <https://doi.org/10.1093/jpepsy/jsp004>
- 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. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>
- Ebert, D. D., Cuijpers, P., Muñoz, R. F., & Baumeister, H. (2017). Internet and mobile-based interventions for the prevention of mental health disorders: A narrative review and recommendations for future research. *Frontiers Psychiatry, 8*, 166. <https://doi.org/10.3389/FPSYT.2017.00116>
- Ebert, D. D., Van Daele, T., Nordgreen, T., Karekla, M., Compare, A. T., Zarbo, C., ... Baumeister, H. (2018). Internet- and mobile-based psychological interventions: Applications, efficacy, and potential for improving mental health. *European Psychologist, 23*, 167–187. <https://doi.org/10.1027/1016-9040/a000318>
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109). Retrieved from <http://www.bmj.com/content/315/7109/629>
- Eisenberg, D., Golberstein, E., & Hunt, J. B. (2009). Mental health and academic success in college. *The BE Journal of Economic Analysis & Policy, 9*(1). <https://doi.org/10.2202/1935-1682.2191>
- Eisenberg, D., Hunt, J., & Speer, N. (2013). Mental health in American colleges and universities: Variation across student subgroups and across

- campuses. *The Journal of Nervous and Mental Disease*, 201(1), 60–67. <https://doi.org/10.1097/NMD.0b013e31827ab077>
- Evans, N. J., Forney, D. S., Guido, F. M., Patton, L. D., & Renn, K. A. (2009). *Student development in college: Theory, research, and practice*. New York: John Wiley & Sons.
- Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90(3), 891–904. <https://doi.org/10.1007/s11192-011-0494-7>
- Freeman, D., Sheaves, B., Goodwin, G. M., Yu, L.-M., Nickless, A., Harrison, P. J., ... Wadekar, V. (2017). The effects of improving sleep on mental health (OASIS): A randomised controlled trial with mediation analysis. *The Lancet Psychiatry*, 4(10), 749–758. [https://doi.org/10.1016/S2215-0366\(17\)30328-0](https://doi.org/10.1016/S2215-0366(17)30328-0)
- Furlan, A. D., Pennick, V., Bombardier, C., & van Tulder, M. (2009). 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine*, 34(18), 1929–1941. <https://doi.org/10.1097/BRS.0b013e3181b1c99f>
- Goldman-Mellor, S. J., Caspi, A., Harrington, H., Hogan, S., Nada-Raja, S., Poulton, R., & Moffitt, T. E. (2014). Suicide attempt in young people: A signal for long-term health care and social needs. *JAMA Psychiatry*, 71(2), 119–127. <https://doi.org/10.1001/jamapsychiatry.2013.2803>
- Heber, E., Ebert, D. D., Lehr, D., Cuijpers, P., Berking, M., Nobis, S., & Riper, H. (2017). The benefit of web- and computer-based interventions for stress: A systematic review and meta-analysis. *Journal of Medical Internet Research*, 19(2), e32. <https://doi.org/10.2196/jmir.5774>
- Hedges, L. V., & Olkin, I. (2014). *Statistical methods for meta-analysis*. Orlando: Academic Press. <https://doi.org/https://www.elsevier.com/books/statistical-methods-for-meta-analysis/hedges/978-0-08-057065-5>
- Higgins, J., & Green, S. (2011). Cochrane handbook for systematic reviews of interventions. Retrieved from https://books.google.de/books?hl=de&lr=&id=NKMg9sMM6GUC&oi=fnd&pg=PT13&dq=Higgins+JPT,+Green+S,+editors.+Cochrane+handbook+for+systematic+reviews+of+interventions.+Chichester:+John+Wiley+%26+Sons%3B+2008.&ots=LH_BOXAyF1&sig=-aPIHN55fHcU3ksvUcLQlluzXRo
- Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539–1558. <https://doi.org/10.1002/sim.1186>
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327, 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hysenbegasi, A., Hass, S. L., & Rowland, C. R. (2005). The impact of depression on the academic productivity of university students. *Journal of Mental Health Policy and Economics*, 8(3), 145–151.
- Ibrahim, A. K., Kelly, S. J., Adams, C. E., & Glazebrook, C. (2013). A systematic review of studies of depression prevalence in university students. *Journal of Psychiatric Research*, 47(3), 391–400. <https://doi.org/10.1016/j.jpsychires.2012.11.015>
- Int'Hout, J., Ioannidis, J. P. A., & Borm, G. F. (2014). The Hartung–Knapp–Sidik–Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian–Laird method. *BMC Medical Research Methodology*, 14(1), 25. <https://doi.org/10.1186/1471-2288-14-25>
- Ishii, T., Tachikawa, H., Shiratori, Y., Hori, T., Aiba, M., Kuga, K., & Arai, T. (2018). What kinds of factors affect the academic outcomes of university students with mental disorders? A retrospective study based on medical records. *Asian Journal of Psychiatry*, 32, 67–72. <https://doi.org/10.1016/j.ajp.2017.11.017>
- Kessler, R. C., Foster, C. L., Saunders, W. B., & Stang, P. E. (1995). Social consequences of psychiatric disorders. I: Educational attainment. *The American Journal of Psychiatry*, 152(7), 1026–1032. <https://doi.org/10.1176/ajp.152.7.1026>
- Kessler, R. C., Van Loo, H. M., Wardenaar, K. J., Bossarte, R. M., Brenner, L. A., Ebert, D. D., ... Sampson, N. A. (2017). Using patient self-reports to study heterogeneity of treatment effects in major depressive disorder. *Epidemiology and Psychiatric Sciences*, 26(1), 22–36. <https://doi.org/10.1017/S2045796016000020>
- Königbauer, J., Letsch, J., Doeblner, P., Ebert, D., & Baumeister, H. (2017). Internet- and mobile-based depression interventions for people with diagnosed depression: A systematic review and meta-analysis. *Journal of Affective Disorders*, 223, 28–40. <https://doi.org/10.1016/j.jad.2017.07.021>
- Kraemer, H. C., & Kupfer, D. J. (2006). Size of treatment effects and their importance to clinical research and practice. *Biological Psychiatry*, 59(11), 990–996. <https://doi.org/10.1016/j.biopsych.2005.09.014>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., ... Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Medicine*, 6(7), e1000100. <https://doi.org/10.1371/journal.pmed.1000100>
- Luedtke, A. R., & van der Laan, M. J. (2017). Evaluating the impact of treating the optimal subgroup. *Statistical Methods in Medical Research*, 26(4), 1630–1640. <https://doi.org/10.1177/0962280217708664>
- Lutz, W., Zimmermann, D., Müller, V. N. L. S., Deisenhofer, A.-K., & Rubel, J. A. (2017). Randomized controlled trial to evaluate the effects of personalized prediction and adaptation tools on treatment outcome in outpatient psychotherapy: Study protocol. *BMC Psychiatry*, 17(1), 306. <https://doi.org/10.1186/s12888-017-1464-2>
- Morris, J., Firkins, A., Millings, A., Mohr, C., Redford, P., & Rowe, A. (2016). Internet-delivered cognitive behavior therapy for anxiety and insomnia in a higher education context. *Anxiety, Stress & Coping: An International Journal*, 29(4), 415–431. <https://doi.org/10.1080/10615806.2015.1058924>
- Niederkröthaler, T., Tinghög, P., Alexanderson, K., Dahlin, M., Wang, M., Beckman, K., ... Mittendorfer-Rutz, E. (2014). Future risk of labour market marginalization in young suicide attempters—A population-based prospective cohort study. *International Journal of Epidemiology*, 43(5), 1520–1530. <https://doi.org/10.1093/ije/dyu155>
- Noone, C., & Hogan, M. (2016). A protocol for a randomised active-controlled trial to evaluate the effects of an online mindfulness intervention on executive control, critical thinking and key thinking dispositions in a university student sample. *BMC Psychology*, 4, 17. <https://doi.org/10.1186/s40359-016-0122-7>
- Olthuis, J. V., Watt, M. C., Bailey, K., Hayden, J. A., & Stewart, S. H. (2015). Therapist-supported Internet cognitive behavioural therapy for anxiety disorders in adults. *The Cochrane Database of Systematic Reviews*, 3, CD011565. <https://doi.org/10.1002/14651858.CD011565>
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61(10), 991–996. <https://doi.org/10.1016/j.jclinepi.2007.11.010>
- R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: Foundation for Statistical Computing.
- Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, 32(4), 329–342. <https://doi.org/10.1016/j.cpr.2012.02.004>
- Ricky, C., & O'Donnell Siobhan, M. N. (2017). Factors associated with delayed diagnosis of mood and/or anxiety disorders. *Health Promotion and Chronic Disease Prevention in Canada: Research, Policy and Practice*, 37(5), 137.
- Royal College of Psychiatrists. (2011). *Mental Health of Students in Higher Education College Report CR166*. London. Retrieved from <http://www.rcpsych.ac.uk/publications/collegereports/cr/cr166.aspx>
- Sander, L., Rausch, L., & Baumeister, H. (2016). Effectiveness of Internet-based interventions for the prevention of mental disorders: A systematic review and meta-analysis. *JMIR Mental Health*, 3(3), e38. <https://doi.org/10.2196/mental.6061>
- Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). *Meta-analysis with R*. Springer. <https://doi.org/10.1007/978-3-319-21416-0>

- Sidik, K., & Jonkman, J. N. (2002). A simple confidence interval for meta-analysis. *Statistics in Medicine*, 21(21), 3153–3159. <https://doi.org/10.1002/sim.1262>
- Simonsohn, U., Nelson, L., & Simmons, J. (2018). p-curve app 4.06. http://p-curve.com/app4/pcurve_app4.06.r
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014a). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, 143(2), 534–547. <https://doi.org/10.1037/a0033242>
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014b). P-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science*, 9(6), 666–681. <https://doi.org/10.1177/1745691614553988>
- van Aert, R. C. M., Wicherts, J. M., & van Assen, M. A. L. M. (2016). Conducting meta-analyses based on p values: Reservations and recommendations for applying p-uniform and p-curve. *Perspectives on Psychological Science*, 11(5), 713–729. <https://doi.org/10.1177/1745691616650874>
- Verger, P., Guagliardo, V., Gilbert, F., Rouillon, F., & Kovess-Masfety, V. (2009). Psychiatric disorders in students in six French universities: 12-month prevalence, comorbidity, impairment and help-seeking. *Social Psychiatry and Psychiatric Epidemiology*, 45(2), 189–199. <https://doi.org/10.1007/s00127-009-0055-z> Epub 2009 Apr 19
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36. <https://doi.org/10.18637/jss.v036.i03>
- Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. <https://doi.org/10.1002/jrsm.11>
- Wang, P. S., Berglund, P., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 603–613. <https://doi.org/10.1001/archpsyc.62.6.603>
- WMH-ICS (2018). The WHO World Mental Health International College Student initiative. <http://www.webcitation.org/6w0NBawLL>
- World Health Organization. (2008). The global burden of disease: 2004 update. Retrieved from http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/index.html
- Yager, Z., & O'Dea, J. A. (2008). Prevention programs for body image and eating disorders on university campuses: A review of large, controlled interventions. *Health Promotion International*, 23(2), 173–189. <https://doi.org/10.1093/heapro/dan004>

SUPPORTING INFORMATION

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

How to cite this article: Harrer M, Adam SH, Baumeister H, et al. Internet interventions for mental health in university students: A systematic review and meta-analysis. *Int J Methods Psychiatr Res*. 2019;28:e1759. <https://doi.org/10.1002/mpr.1759>