Received: 5 July 2021 | Revised: 28 December 2021 | Accepted: 5 January 2022

DOI: 10.1111/jpm.12818

ORIGINAL ARTICLE

WILEY

Effect of horticultural therapy on mental health: A meta-analysis of randomized controlled trials

Hung-Ming Tu

Department of Horticulture, National Chung Hsing University, Taichung, Taiwan

Correspondence
Hung-Ming Tu, Associate Professor, Department of Horticulture, National Chung Hsing University, Taichung, Taiwan.
Email: hmtu@dragon.nchu.edu.tw

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

Accessible summary

What is known on the subject?

- Mental health assessment is a common topic in horticultural therapy (HT).
- Some study findings were contrary to positive findings, although several studies have explored the effect of HT on mental health and reported positive results.
- There is weak evidence in the benefits of HT based on evidence-based medicine. Stronger evidence of the link between HT and mental health will help in policy making by health policy makers and governments.

What the paper adds to existing knowledge?

- The purpose of this meta-analysis of randomized controlled trials (RCTs) was to explore the effect of HT on mental health.
- This study is the first meta-analysis of RCTs to provide critical evidence that HT has a positive effect on mental health.
- The effect size of HT on mental health was medium (0.55). The included studies had no negative effects on mental health.

What are the implications for practice?

- HT should be considered a useful therapy to be integrated in healthcare settings by horticultural therapists to improve mental health.
- In the HT program, indoor and outdoor plant activities, plant-related arts, crafts and other activities were common interventions to improve specific mental health problems through at least eight sessions.

Abstract

Introduction: HT has been widely used to promote mental health. However, heterogeneity and sample size issues of randomized controlled trials made it challenging to illustrate effect sizes across the evidence.

Aim: The purpose of this meta-analysis was to explore the effect of HT on mental health.

Methods: We used the PRISMA framework. A keyword search of Web of Science, PubMed, ProQuest and Cochrane was performed. The inclusion criteria were HT with RCTs and mental health assessments. A random-effects model was used to perform the meta-analysis.

Results: A total of 1,056 records were searched, and 18 eligible studies were extracted. The included RCTs had no statistical heterogeneity and publication bias. The meta-analysis showed that the HT experimental groups had a significant and positive impact on mental health compared with the control groups (effect size = 0.55).

Discussion: HT should be considered for enhancing mental health. The included studies had no negative outcomes and the most common HT intervention was at least eight sessions. Therefore, HT should be considered to enhance mental health. However, the concealed allocation and blinding processes should be improved in future studies.

Implications for Practice: This study recommends that HT should be integrated into healthcare settings to improve mental health.

KEYWORDS
anxiety, depression, mental health promotion, meta-analysis, therapy

1 | INTRODUCTION

The inclusive definition of mental health is a dynamic state of internal equilibrium that enables individuals to use their abilities in harmony with the universal values of society (Galderisi et al., 2015). This is possible because mentally healthy people may experience emotions that enables them to respect and care for themselves, other living beings, relationships, the environment and value freedom (Galderisi et al., 2015). The important components of mental health include cognitive and social skills, emotional regulation, empathy, flexibility, social roles and a harmonious relationship between the body and mind (Galderisi et al., 2015). Depression is a common mental health problem that affects personal development (World Health Organization, 2020). Some diseases and ageing are usually accompanied by mental problems (Parpa et al., 2015) that are easily overlooked. HT is a non-pharmacological intervention strategy for promoting mental health and reducing depression (Makizako et al., 2020). When presented by a professional horticultural therapist, it involves a treatment program with horticultural activities intended to achieve specific goals (American Horticultural Therapy Association, 2020a). Furthermore, HT is widely used in rehabilitative, vocational and community settings, since horticultural plants and activities promote mental health and offer multiple benefits (American Horticultural Therapy Association, 2020b; Relf, 1992; Relf & Lohr, 2003). Therefore, several countries have started implementing HT to improve mental health in healthcare and community programs.

The mechanism linking HT and mental health is the belief that horticultural activity facilitates well-being and is based on the biophilic tendency among humans to be in touch with plants and natural environments (Chen et al., 2013). According to the attention restoration theory, HT program activities involve being in touch with plants and nature to move attention away from negative emotions and reduce feelings of depression and anxiety (Kenmochi et al., 2019). This type of leisure activity in a natural environment produces a feeling of distancing from daily life or the urban environment, which is critical in promoting mental health and reducing mental fatigue (von Lindern, 2017). HT improves decision-making and bestows a sense of perceived control and empowerment, which acts as a buffer against negative mental problems (Kenmochi et al., 2019). In addition, physical activities and social cohesion and connectedness are important mechanisms linking greenspace and mental health (Ng et al., 2021; Thompson Coon et al., 2011; Van den Berg et al., 2019). HT may benefit mental health due to its biophilic and physical nature, ability to restore attention, and promote social cohesion and connectedness.

Mental health assessment is a common topic in HT studies. Several studies have explored the effect of HT on mental health and reported positive results (Kenmochi et al., 2019; Zhu et al., 2016). However, some studies presented results to the contrary (Luk et al., 2011; Nicholas et al., 2019). Results are inconsistent because of the different tools used to measure mental health (Kam & Siu, 2010). Study results may also have been affected by factors such as study participants, report quality, suitable control group randomization and blinding methods. Although some HT studies provide high-quality evidence from RCTs, Kamioka et al. (2014) explained that poor methodological and reporting quality and study heterogeneity resulted in insufficient evidence in favour of HT. A systematic review by Nicholas et al. (2019) indicated that RCTs of HT lacked robustness of evidence and led to non-significant results.

Small sample sizes in control and experimental groups were common characteristics in HT studies, such as Luk et al. (2011). Unfortunately, such sample sizes may induce statistically non-significant treatment effects (Borenstein et al., 2011). Some meta-analyses integrated the sample numbers of several studies to solve the problems related to small sample size. This was done by using search strategies, data extraction, quality assessment and data analysis (Borenstein et al., 2011). Soga et al. (2017) performed a meta-analysis of non-RCTs and RCTs to examine the overall health benefits of gardening and HT. Still, past studies rarely used meta-analysis to explore the effects of RCTs of HT. For example, Kamioka et al. (2014) and Nicholas et al. (2019) did not use meta-analysis to explain the effect of RCTs of HT due to heterogeneity of HT studies. Therefore, meta-analysis of RCTs remains an important research gap in studies about HT, especially mental health. Recently, the RCTs on HT in mental health assessment have increased rapidly. Thus, selecting

RCTs, using study quality assessment, and including a large number of studies with small sample sizes, could enhance the robustness of evidence. Therefore, the purpose of this meta-analysis of RCTs was to explore the effects of HT on mental health. This study hypothesized that HT has a positive effect on mental health.

2 | METHODS

2.1 | Search strategy, inclusion criteria, data extraction and quality assessment

Our search strategy, inclusion criteria, data extraction and quality assessment were based on the preferred reporting items for systematic reviews and meta-analyses (PRISMA) (Moher et al., , 2010). Table A.1 presents the PRISMA checklist. Studies were identified, screened for eligibility, and included in accordance with the PRISMA guidelines (Moher et al., , 2010). We searched the following electronic databases: Web of Science, PubMed, ProQuest and Cochrane (The Cochrane Collaboration Central Register of Controlled Clinical Trials, and Cochrane Database of Systematic Reviews). Relevant reviews were used to search for further references through other sources, such as Google Scholar. The search was conducted using the keywords "horticultural therapy" AND "randomized controlled trial," "horticul*" AND "randomized controlled trial," and "therapeutic horticulture" AND "randomized controlled trial" during September 2021. This study did not use the keyword "mental health" because of the multiple components of mental health. The searched articles were assessed to match the definition of mental health after identification and screening. The publication dates of the references and language were unrestricted.

During screening, duplicate references were removed by checking for identical titles, authors and publication dates. The first author also screened the titles and abstracts of articles to identify only relevant studies on HT. This study used three inclusion criteria. First, the study design had to be an RCT. Second, the study assessed mental health based on the inclusive definition of mental health, which covers cognitive and social skills, emotional regulation, empathy, flexibility, social roles and a harmonious relationship between body and mind (Galderisi et al., 2015). Thus, studies with a non-RCT design, review articles, study protocols and those that did not assess mental health were excluded. Third, studies had to have involved a series of HT treatment programs aimed at specific goals that were based on the definition of HT. Among these programs, leisure gardening was excluded as it did not match the definition of HT. Our meta-analysis had two exclusion criteria. First, articles with unavailable full-text or abstracts were excluded. Second, if an article had insufficient information for meta-analysis, we contacted the corresponding author to request statistical information via email. When the authors did not reply or provide adequate information the relevant articles were excluded.

Thereafter, the first author and one independent reviewer experienced in conducting meta-analyses assessed the quality of the RCT design of each eligible study. First, the first author and the reviewer independently used a standardized form to check and extract the information of each eligible study. The form included the names of the authors and the journal, sample size and age of controls and the experimental group, sample characteristics, HT program (frequency, duration, and type), mental health measurement, country, language of article and quality assessment score (Table 1). Second, the Physiotherapy Evidence Database (PEDro) scale (Morton, 2009; Moseley et al., 2002) was used to assess RCT quality. One systematic review of HT also used the PEDro scale to assess the quality of RCTs (Nicholas et al., 2019). The PEDro scale consists of 11 qualitative items: eligibility, random and concealed allocation, baseline similarity, subject, therapist, and assessor blinding, more than 85% follow-up, intention-to-treat analysis, between-group statistical comparison, and point and variability measures (Morton, 2009; Moseley et al., 2002). The first item, eligibility, assessed the external validity of the trial and was not scored. The remaining 10 items assessed the internal or statistical validity of the trial and assigned a score of 0 to 10 based on yes (1) or no (0) answers to each item (Morton, 2009; Moseley et al., 2002). The mean PEDro Scale score was about 5.1 to 5.7 (Morton, 2009; Moseley et al., 2002; Physiotherapy Evidence Database, 2020). A score of more than 6 meant that the study was of moderate to high quality (Physiotherapy Evidence Database, 2020). The first author and reviewer independently evaluated the PEDro Scale scores; disagreements were resolved by consensus. We used Cohen's kappa coefficient to assess the inter-coder reliability to ensure accurate and consistent quality assessment.

3 | META-ANALYSIS

We used Comprehensive Meta-Analysis Version 3 (Borenstein et al., 2013) to perform the meta-analysis. We evaluated the Hedges' g effect size on mental health in the control and experimental groups. The mental health level increased and decreased in the experimental group relative to the control group when the Hedges' g effect size was greater or less than zero respectively. In each included study, four data formats were used to calculate the Hedges' g effect size: (1) mean and standard deviation in each group; (2) mean change and difference in standard deviation in each group; (3) F-value for difference between changes; and (4) mean change and p-value in each group.

In the effect model, we assumed that no common true effect size was present in the studies included, as the studies included had a variety of HT frequency, duration, and type, mental health measurements, sample age and diseases. The fixed-effect model assumes that all studies share and generalize the same effect size and population variance (Borenstein et al., 2010; Cheung et al., 2012). In contrast, the random-effects model assumes that the included studies are random samples from a larger population to generalize the grand effect size (Borenstein et al., 2010; Loh et al., 2020). Therefore, we selected the random-effects model to evaluate the overall Hedges' g effect size, lower limit, upper limit, Z-value and p-value.

TABLE 1 Characteristics of eligible studies

No	Study and journal name	Sample size	Age (year)	Sample characteristics	HT Frequency and duration
1	D'Andrea et al. (2007) J. Ther. Hortic.	19 Experimental 19 Control	84.2 ± 8.9 83.8 ± 7.5	Older with Alzheimer's disease	2 times a week (for 12 weeks)
2	Kam and Siu (2010) Hong Kong J. Occup. Ther.	10 Experimental 12 Control	45.3 ± 10.4 43.3 ± 11.7	Psychiatric illness	10 times in 2 weeks (for 2 weeks)
3	Tse (2010) J. Clin. Nurs.	26 Experimental 27 Control	85.2 ± 5.2 83.0 ± 7.9	Older	once a week (for 8 weeks)
4	Luk et al. (2011) Int. J. Geriatr. Psychiatry	7 Experimental 6 Control	84.9 ± 8.3#	Older with dementia	2 times a week (for 6 weeks)
5	Detweiler et al. (2015) Altern. Ther. Health Med.	13 Experimental 11 Control	46.4 ± 11.9#	Veteran	5 times a week (for 3 weeks)
6	Zhu et al. (2016) Shanghai Arch. Psychiatry	52 Experimental 52 Control	46.5 ± 9.0#	Schizophrenia	3 times a week (for 12 weeks)
7	Vujcic et al. (2017) Environ. Res.	16 Experimental 14 Control	45.6 ± 10.2	Psychiatric illness	3 times a week (for 4 weeks)
8	Cha and Lee (2018) J. Korea Acad. Industr. Coop. Soc.	24 Experimental 29 Control	31.8 ± 5.7 30.1 ± 3.6	White-collar workers	Once a week (for 8 weeks)
9	Kim and Park (2018) Complement. Ther. Med.	18 Experimental 18 Control	40–59#	Middle-aged women	2 times a week (for 6 weeks)
10	Lee et al. (2018) Complement. Ther. Med.	10 Experimental 10 Control	10.4 10.5	Maladjusted elementary school children	Once a week (for 9 weeks)
11	Mochizuki-Kawai et al. (2018) Front. Psychol.	16 Experimental 11 Control	43.8 ± 10.4 40.6 ± 12.3	Neurocognitive disorder	Three classes within eight days in each phase Two phases with two weeks intervals
12	Ng et al. (2018) Int. J. Environ. Res. Public Health	29 Experimental 30 Control	67.2 ± 4.5 67.0 ± 4.2	Older	Once a week (for 12 weeks) plus once a month (during 3 months)
13	Sia et al. (2018) J. Ther. Hortic.	29 Experimental 30 Control	67.2 ± 4.5 67.0 ± 4.2	Older	Once a week (for 15 weeks)
14	Chu et al. (2019) Geriatr. Nurs.	75 Experimental 75 Control	79.2 77.9	Older	Once a week (for 8 weeks)
15	Makizako et al. (2020) J. Clin. Med.	26 Experimental 28 Control	73.1 ± 5.6 73.0 ± 5.9	Older	20 times a week (for 1 year)
16	Kim et al. (2020) J. People Plants Environ.	10 Experimental 9 Control	58.8 ± 9.6 61.3 ± 12.4	Caregiver	2 times a week (for 4 weeks)
17	Siu et al. (2020) Int. J. Environ. Res. Public Health	41 Experimental 41 Control	50.8 ± 10.5 49.7 ± 8.7	Mental Illness	Once a week (for 8 weeks)
18	Palsdottir et al. (2020) J. Rehabil. Med.	51 Experimental 50 Control	67 (47–79) 66 (48–80)	Stroke	2 times a week (for 10 weeks)
19	Yang et al. (2021) Aging Ment. Health	16 Experimental 16 Control	84.5 ± 6.0 85.0 ± 11.0	Older with dementia	Once a week (for 10 weeks)

#The statistic in all groups.

*Quality scores derived from Physiotherapy Evidence Database Scale (PEDro); AC, plant-related arts and craft activities; AES-I, Apathy Evaluation Scale-Informant Version; AS, Apathy Scale; BI, Barthel Index; BT, Block-Tapping; CES-D, Center for Epidemiologic Studies Depression Scale; CMAI, Cohen-Mansfield Agitation Inventory; C-SWEMWBS, Chinese version of the Short Warwick-Edinburgh Mental Well-Being Scale; DASS, Depression Anxiety Stress Scale; DS, Digit Span; GDS, Geriatric Depression Scale; HAD, Hospital Anxiety and Depression Scale; IP, indoor plant activity; LMS-WMS-R, Logical Memory Subtests of the Wechsler Memory Scale-Revised; LSI, Life Satisfaction Index; LSNS, Lubben Social Network Scale; MFS, Mental Fatigue Scale; MMSE, Mini-Mental State Examination; OP, outdoor plant activity; PANSS, Positive and Negative Syndrome Scale; PCL, Posttraumatic Stress Disorder Checklist; PR, other plant-related activities; PSS, Perceived Stress Scale; PWB, Ryff's Psychological Well-Being; PWI, Personal Wellbeing Index; QLESQ, Quality of Life Enjoyment and Satisfaction Questionnaire; QoL-AD, Quality of Life in Alzheimer's disease scale; SDS, Self-rating Depression Scale; SIS, Self-identity scale; STAI, State-Trait Anxiety Inventory; TSI, Test for Severe Impairment; UCLA-LS, UCLA Loneliness Scale; WMT, Wordlist Memory Tasks.

Type	Control group: Interventions	Measure	Country	Language of article	Quality Assessment*
IP, PR	Normal daily activities and regular therapeutic recreation programs (e.g. music programs, social hour)	TSI	USA	English	6/10
IP, OP, PR	Conventional work-related training	DASS, PWI	Hong Kong	English	7/10
IP	Conventional therapy	BI, LSI, UCLA-LS, LSNS	Hong Kong	English	5/10
OP	Sensory stimulation and social interaction (Craft activities)	CMAI	Hong Kong	English	7/10
OP	Occupational Therapy (Craft activities)	CES-D, PCL, QLESQ	US	English	5/10
IP, OP, PR	Conventional treatment (Medication)	PANSS	China	English	7/10
OP, PR	Conventional therapy (e.g. occupational and art therapy)	DASS	Serbia	English	6/10
IP, PR	No intervention	CES-D, Stress	South Korea	Korean	6/10
IP, AC	No description	SDS, SIS, STAI	South Korea	English	5/10
IP, AC	Activities (e.g. reading)	PSS	South Korea	English	4/10
AC	Daily activities (e.g. singing, cooking, light exercise)	AS, DS, BT	Germany	English	6/10
IP, PR	No description	PWB	Singapore	English	6/10
IP, OP, PR	No description	PWB	Singapore	English	6/10
IP, OP, PR	Daily activities (e.g. watching TV, listening to music, reading newspapers)	GDS, UCLA-LS	Taiwan	English	7/10
IP, OP, PR	Education classes (e.g. traffic safety or disaster prevention)	GDS, LMS-WMS-R, WMT	Japan	English	7/10
AC	No description	CES-D	South Korea	English	5/10
IP, AC, PR	Work-related skills training (e.g. crafts or manufacturing work)	DASS C-SWEMWBS	Hong Kong	English	6/10
OP, PR	Standard care (e.g. physiotherapy, occupational therapy, or speech therapy, mental health care, or stroke rehabilitation)	MFS, HAD	Sweden	English	6/10
IP, AC, PR	Usual activities (e.g. singing, calisthenics, and puzzle games)	AES-I, BI, MMSE, QoL-AD	China	English	6/10

Following Borenstein et al. (2011), Cooper (2015), and Tu and Chiu (2020), this study assessed heterogeneity, sensitivity analysis and publication bias in the included studies. We used the *Q* statistic test, I^2 statistic and τ^2 to examine the heterogeneity of the included studies (Borenstein et al., 2011). In sensitivity analysis, the cumulative meta-analysis of publication year involves a meta-analysis performed first with one study, then two, until all studies are included; it tests the statistical stability of the results (Borenstein et al., 2011). The study-removed meta-analysis was used to delete each study or a set of studies from the total pool repeatedly and observe the change in results (Borenstein et al., 2011). We repeated the process of removing one study at a time and performing the meta-analysis again. Thus, if the number of included studies is X, the analysis will have been performed X times. We also analysed the sensitivity of quality by deleting low-quality studies. We used six methods to test publication bias: (1) funnel plots; (2) Duval and Tweedie's trim-and-fill; (3) classic (Rosenthals) fail-safe *N*; (4) Orwin's fail-safe *N*; (5) Begg and Mazumdar's rank correlation; and (6) Egger's regression test (Begg & Mazumdar, 1994; Borenstein et al., 2011; Duval & Tweedie, 2000; Egger et al., 1997; Rosenthal, 1979). Lastly, funnel plots and one study-removed meta-analysis were used for outlier analysis to investigate the influence of outliers.

4 | RESULTS

4.1 | Study characteristics

A total of 1,056 records were searched in Web of Science, PubMed, ProQuest, Cochrane and other sources. We identified and removed 317 duplicates and obtained 739 studies (Figure 1). After excluding 679 studies on irrelevant topics, we obtained 60 studies on HT. Nineteen articles, including RCTs and mental health assessments, matched the inclusion eligibility criteria. Forty-one articles (26 non-RCTs; six studies lacking measurement for mental health; five abstracts; two articles with insufficient information for meta-analysis; two articles whose full text was unavailable) were excluded.

Data were extracted from 19 eligible studies, and their study quality was assessed (Table 1). Tables A.2 and A.3 present the original quality assessment scores for each item of the eligible studies from the first author and independent reviewer. The intercoder reliability of the quality assessment was good (Cohen's kappa coefficient, 0.87). Tables 1 and A.4 show that the quality scores of the 19 eligible articles ranged from 4 to 7, indicating that these trial reports were of moderate quality.

4.2 | Results of the meta-analysis

A funnel plot detected one study as an outlier (Figure a.1). This study (Chu et al., 2019) recorded larger means and small standard deviations to estimate the large effect size (Hedges' $g = 17.04$). The one study-removed meta-analysis showed that the Hedges' *g* effect size descended from 1.16 (95% CI = 0.59–1.74) to 0.55 (95% CI = 0.38–0.72) without this study (Figure a.2). Therefore, we performed an additional analysis without the study.

Table A.5 presents the original data of meta-analysed studies. The heterogeneity of the 18 studies included was not significant ($Q = 22.75$, df = 17, $p = .158$, $I^2 = 25.27\%$, $\tau^2 = 0.03$), indicating that the overall effect size of the studies included can be combined. Our meta-analysis of the final 18 studies included showed that HT experimental groups had a significant and positive impact on mental health compared with control groups (Hedges' $g = 0.55$, 95% CI = 0.38–0.72, $Z = 6.44$, $p < .001$) (Figure 2), which supports the hypothesis of this study.

Figure 3 presents the results of the quality assessment of the final 18 studies included. Half of the studies included did not fulfil four quality criteria (e.g. concealed allocation, subject, therapist and assessor blinding), indicating a risk of bias. However, several studies matched six quality criteria of eligibility, random allocation, similarity at baseline, intention-to-treat analysis, between-group statistical comparison, and point and variability measures.

4.3 | Sensitivity analyses and publication bias

In the sensitivity analysis, our cumulative meta-analysis showed that the Hedges' *g* effect size was stable when the publication year and number of samples were gradually increased (Figure 4a). No temporal bias was observed. In the one study-removed meta-analysis, Hedges' *g* effect size after removing any single study was stable and varied from 0.50 to 0.59 (Figure 4b), indicating that removing any one study did not affect the results. The Hedges' *g* effect size of studies with appropriate quality was 0.47 (95% CI = 0.29–0.64) by deleting the low-quality studies (lower than 6) (Figure 4c). The result was stable and significant, although the effect size of studies with appropriate quality was lower than that of the total studies.

The funnel plot (Figure 5) indicated that five studies with negative point estimates were missing from the meta-analysis. The imputed point estimate was 0.42 (95% CI = 0.24–0.60) as Duval and Tweedie's trim-and-fill method imputed five missing studies based on a random-effects model. Second, the classic fail-safe N showed that the result of this meta-analysis required 255 missing non-significant studies to change the significant *p*-value into a non-significant *p*-value ($Z = 7.62$, $p < .001$). Orwin's fail-safe N showed that the result of this meta-analysis required 945 studies with mean 0.00 Hedges' *g* to change the Hedges' *g* into the 0.01 Hedges' *g*. Neither Begg and Mazumdar's rank correlation (Kendall's $\tau = 0.27$; $p = .120$) nor Egger's regression test ($t = 1.40$, df = 16, $p = .182$) was significant. There was no indication that the study results were affected by the publication bias.

5 | DISCUSSION

This is the first meta-analysis of RCTs relating to the effects of HT on mental health. Our meta-analysis provides critical evidence that HT has a significantly positive effect on mental health. The effect

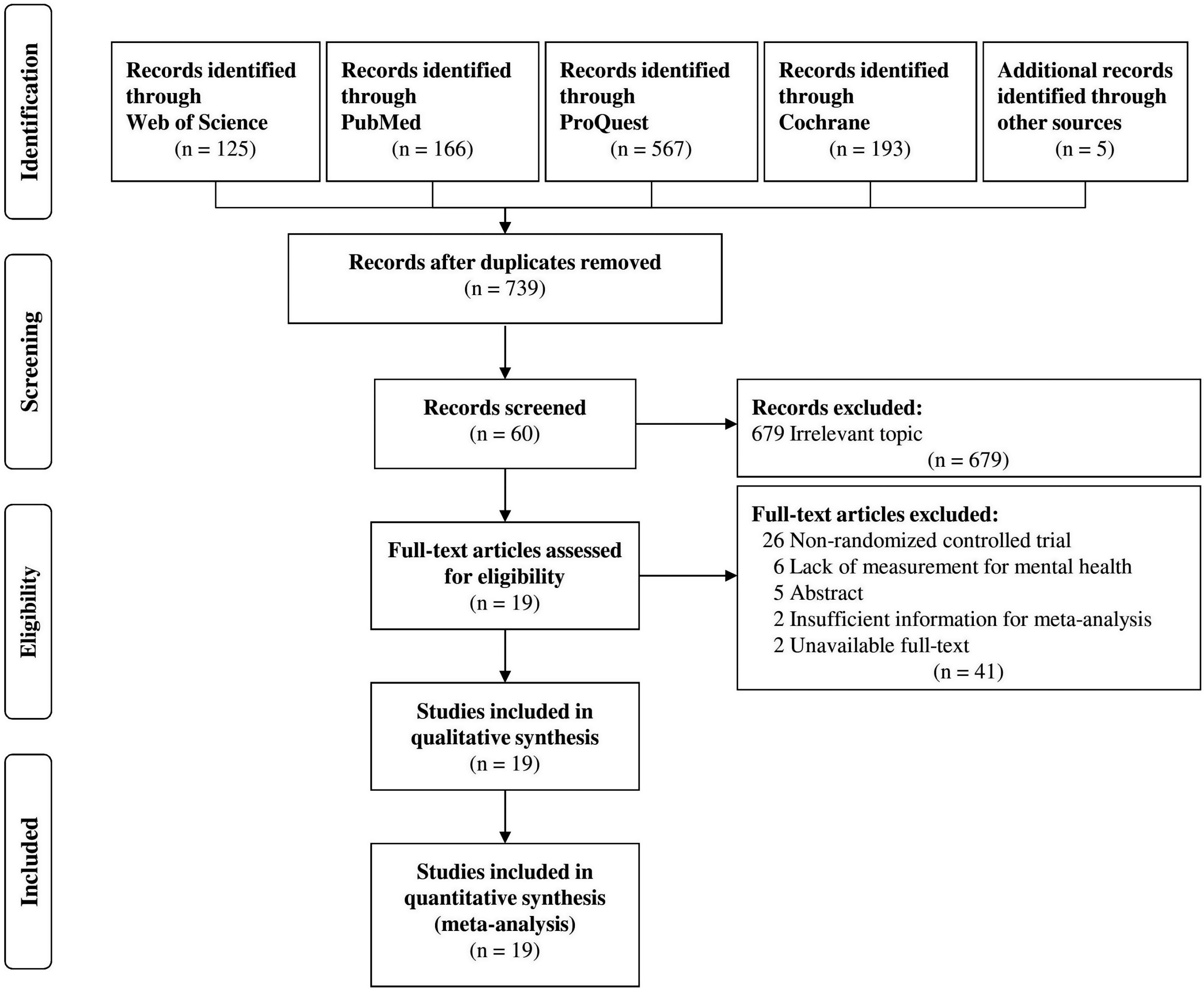


FIGURE 1 PRISMA flow diagram for study selection

size of HT on mental health was medium (0.55). Furthermore, our results support the mental health benefits of HT and may be of critical importance in practice and public mental health policies. HT should be considered useful for improving mental health and may be used in healthcare and community settings by a certified horticultural therapist. Not many countries use HT, although there are certified horticultural therapist professionals in America. This is because mainstream medicine practitioners feel that there is no strong evidence in favour of the beneficial effects of HT based on evidence-based medicine. However, our findings indicate that HT may be useful for mental health. We recommend that HT be integrated into medical, healthcare and community settings to improve mental health and that HT be formally certified.

The increase in RCTs of HT promotes the possibility of meta-analysis of RCTs. In 2014, Kamioka et al. (2014) discovered four RCTs of HT through a systematic review and did not implement a meta-analysis. Of these, two (Kam & Siu, 2010; Tse, 2010) was included in our meta-analysis, while the other two studies were excluded due to insufficient information required for meta-analysis and unavailable full-text respectively. Our study found that 18 RCTs of HT focused on mental health and were sufficient to implement a meta-analysis. Interestingly, the RCTs on HT increased rapidly from 2014 to 2021. Consequently, the total number of RCTs of HT was more than eighteen. Some of the RCTs focused on muscle strength (Yan et al., 2019) and biomarkers (Wong et al., 2021). Muscle strength and immune systems are the focus of future research directions in HT studies owing to scarce evidence.

The results of this study have several practical implications. First, the indoor and outdoor plant activities, plant-related arts, crafts and other activities were common interventions to improve specific mental health problems through at least eight consecutive sessions in the horticulture therapy program from the included studies. The number of sessions ranged from 8 to 36 except the study of Mochizuki-Kawai et al. (2018) (Table 1). Second, the program frequency and duration were determined for the specific purpose of the HT program because the included studies presented multiple

Study name	Hedges's g	Lower limit	Upper limit	Z-value	p-value
D'Andrea et al. (2007)	1.04	0.38	1.71	3.07	.002
Kam & Siu (2010)	0.92	0.03	1.81	2.04	.042
Tse (2010)	0.84	0.28	1.40	2.92	.003
Luk et al. (2011)	0.73	-0.33	1.78	1.35	.176
Detweiler et al. (2015)	0.45	-0.41	1.31	1.02	.307
Zhu et al. (2016)	0.80	0.41	1.20	3.96	<.001
Vujcic et al. (2017)	0.53	-0.18	1.24	1.46	.144
Cha & Lee (2018)	0.27	-0.26	0.81	1.00	.320
Kim & Park (2018)	1.39	0.68	2.11	3.81	<.001
Lee et al. (2018)	0.83	-0.06	1.72	1.83	.068
Mochizuki-Kawai et al. (2018)	0.31	-0.44	1.06	0.82	.411
Ng et al. (2018)	0.81	0.29	1.34	3.04	.002
Sia et al. (2018)	0.37	-0.14	0.88	1.44	.151
Makizako et al. (2020)	0.07	-0.46	0.59	0.25	.805
Kim et al. (2020)	0.61	-0.27	1.49	1.35	.177
Siu et al. (2020)	0.22	-0.24	0.68	0.93	.351
Palsdottir et al. (2020)	0.16	-0.25	0.58	0.77	.442
Yang et al. (2021)	0.45	-0.24	1.13	1.28	.202
Overall (Random-effects)	0.55	0.38	0.72	6.44	<.001

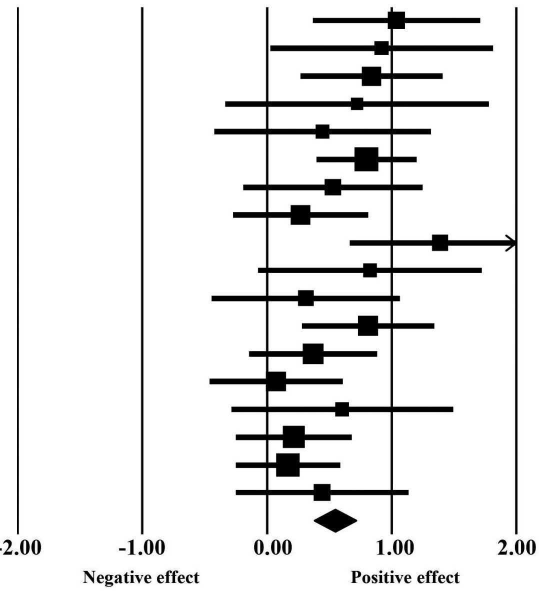


FIGURE 2 Forest plot: meta-analysis comparing the effect of HT on mental health

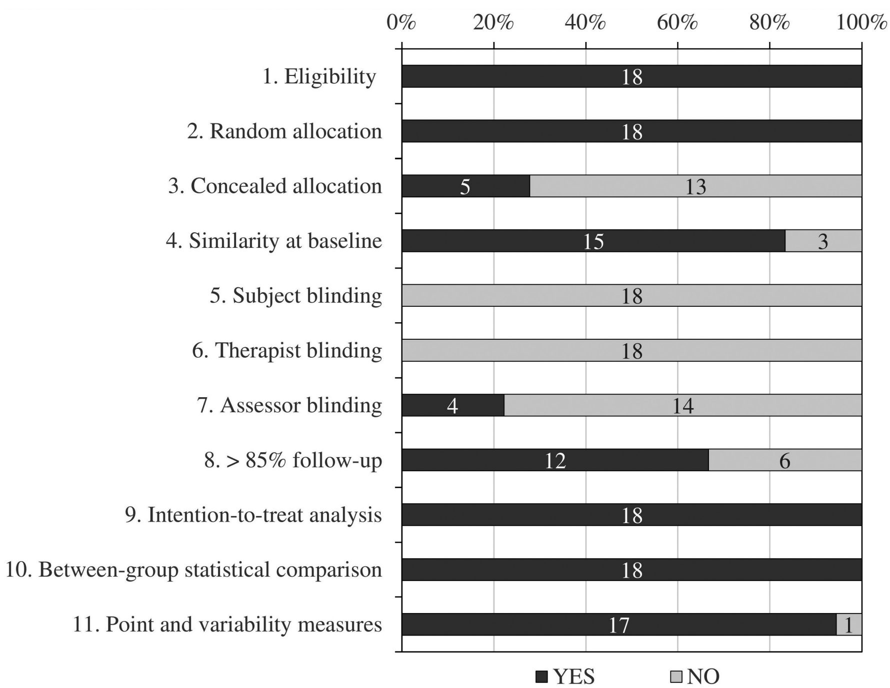


FIGURE 3 Quality assessment of the studies included in the meta-analysis using the PEDro Scale

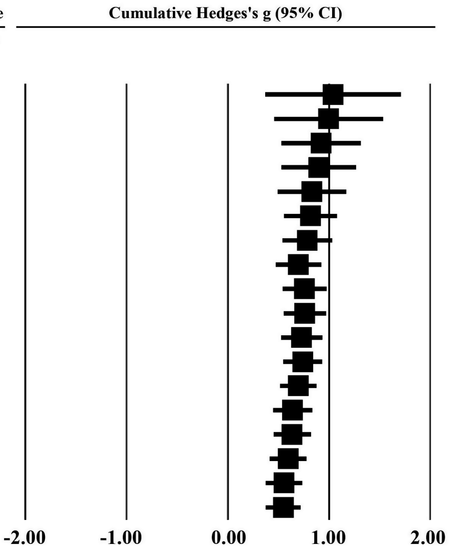
frequencies and durations. Third, older people and people with mental illness were the major subjects of HT for improving mental health problems. Finally, the interventions of the control group were conventional treatment or training, including work-related skills training, physiotherapy, occupational therapy, art therapy, speech therapy, mental health care, stroke rehabilitation, etc. Therefore, standard healthcare and treatment can combine HT to provide better healthcare.

6 | LIMITATIONS

Our study had some limitations. First, multiple sample characteristics, HT programs and mental health assessment measures may produce heterogeneity despite the lack of statistical heterogeneity in the included studies. This is because mental health is a common health problem with different sample characteristics. Different HT programs may produce similar mental health benefits. The different

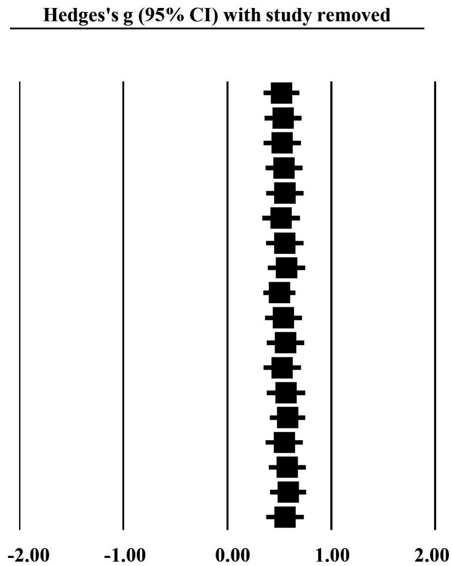
(a) Cumulative meta-analysis: publication year

Study name	Statistics for cumulative study					Cumulative sample size	
	Hedges's g	Lower limit	Upper limit	Z-value	p-value	Control group	Experimental group
D'Andrea et al. (2007)	1.04	0.38	1.71	3.07	.002	19	19
Kam & Siu (2010)	1.00	0.47	1.53	3.68	<.001	31	29
Tse (2010)	0.92	0.54	1.31	4.68	<.001	58	55
Luk et al. (2011)	0.90	0.54	1.26	4.86	<.001	64	62
Detweiler et al. (2015)	0.83	0.50	1.17	4.87	<.001	73	73
Zhu et al. (2016)	0.82	0.56	1.08	6.28	<.001	125	125
Vujcic et al. (2017)	0.79	0.55	1.03	6.41	<.001	139	141
Cha & Lee (2018)	0.70	0.48	0.92	6.25	<.001	168	165
Kim & Park (2018)	0.76	0.55	0.97	7.08	<.001	186	183
Lee et al. (2018)	0.76	0.56	0.97	7.32	<.001	196	193
Mochizuki-Kawai et al. (2018)	0.73	0.54	0.93	7.28	<.001	207	209
Ng et al. (2018)	0.74	0.56	0.93	7.89	<.001	237	238
Sia et al. (2018)	0.70	0.53	0.87	7.90	<.001	267	267
Makizako et al. (2020)	0.64	0.45	0.83	6.66	<.001	295	293
Kim et al. (2020)	0.64	0.46	0.82	7.04	<.001	304	303
Siu et al. (2020)	0.60	0.42	0.77	6.63	<.001	340	340
Palsdottir et al. (2020)	0.56	0.38	0.73	6.19	<.001	383	387
Yang et al. (2021)	0.55	0.38	0.72	6.44	<.001	399	403



(b) One study-removed meta-analysis

Study name	Statistics with study removed				
	Hedges's g	Lower limit	Upper limit	Z-value	p-value
D'Andrea et al. (2007)	0.52	0.36	0.69	6.14	<.001
Kam & Siu (2010)	0.54	0.37	0.71	6.15	<.001
Tse (2010)	0.53	0.36	0.70	5.99	<.001
Luk et al. (2011)	0.55	0.37	0.72	6.19	<.001
Detweiler et al. (2015)	0.56	0.38	0.73	6.23	<.001
Zhu et al. (2016)	0.52	0.35	0.70	5.83	<.001
Vujcic et al. (2017)	0.55	0.38	0.73	6.14	<.001
Cha & Lee (2018)	0.57	0.40	0.75	6.39	<.001
Kim & Park (2018)	0.50	0.35	0.65	6.62	<.001
Lee et al. (2018)	0.54	0.37	0.72	6.14	<.001
Mochizuki-Kawai et al. (2018)	0.56	0.39	0.74	6.31	<.001
Ng et al. (2018)	0.53	0.36	0.71	5.95	<.001
Sia et al. (2018)	0.57	0.39	0.74	6.22	<.001
Makizako et al. (2020)	0.58	0.42	0.75	6.93	<.001
Kim et al. (2020)	0.55	0.38	0.73	6.17	<.001
Siu et al. (2020)	0.58	0.41	0.75	6.58	<.001
Palsdottir et al. (2020)	0.59	0.42	0.76	6.91	<.001
Yang et al. (2021)	0.56	0.38	0.73	6.19	<.001



(c) Study-removed meta-analysis by deleting low-quality studies (lower than 6)

Study name	Statistics for each study				
	Hedges's g	Lower limit	Upper limit	Z-value	p-value
D'Andrea et al. (2007)	1.04	0.38	1.71	3.07	.002
Kam & Siu (2010)	0.92	0.03	1.81	2.04	.042
Luk et al. (2011)	0.73	-0.33	1.78	1.35	.176
Zhu et al. (2016)	0.80	0.41	1.20	3.96	<.001
Vujcic et al. (2017)	0.53	-0.18	1.24	1.46	.144
Cha & Lee (2018)	0.27	-0.26	0.81	1.00	.320
Mochizuki-Kawai et al. (2018)	0.31	-0.44	1.06	0.82	.411
Ng et al. (2018)	0.81	0.29	1.34	3.04	.002
Sia et al. (2018)	0.37	-0.14	0.88	1.44	.151
Makizako et al. (2020)	0.07	-0.46	0.59	0.25	.805
Siu et al. (2020)	0.22	-0.24	0.68	0.93	.351
Palsdottir et al. (2020)	0.16	-0.25	0.58	0.77	.442
Yang et al. (2021)	0.45	-0.24	1.13	1.28	.202
Overall (Random-effects)	0.47	0.29	0.64	5.23	<.001

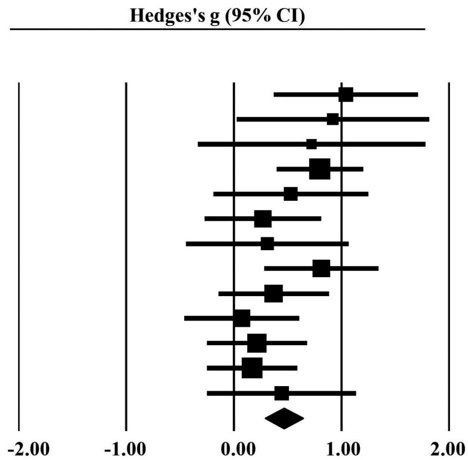


FIGURE 4 Forest plot: (a) cumulative meta-analysis; (b) one study-removed meta-analysis; and (c) study-removed meta-analysis by deleting low-quality studies (lower than 6)

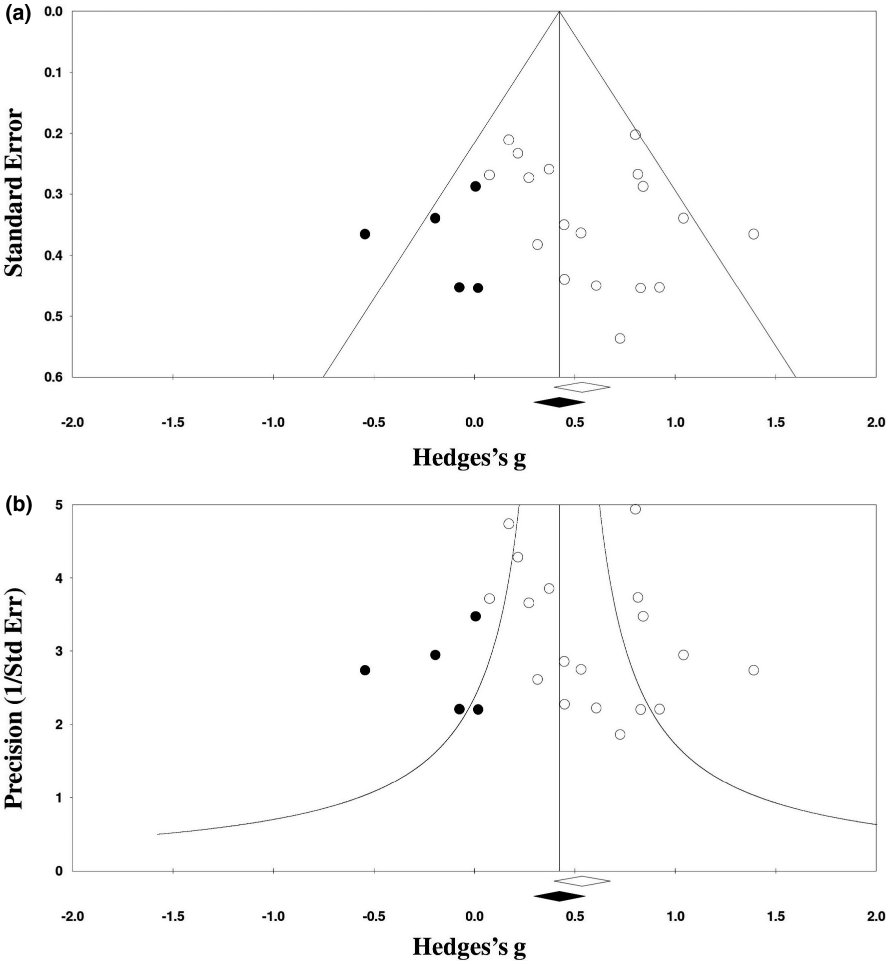


FIGURE 5 Evaluation of publication bias using funnel plots: (a) distribution of effect sizes (Hedges' g) and standard error; and (b) distribution of effect sizes (Hedges' *g*) and precision

assessment measures of the included studies correlated with mental health and probably produced no statistical heterogeneity. There should be more RCTs so that meaningful meta-analyses can be conducted. For example, a frequent question is the appropriate duration of HT programs. One study indicated that HT program duration may require more than four weeks to affect gardening behaviour due to familiarization with the therapists and horticultural activities (Jarrott et al., 2002). The duration of the HT programs included in our meta-analysis varied widely from two weeks to one year. Only Kam and Siu (2010) performed HT ten times in two weeks. The one study-removed meta-analysis showed a positive effect size without the study of Kam and Siu (2010). HT programs may require at least four weeks to produce benefits. Future studies could categorize studies according to the length of the HT program (less than or more than 4 weeks) to determine the appropriate duration of the intervention provided there are a sufficient number of RCTs. A meta-analysis of non-RCT studies may also be performed for practical purposes, given the paucity of RCTs on HT. The significant issues were sample, HT program and mental health assessment.

Second, older people or people suffering from mental illness were the main participants of the included studies. Only six of the included RCT studies explored the effect of HT on groups such as veterans, white-collar workers, middle-aged women, maladjusted elementary school children and stroke caregivers. Further RCT studies related to HT should explore groups such as children, people who suffer from physical disorders, are middle-aged, or free from mental health struggles.

Third, none of the studies included reached the highest score on the PEDro scale of ten; the highest score of the studies included was seven points. More high-quality RCTs are required to conduct a meaningful meta-analysis. Although the studies included were of moderate to high quality, they did not perform subject blinding or therapist blinding. Half of the studies failed to ensure concealed allocation and assessor blinding. Future HT studies should consider concealed allocation, assessor blinding, subject blinding and therapist blinding to improve study quality. Some studies have also indicated that concealment and blinding are lacking. Several quality assessment tools could be used to promote the quality of RCTs even when the studies lacked information on the concealment and blinding processes used. RCTs on HT should use appropriate quality scales, such as the PEDro scale, the Jadad scale (Jadad et al., 1996) or Cochrane Collaboration's tool (Higgins et al., 2011). Some HT studies have mentioned randomization, concealment, whether single-blind or double-blind, but have not explained in detail the processes involved and were therefore categorized as low-quality studies. These findings can be explained in detail. Two RCT protocols on HT clearly mention blinding, and other important steps (Chan et al., 2017; Makizako et al., 2015) can be referenced.

Sequentially numbered opaque sealed envelopes (SNOSE) is a common method for randomly assigning participants to control and experimental groups in allocation concealment (Clark et al., 2016). Some HT studies have implemented SNOSE (e.g. Palsdottir et al., 2020). The process involves a rigorous methodology and article publication interpretation process that includes an explanation of security measures, and who prepared and opened the envelopes (Clark et al., 2016).

In blinding, some HT studies have conducted blinding for assessors or research assistants (e.g. Kam & Siu, 2010; Luk et al., 2011; Makizako et al., 2015; Palsdottir et al., 2020; Zhu et al., 2016). The main problem regarding blinding is that the placebo effect for participants may taint the HT results. For example, participants' enhanced expectations of HT may result in positive outcomes (Cuijpers & Cristea, 2015). To exclude the placebo effect, Zhu et al. (2016) performed conventional treatment in both HT treatment and control groups. Vujcic et al. (2017) designed conventional therapy and art therapy without plants for the control group, thereby potentially reducing the placebo effect. We propose that both the HT treatment and control groups should take part in the same conventional therapy. For the control group, similar HT treatment activities without plants can be considered to reduce the potential placebo effect.

7 | CONCLUSIONS

Our meta-analysis provides critical evidence that HT has a significant and positive effect on mental health. The effect size of HT on mental health was moderate. The results were stable and unaffected by publication bias, publication year or the removal of any single study. HT should be considered in order to improve mental health and subsequently integrated into healthcare and community settings.

8 | RELEVANCE STATEMENT

There is sparse evidence regarding the benefits of HT based on evidence-based medicine. However, our results provide evidence of RCTs that showed the mental health benefits of HT. This suggests that HT should be considered as a therapy option in healthcare and community settings to improve mental health.

ACKNOWLEDGEMENTS

None.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTION

Hung-Ming Tu designed the study, searched and analysed the data, wrote, drafted and revised the manuscript.

ETHICS STATEMENTS

None.

DATA AVAILABILITY STATEMENT

Data available in article supplementary material

ORCID

Hung-Ming Tu https://orcid.org/0000-0002-0436-0789

REFERENCES

American Horticultural Therapy Association (2020). *Definitions and positions*. Retrieved June 24, 2020 from https://www.ahta.org/ahta-definitions-and-positions

American Horticultural Therapy Association (2020). *Horticultural therapy*. Retrieved June 24, 2020 from https://www.ahta.org/what-is-horticultural-therapy

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*, 1088–1101. https://doi.org/10.2307/2533446

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*, 97–111. https://doi.org/10.1002/jrsm.12

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. , NJ: Wiley.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2013). *Comprehensive Meta-Analysis Version 3*. , NJ: Biostat.

Cha, H. S., & Lee, S. (2018). The Effects of Horticulture Activity Caring Program on Stress, Depression, Blood Glucose, Lipids for Major Company White-Collar Workers.: A Randomized Controlled Trial. *Journal of the Korea Academia-Industrial Cooperation Society*, *19*, 231–242. https://doi.org/10.5762/KAIS.2018.19.10.231

Chan, H. Y., Ho, R. C. M., Mahendran, R., Ng, K. S., Tam, W. W. S., Rawtaer, I., Tan, C. H., Larbi, A., Feng, L., Sia, A., Ng, M. K. W., Gan, G. L., & Kua, E. H. (2017). Effects of horticultural therapy on elderly' health: protocol of a randomized controlled trial. *BMC Geriatrics*, *17*, 192. https://doi.org/10.1186/s12877-017-0588-z

Chen, H. M., Tu, H. M., & Ho, C. I. (2013). Understanding biophilia leisure as facilitating well-being and the environment: An examination of participants' attitudes toward horticultural activity. *Leisure Sciences*, *35*(4), 301–319. https://doi.org/10.1080/01490400.2013.797323

Cheung, M. W. L., Ho, R. C., Lim, Y., & Mak, A. (2012). Conducting a meta-analysis: basics and good practices. *International Journal of Rheumatic Diseases*, *15*, 129–135. https://doi.org/10.1111/j.1756-185X.2012.01712.x

Chu, H. Y., Chen, M. F., Tsai, C. C., Chan, H. S., & Wu, T. L. (2019). Efficacy of a horticultural activity program for reducing depression and loneliness in older residents of nursing homes in Taiwan. *Geriatric Nursing*, *40*(4), 386–391. https://doi.org/10.1016/j.gerinurse.2018.12.012

Clark, L., Fairhurst, C., & Torgerson, D. J. (2016). Allocation concealment in randomised controlled trials: are we getting better? *BMJ*, *355*, https://doi.org/10.1136/bmj.i5663

Cooper, H. (2015). *Research synthesis and meta-analysis: A step-by-step approach*. Sage.

Cuijpers, P., & Cristea, I. A. (2015). What if a placebo effect explained all the activity of depression treatments? *World Psychiatry*, *14*, 310–311. https://doi.org/10.1002/wps.20249

D'Andrea, S. J., Batavia, M., & Sasson, N. (2007). Effect of Horticultural Therapy on Preventing the Decline of Mental Abilities of Patients with Alzheimer's Type Dementia. *Journal of Therapeutic Horticulture*, *18*, 8–17.

Detweiler, M. B., Lane, S., Spencer, L., Lutgens, B., Halling, M. H., Rudder, T. F. et al (2015). Horticultural therapy: a pilot study on modulating cortisol levels and indices of substance craving, posttraumatic stress disorder, depression, and quality of life in veterans. *Alternative Therapies in Health and Medicine*, *21*, 36–41.

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*, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629

Galderisi, S., Heinz, A., Kastrup, M., Beezhold, J., & Sartorius, N. (2015). Toward a new definition of mental health. *World Psychiatry*, *14*(2), 231–233. https://doi.org/10.1002/wps.20231

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savovic, J., Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, *343*, https://doi.org/10.1136/bmj.d5928. d5928.

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials*, *17*, 1–12. https://doi.org/10.1016/0197-2456(95)00134-4

Jarrott, S. E., Kwack, H. R., & Relf, D. (2002). An observational assessment of a dementia-specific horticultural therapy program. *Horttechnology*, *12*, 403–410. https://doi.org/10.21273/HORTTECH.12.3.403

Kam, M. C. Y., & Siu, A. M. H. (2010). Evaluation of a horticultural activity programme for persons with psychiatric illness. *Hong Kong Journal of Occupational Therapy*, *20*(2), 80–86. https://doi.org/10.1016/S1569-18611170007-9

Kamioka, H., Tsutani, K., Yamada, M., Park, H., Okuizumi, H., Honda, T., Okada, S., Park, S. J., Kitayuguchi, J., Abe, T., Handa, S., & Mutoh, Y. (2014). Effectiveness of horticultural therapy: a systematic review of randomized controlled trials. *Complementary Therapies in Medicine*, *22*, 930–943. https://doi.org/10.1016/j.ctim.2014.08.009

Kenmochi, T., Kenmochi, A., & Hoshiyama, M. (2019). Effects of horticultural therapy on symptoms and future perspective of patients with schizophrenia in the chronic stage. *Journal of Therapeutic Horticulture*, *29*, 1–10.

Kim, K. H., & Park, S. A. (2018). Horticultural therapy program for middle-aged women's depression, anxiety, and self-identify. *Complementary Therapies in Medicine*, *39*, 154–159. https://doi.org/10.1016/j.ctim.2018.06.008

Kim, Y. H., Park, C. S., Bae, H. O., Lim, E. J., Kang, K. H., Lee, E. S., Jo, S. H., & Huh, M. R. (2020). Horticultural therapy programs enhancing quality of life and reducing depression and burden for caregivers of elderly with dementia. *Journal of People Plants Environment*, *23*, 305–320. https://doi.org/10.11628/ksppe.2020.23.3.305

Lee, M. J., Oh, W., Jang, J. S., & Lee, J. Y. (2018). A pilot study: Horticulture-related activities significantly reduce stress levels and salivary cortisol concentration of maladjusted elementary school children. *Complementary Therapies in Medicine*, *37*, 172–177. https://doi.org/10.1016/j.ctim.2018.01.004

Loh, A. Z. H., Tan, J. S., Tam, J. K. C., Zhang, M. W., Ho, C. S. H., & Ho, R. C. (2020). Postoperative psychological disorders among heart transplant recipients: A meta-analysis and meta-regression. *Psychosomatic Medicine*, *82*, 689–698. https://doi.org/10.1097/PSY.0000000000000833

Luk, K. Y., Lai, K. Y. C., Li, C. C., Cheung, W. H., Lam, S. M. R., Li, H. Y., Ng, K. P., Shiu, W. H., So, C. Y., & Wan, S. F. (2011). The effect of horticultural activities on agitation in nursing home residents with dementia. *International Journal of Geriatric Psychiatry*, *26*, 435–436. https://doi.org/10.1002/gps.2493

Makizako, H., Tsutsumimoto, K., Doi, T., Hotta, R., Nakakubo, S., Liu-Ambrose, T., Shimada, H. (2015). Effects of exercise and horticultural intervention on the brain and mental health in older adults with depressive symptoms and memory problems: study protocol for a randomized controlled trial [UMIN000018547]. *Trials*, *16*, 499. https://doi.org/10.1186/s13063-015-1032-3

Makizako, H., Tsutsumimoto, K., Doi, T. Makino, K., Nakakubo, S., Liu-Ambrose, T., & Shimada, H. (2020). Exercise and horticultural programs for older adults with depressive symptoms and memory problems: A randomized controlled trial. *Journal of Clinical Medicine*, *9*, 99. https://doi.org/10.3390/jcm9010099

Mochizuki-Kawai, H., Kotani, I., Mochizuki, S., & Yamakawa, Y. (2018). Structured floral arrangement program benefits in patients with neurocognitive disorder. *Frontiers in Psychology*, *9*, 1328. https://doi.org/10.3389/fpsyg.2018.01328

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery*, *8*, 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007

Morton, N. A. (2009). The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *Australian Journal of Physiotherapy*, *55*, 129–133. https://doi.org/10.1016/S0004-9514(09)70043-1

Moseley, A. M., Herbert, R. D., Sherrington, C., & Maher, C. G. (2002). Evidence for physiotherapy practice: a survey of the Physiotherapy Evidence Database (PEDro). *Australian Journal of Physiotherapy*, *48*, 43–49. https://doi.org/10.1016/S0004-9514(14)60281-6

Ng, K., Sia, A., Ng, M., Tan, C., Chan, H., Tan, C., Rawtaer, I., Feng, L., Mahendran, R., Larbi, A., Kua, E., & Ho, R. (2018). Effects of horticultural therapy on Asian older adults: A randomized controlled trial. *International Journal of Environmental Research and Public Health*, *15*, 1705. https://doi.org/10.3390/ijerph15081705

Ng, T. K. S., Gan, D. R. Y., Mahendran, R., Kua, E. H., & Ho, R. C. M. (2021). Social connectedness as a mediator for horticultural therapy's biological effect on community-dwelling older adults: Secondary analyses of a randomized controlled trial. *Social Science & Medicine*, *284*, https://doi.org/10.1016/j.socscimed.2021.114191. 114191.

Nicholas, S. O., Giang, A. T., & Yap, P. L. K. (2019). The effectiveness of horticultural therapy on older adults: a systematic review. *Journal of the American Medical Directors Association*, *20*, 1351.e1–1351.e11. https://doi.org/10.1016/j.jamda.2019.06.021

Palsdottir, A. M., Stigmar, K., Norrving, B., Petersson, I. F., Åström, M., & Pessah-Rasmussen, H. (2020). The nature stroke study; Nastru: A randomized controlled trial of nature-based post-stroke fatigue rehabilitation. *Journal of Rehabilitation Medicine*, *52*, 1–7. https://doi.org/10.2340/16501977-2652

Parpa, E., Tsilika, E., Gennimata, V., & Mystakidou, K. (2015). Elderly cancer patients' psychopathology: a systematic review: aging and mental health. *Archives of Gerontology and Geriatrics*, *60*, 9–15. https://doi.org/10.1016/j.archger.2014.09.008

Physiotherapy Evidence Database (2020). PEDro statistics. Retrieved June 20, 2020 from https://www.pedro.org.au/english/downloads/pedro-statistics/

Relf, D. (1992). Human issues in horticulture. *Horttechnology*, *2*, 159–171. https://doi.org/10.21273/HORTTECH.2.2.159

Relf, P. D., & Lohr, V. I. (2003). Human issues in horticulture. *HortScience*, *38*, 984–993. https://doi.org/10.21273/HORTSCI.38.5.984

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Sia, A., Ng, K. S. T., Ng, M. K., Chan, H. Y., Tan, C. H., Rawtaer, I. et al (2018). The effect of therapeutic horticulture on the psychological wellbeing of elderly in Singapore: A randomised controlled trial. *Journal of Therapeutic Horticulture*, *28*, 1–10.

Siu, A. M. H., Kam, M., & Mok, I. (2020). Horticultural therapy program for people with mental illness: A mixed-method evaluation.

International Journal of Environmental Research and Public Health, *17*, 711. https://doi.org/10.3390/ijerph17030711

Soga, M., Gaston, K. J., & Yamaura, Y. (2017). Gardening is beneficial for health: A meta-analysis. *Preventive Medicine Reports*, *5*, 92–99. https://doi.org/10.1016/j.pmedr.2016.11.007

Thompson Coon, J., Boddy, K., Stein, K., Whear, R., Barton, J., & Depledge, M. H. (2011). Does participating in physical activity in outdoor natural environments have a greater effect on physical and mental wellbeing than physical activity indoors? A systematic review. *Environmental Science & Technology*, *45*(5), 1761–1772. https://doi.org/10.1021/es102947t

Tse, M. M. Y. (2010). Therapeutic effects of an indoor gardening programme for older people living in nursing homes. *Journal of Clinical Nursing*, *19*(7–8), 949–958. https://doi.org/10.1111/j.1365-2702.2009.02803.x

Tu, H. M., & Chiu, P. Y. (2020). Meta-analysis of controlled trials testing horticultural therapy for the improvement of cognitive function. *Scientific Reports*, *10*(1), 14637. https://doi.org/10.1038/s41598-020-71621-7

Van den Berg, M. M., van Poppel, M., van Kamp, I., Ruijsbroek, A., Triguero-Mas, M., Gidlow, C. Nieuwenhuijsen, M. J., Gražulevičiene, R., van Mechelen, W., Kruize, H., & Maas, J. (2019). Do physical activity, social cohesion, and loneliness mediate the association between time spent visiting green space and mental health? *Environment & Behavior*, *51*(2), 144–166. https://doi.org/10.1177/0013916517738563

von Lindern, E. (2017). Perceived interdependencies between settings as constraints for self-reported restoration. *Journal of Environmental Psychology*, *49*, 8–17. https://doi.org/10.1016/j.jenvp.2016.11.004

Vujcic, M., Tomicevic-Dubljevic, J., Grbic, M., Lecic-Tosevski, D., Vukovic, O., & Toskovic, O. (2017). Nature based solution for improving mental health and well-being in urban areas. *Environmental Research*, *158*, 385–392. https://doi.org/10.1016/j.envres.2017.06.030

Wong, G. C. L., Ng, T. K. S., Lee, J. L., Lim, P. Y., Chua, S. K. J., Tan, C. Chua, M., Tan, J., Lee, S., Sia, A., Ng, M. K. W., Mahendran, R., Kua, E. H., Ho, R. C. M., & Larbi, A. (2021). Horticultural therapy reduces biomarkers of immunosenescence and inflammaging in community-dwelling older adults: A feasibility pilot randomized controlled trial. *The Journals of Gerontology: Series A*, *76*, 307–317. https://doi.org/10.1093/gerona/glaa271

World Health Organization (2020). Depression. Retrieved June 24, 2020 from https://www.who.int/health-topics/depression#tab=tab_1

Yan, W., Liu, L., Murong, J., Wang, Z., Cui, S., & Huang, W. (2019). Environmental Effect of Horticultural Therapy on Improving Muscle Strength after ACL Reconstruction. *Ekoloji*, *28*(107), 4289–4293.

Yang, Y., Kwan, R. Y., Zhai, H. M., Xiong, Y., Zhao, T., Fang, K. L., & Zhang, H. Q. (2021). Effect of horticultural therapy on apathy in nursing home residents with dementia: a pilot randomized controlled trial. *Aging & Mental Health*, *1–9*, https://doi.org/10.1080/13607863.2021.1907304

Zhu, S., Wan, H., Lu, Z., Wu, H., Zhang, Q., Qian, X., & Ye, C. (2016). Treatment effect of antipsychotics in combination with horticultural therapy on patients with schizophrenia: a randomized, double-blind, placebo-controlled study. *Shanghai Archives of Psychiatry*, *28*, 195–203. https://doi.org/10.11919/j.issn.1002-0829.216034

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Tu, H.-M. (2022). Effect of horticultural therapy on mental health: A meta-analysis of randomized controlled trials. *Journal of Psychiatric and Mental Health Nursing*, 29, 603–615. https://doi.org/10.1111/jpm.12818